

# **Associations between young learners' English language performance and teacher proficiency and experience with English**

**Laura Loder Buechel**

Dissertation zur Erlangung der Doktorwürde an der Philosophischen Fakultät der Universität Freiburg in der Schweiz.

*Genehmigt von der Philosophischen Fakultät auf Antrag der Professoren Raphael Berthele (1. Gutachter) und David Singleton (2. Gutachter).*

*Freiburg, den 16.12.2014.*

*Prof. Dr. Marc-Henry Soulet, Dekan.*

Cite as:

Loder Buechel, Laura (2015). Associations between young learners' English language performance and teacher proficiency and experience with English PhD thesis. University of Fribourg (Switzerland).

Ich erkläre ehrenwörtlich, dass ich meine Dissertation selbständig und ohne unzulässige fremde Hilfe verfasst habe und sie noch keiner anderen Fakultät vorgelegt habe.

## Acknowledgements

First and foremost, I'd like to offer up a hearty thank you to the teachers who participated in the pilot test and the final study and made it all possible – thank you!

Secondly, without Nora Fenske, I would not have believed in what my data said to me. Thank you ever so much for your R syntax support and helping me to prepare the data. Along these same lines, I have to thank the Centre for Multilevel Modelling at the University of Bristol. You don't know me and I don't know you, but your valuable videos and explanations were essential to my understanding.

A warm thank you goes to Raphael Berthele – thank you for your patience, support, advice and the chance to come to Fribourg for some inspiration.

Were it not for Patrick and Fredi Büchel's unwavering belief in me, for their ideas and constructive criticisms, the last four years might have been a lot easier but I may not have learned as much and confusion does indeed lead to learning, n'est pas?

To Martina Defuns, I would like to say thank you for your valuable work in scoring hundreds of written samples.

To Barbara Houtz – I so much appreciate your valuable proofreading skills and the time you took to do this. Thank you for helping me to improve my long-lost English language skills. Merci!

Finally, to my workmates at the Zurich University of Teacher Education: Thank you Kathleen Gallagher for being my friend; thank you Daniel Stotz for getting me going on the whole project and remaining curious; thank you Alex Buff for your statistical support and letting me pick your brain; thank you Christoph Suter and Annette Landau for giving me the chance to finish by approving my request for financial support and freeing up my schedule.

## Abstract

Elementary school English language teachers in Swiss public schools often undergo criticism due to their supposed lack of English language proficiency though this is only one small part of the full set of skills necessary for certification. This study set out to discover if young learner performance in reading, writing and listening after one or two years of instruction with the same teacher is associated with the teacher's measured language proficiency, with the teacher's feelings of improvement, with the teacher's contact with English outside the classroom as well as with teacher profiles of time spent on reading, writing, speaking and listening skills in the classroom.

Findings from this study suggest that neither teacher language proficiency nor teacher exposure to English outside of the classroom are determining factors in learner performance in the first two years of English instruction - indeed high levels of teacher proficiency in speaking and grammar here are negatively associated with learner performance. In this study, what positively associates with learner performance is a teacher's feeling of continually improving his or her speaking skills and instructional time on certain combinations of language skills.

Implications from this research should help improve English language teaching effectiveness by refining priorities in teacher training.

*Keywords: teacher effectiveness; early foreign language instruction; English language teaching; teacher qualification; teacher training*

## Table of Contents

Acknowledgements.....	iii
Abstract.....	iv
List of Tables .....	vii
List of Figures .....	viii
I. Introduction .....	1
1.1 Motivation and relevance .....	2
1.2 Overview.....	6
II. Literature Review / Theoretical Framework .....	8
2.1. English language teaching in Europe and Switzerland .....	9
2.1.1 The European context.....	9
2.1.2 The Swiss Context.....	12
2.2 Non-Native Language Teacher Proficiency.....	19
2.2.1 What does “highly proficient” mean? .....	22
2.2.2 The Swiss definition of highly proficient .....	26
2.2.3 The quality of the linguistic model or the teaching skills?.....	31
2.2.4 Can we teach what we don’t know?.....	35
2.3 Different types of teachers, different types of knowledge .....	38
2.3.1 General models of teacher knowledge from foreign language teaching.....	41
2.3.2 Teacher qualification and content knowledge .....	45
2.3.2.1 Teacher qualification .....	46
2.3.2.2 Teacher content knowledge .....	50
2.3.2.3 Language and foreign language content knowledge.....	52
2.3.3: Pedagogical content knowledge and instructional practices .....	57
2.3.3.1 Pedagogical Content Knowledge.....	58
2.3.3.2 Instructional practices and their effect on achievement.....	60
2.4 Consolidation of the literature and open questions .....	74
2.5 Working hypotheses.....	77
2.5.1 Teacher proficiency level.....	77
2.5.2 Teacher learning curve .....	78
2.5.3 Teacher exposure to English .....	79
2.5.4 Teacher time teaching language skills profile .....	79
2.5.5 Combined / multivariable model.....	80
III. Development and preliminary results .....	81
3.1 Learner subtests .....	82
3.1.1 Reading and listening test development .....	82
3.1.2 Writing test development .....	88
3.1.3 Learner test results.....	94
3.2 Teacher survey.....	97
3.2.1 Teacher survey overview.....	97
3.2.2 Teacher survey findings.....	99
3.2.2.1 Survey section 1 results – Teacher language skills .....	99
3.2.2.2 Survey section 2 results – Teacher contact with English.....	102
3.2.2.3 Survey section 3 results – Instructional practices .....	102

IV. Associations between teacher factors and learner performance .....	106
4.1 Regression Methods .....	107
4.2 Results for individual predictors.....	109
4.2.1 Teacher proficiency level.....	109
4.2.2 Teacher learning curve .....	116
4.2.3 Teacher exposure to English .....	123
4.2.4 Teacher classroom skills profile .....	132
4.2.5 Summary of individual regressions .....	138
4.3 Results for a combination of predictors – linear mixed models.....	140
4.3.1 Method.....	140
4.3.2 Results.....	145
4.3.3 Summary of findings from the linear mixed models .....	150
4.4 Consolidation of findings.....	152
V. Discussion .....	155
5.1 General discussion .....	156
5.2 Way over their heads or way below par? .....	161
5.3. Do what and how matter? .....	165
5.4. Learning along with one's learners .....	168
VI. Implications.....	170
6.1. In-service and pre-service teacher training .....	171
6.2 Messages to teachers? .....	175
6.3. Selection.....	182
VII. Conclusions.....	185
7.1 Looking back.....	186
7.2 Looking forward.....	189
Appendices.....	191
8.1 Teacher survey and survey summary data .....	192
8.2 Learner test.....	196
8.3 Instructions to teachers .....	207
8.4 Sample R Codes.....	209
8.5 Associations – Teacher feeling of change (total) and learner subtest scores.....	210
References .....	211

## List of Tables

Table 1: Status of English/French in EDK East Cantons, 2014/2015 school year.....	13
Table 2: Teaching English final exam grading rubric at the PHZH.....	17
Table 3: Working definition of language proficiency needed to teach elementary school English and rationale .....	30
Table 4: Definitions of pedagogical content knowledge for foreign language teachers .....	59
Table 5: Reading test overview .....	84
Table 6: Listening test overview .....	85
Table 7: Children’s self-assessment of listening skills (example) .....	87
Table 8: Writing – scoring rubric.....	89
Table 9: Interrater reliability – learners writing test scoring.....	91
Table 10: Descriptive statistics of the learner subtests and total scores .....	95
Table 11: Cambridge Exams Overview .....	97
Table 12: Conversion chart and teacher scores .....	100
Table 13: Teacher subtest scores .....	101
Table 14: Teachers’ feeling of skill change since exam.....	101
Table 15: Teacher contact with English outside of the classroom .....	102
Table 16: Class activities .....	103
Table 17: Learner total scores grouped by teacher total scores .....	109
Table 18: Learner listening, reading and writing scores grouped by teacher total scores .	110
Table 19: Linear mixed models for teacher and learner scores .....	111
Table 20: Learner total scores grouped by teacher feeling of improvement on all skills....	116
Table 21: Associations between teacher feeling of improvement since the official exam and learner scores.....	118
Table 22: Learner total scores grouped by teacher time abroad.....	124
Table 23: Association between teacher years abroad and learner performance.....	124
Table 24: Association between teacher holiday contact and learner scores .....	125
Table 25: Learner total scores grouped by teacher frequency of contact in specific language skills .....	126
Table 26: Association between teacher regular weekly contact and learner performance	127
Table 27: Teacher total contact and learner total scores .....	130
Table 28: Association between total teacher regular weekly contact and learner performance .....	130
Table 29: Learner total scores based on teacher time in German per lesson .....	133
Table 30: Association between teacher time in German and learner performance .....	133
Table 31: Descriptive statistics for teacher types.....	134
Table 32: Association between teacher type and learner performance .....	135
Table 33: Variable selection for the final mixed model.....	142
Table 34: Full and reduced multivariable mixed models .....	146

## List of Figures

Figure 1: Andrews' 2001 model of teacher knowledge (p.10) .....	41
Figure 2: Freeman, Orzulak and Morissey (2009, p. 83) – Elaborated frame .....	42
Figure 3: Freeman, Orzulak and Morissey (2009, p. 86) – Emergent frame .....	43
Figure 4: Writing sample from high scoring learners .....	92
Figure 5: Writing sample from low scoring learners.....	93
Figure 6: VRIPQ model (Cambridge, 2013, p. 10).....	98
Figure 7: Association between teacher total scores and learner total scores.....	113
Figure 8: Association between teacher Use of English scores and learner listening scores.....	113
Figure 9: Association between teacher speaking scores and learner listening scores .....	114
Figure 10: Association between teacher total change and learner total scores.....	119
Figure 11: Association between teacher speaking change and learner total scores .....	120
Figure 12: Association between teacher speaking change and learner listening scores...	120
Figure 13: Association between teacher Use of English change and learner listening scores.....	121
Figure 14: Association between teacher listening contact and learner total scores .....	128
Figure 15: Association between teacher reading contact and learner total scores.....	128
Figure 16: Association between teacher writing contact and learner total scores .....	129
Figure 17: Association between teacher speaking contact and learner total scores .....	129
Figure 18: Association between teacher balance of skills and learner total scores.....	135
Figure 19: Association between teacher balance of skills and learner listening scores.....	136
Figure 20: Association between teacher balance of skills and learner reading scores.....	136
Figure 21: Association between teacher balance of skills and learner writing scores .....	137



## I. Introduction

In my role as a teacher trainer, I attempt to provide theory-driven practical ideas to my future and current teachers who are often fairly allergic to the word ‘theory’. As Berthele (2009) puts it:

Etwas karikierend könnte man sagen, dass sich die Forschung zu 90% für das Problematisieren interessiert und im besten Fall 10% Empfehlungen abgeben kann, während die Praxis sich zu mindestens 90% für möglichst praktikable Empfehlungen und Lösungen interessiert und sich allenfalls zu 10% problemorientiert reflektieren kann.<sup>1</sup> (p. 11)

What teachers want to know is how they can or should “teach” – what is effective, ineffective, and they ask for my suggestions given a certain situation. My students and teachers want solutions that support their gut feelings which stem from their own subjective experiences and notions of how things should be in the classroom – sometimes justification for fine-tuning their gut-feelings can be found through the research, other times a full paradigm shift might be required in their habits and teaching behaviors and these are the cases that are the most difficult to change, or as Bruner (1996) states: “Knowledge helps only when it descends into habits” (p. 23).

Thus over the years, I have become dissatisfied with answers I have given to questions concerning the level of English teachers need to teach English. I have heard myself say to one person “it doesn’t matter, you just have to be a good teacher!!” and then a few moments later, I encouraged another person to enroll in a language course as their skills (in my humble opinion) were not up to par and their learners would soon “speak better than they do”. What level of proficiency do teachers really need to teach English? Isn’t their exposure to the language outside the classroom and their motivation to be good teachers enough to maintain their skills to teach? Can they teach well if they don’t have the “right” level?

These questions were the motivation to try to find some solutions or indications of what might be effective teacher attributes in foreign language teaching to perhaps

---

<sup>1</sup> All translations henceforth appear in italics in the footnotes and are my own. [To slightly exaggerate, it could be said that research is 90% interested in pointing out problems and can at best provide suggestions 10% of the time whilst the practice is at least 90% interested in recommendations and solutions which are as practicable as possible but is, at best, capable of problem-oriented reflection only 10% of the time. (Berthele, 2009, p.11)]

narrow the 90% question / 10% solution gap that Berthele (2009) mentions, with the full knowledge that “further research is [always] necessary” (p. 11). The purpose of this research is to delve into the topic of teacher effectiveness in the young learners’ English language classroom in Switzerland - which elements of teacher qualification, teacher use of English outside of the classroom and teacher estimation of time spent teaching particular skills (reading, writing, speaking and listening) in the classroom can be associated with learner performance? These questions are relevant to teacher trainers and to policy makers who may want to know more about reasonable requirements for teacher certification and want to know upon what to focus in teacher development.

### 1.1 Motivation and relevance

**Motivation:** The motivation for diving into the topic of teacher effectiveness in the Swiss elementary school English classroom stems from side-comments overheard and those blatantly stated time and time again in compulsory in-service<sup>2</sup> English teaching methodology courses in the Swiss cantons<sup>3</sup> of Zurich and Schaffhausen. Teachers question their own ability to teach English as they do not always feel comfortable in the language and worry about the effect of their “imperfect” language skills on their learners. Pre-service teachers in their student teaching field experiences also pick up on comments from their mentor teachers: “Ein paar Englischlehrpersonen scheinen an ihre eigenen sprachlichen Grenzen zu gelangen und suchen nach eigenen Wegen, um die Probleme zu überwinden”<sup>4</sup> (Hasler, 2006).

In thinking about teaching English, both in-service and pre-service teachers are principally concerned with their own linguistic ability to be able to teach in English, on whether or not their own language performance affects that of their learners and they less often show concern about how they are actually teaching or how they plan to teach - they have or are pursuing degrees in education after all, so teaching is not really “the problem”. Through articles found in *Non-Native Language Teachers*:

---

<sup>2</sup> In-service teachers are fully qualified working teachers who are pursuing a continued education course, here in order to teach English, as they have not yet been qualified to teach this specific subject. Pre-service teacher-training refers to students who are in the process of becoming fully certified teachers.

<sup>3</sup> Cantons are Swiss federal subdivisions; twenty-six cantons comprise Switzerland.

<sup>4</sup> [Some teachers seem to have reached the limits of their linguistic ability and are trying to find their own ways of solving this problem. (Hasler, 2006)]

*Perceptions, Challenges and Contributions to the Profession* (Llurda, 2006), we see that this is not uncommon for non-native speaking foreign language teachers.

This study aims at discovering if these feelings of inferiority are justified, at providing some concrete advice in teacher training and while the focus is set on English, it may also be relevant to any modern foreign language taught by a non-native teacher and where foreign languages are a compulsory or even elective part of the elementary school curriculum. As a teacher trainer, I want to be able to provide my students with some concrete suggestions. Should my teachers invest as much time and money in foreign language proficiency as often they do? Are they rightfully criticized when their language skills are considered “sub-standard”? Don't their teaching skills and effects of these on learners merit more scrutiny than their language proficiency?

Certainly countries which invest in the development of their teachers often have successful school systems, and teachers who invest in their skills demonstrate motivation to teach (which would hopefully have a positive effect on the learners). And clearly many studies (e.g. Bell, 2005; Huang & Moon, 2009) show that teacher attributes and qualities in the classroom are more important to learner performance than teacher language proficiency itself - so where should the focus in training be placed, on language proficiency or on teaching skills or on something else?

**Relevance:** First and foremost, this study has implications on pre-service and in-service teacher-training: knowing more about which factors attributed to the teacher interact with learner performance can help define elements to emphasize during teacher training. For example, if becoming aware that one specific skill, such as teacher speaking ability, is more important than another skill (teacher writing ability, for example) on measures of learner performance in the target language, then it might be worth placing more of an emphasis on teacher speaking skills in teacher training programs than is currently being done. Likewise, knowing if teachers who performed poorly on measures of reading have classes who performed just as well as teachers who performed exceptionally well on reading, then it might be said that a teacher's reading skill ability does not interact with learner performance, thus it is not a skill worth overly emphasizing in teacher training, even though it is a skill that is indeed currently quite at the forefront of curriculum design in some universities of teacher education. Moreover, similar associations can be studied between a teacher's time spent abroad and regular contact with English as well as a teacher's

time spent teaching each reading, writing, speaking and listening skill or combination of skills and learner performance. Findings in these regards have the aim of improving teacher education – knowing what is important to learner performance can guide teachers in choosing appropriate continuing education courses for English and other topics.

Secondly, this study has implications on teaching and testing in the classroom. There are few tests of learner performance for young learners of English and teachers shy away from testing. Thus the test developed for learners perhaps provides a model for teachers to analyze during training courses and can help them see that test-design does not have to be translating word lists. The results of the learner test provide a snapshot of what is possible for third and fourth graders to achieve in reading, writing and listening to English. Furthermore, knowing what balance of teaching time on specific skills' (reading, writing, speaking, listening) instruction in the classroom situation leads to better results in learner performance on specific skills can indicate whether a focus on a specific skill or a combination thereof should be used when teaching.

Thirdly, this study might provide support for decisions made and standards set by cantonal boards of education. Since the integration of English into the curriculum of many Swiss cantons, the question about the level of competence needed for qualification to teach in English has undergone much scrutiny. The Swiss Conference of Cantonal Ministers of Education (EDK)<sup>5</sup> recommends the Common European Framework of Reference (CEFR) (Council of Europe Council for Cultural Co-operation, 2001) level of C1 (proficient user – effective operational proficiency) for elementary school teachers and C2 (proficient user – mastery) for secondary school teachers. However, the author's discussions with teachers indicate that teachers themselves believe that such a high level is not necessary, that a B2 (independent user – upper intermediate) level would suffice. Academics and native speaking teachers show quite a bit of disparity in their arguments – some claim that “teachers don't realize that children will soon know English better than their teachers”, while others say it is teaching that matters, not a teacher's language skills. So what does “enough language to teach” mean? Under what criteria are these

---

<sup>5</sup> The EDK is an inter-cantonal commission of educators and cantonal board of education directors that sets guidelines, not legislation, for standards in education on the primary, secondary and tertiary levels. Refer to <http://www.edk.ch> for more information.

standards set? If it is shown that a teacher's subject-specific performance is not associated with learners' overall performance, then it might be said, for example, that within the range of B2/C1, having teachers with lower or higher performance is of no consequence to learner performance. Furthermore, if it is shown that a short stay abroad has just as much relationship to a learner's performance as a longer stay abroad, then this can help set qualification requirements in teacher certification.

## 1.2 Overview

In an attempt to answer these questions, or at least better define possible answers, an overview of expected teaching qualifications in the Swiss setting and existing research from various countries will provide the background as to what has and has not been studied about the effect of teacher subject-specific and pedagogical knowledge<sup>6</sup> on learners. Subsequently, the empirical part of this project aims at determining if there is any difference in beginning learners' performance on an end-of-year reading, listening and writing test amongst classes having teachers: with higher or lower levels of English; spending more or less time in contact with English outside of the classroom; and spending more or less time on specific skills during the lessons.

Chapter I has set the stage and provided the motivation for this study, provided an introduction to the issues at hand, and described the relevance of this study to teacher training and teaching.

Chapter II presents the state of and some research related to English language teaching in Switzerland and Europe and some of the policy decisions about teacher qualifications. Chapter II covers the rather large ground of teacher knowledge – in general, in language teaching and in the context of non-native foreign language teachers and what it means to know a language that is not one's mother tongue. Similar discussions around subject-matter knowledge and of the effects of various types of knowledge on learner performance will be exemplified. Finally, research around instructional design and the effects on learner performance will be touched upon. This chapter ends with the research questions and some hypotheses in light of the literature.

Chapter III describes the preliminary research findings and the development of the research instruments – the reading, writing and listening test for the learners and the survey for the teachers. Furthermore, statistics are presented from the official round of data collection in 2012 and 2013. This detailed information provides data on a sampling of learners and teachers in the eastern part of Switzerland. This section

---

<sup>6</sup> Generally, in this context, subject-specific knowledge is a teacher's measured proficiency in English and pedagogical knowledge is a reflection of a certain teaching style. These terms will be defined in Chapter II, part 2.3.

is useful for those looking into test development and those wishing to have a description of young learner performance in English.

Chapter IV provides the inferential results and findings to the research questions and hypotheses, including which factors on the part of the teacher are associated with the performance of the learners. The regression models used are presented here and the results from the teacher survey and learner test are combined.

Chapter V discusses the relevant findings of the regression models in Chapter IV and the data gathered in Chapter III.

Chapter VI covers the implications of this study on in-service and pre-service teacher training. The associations between certain teacher factors on learner performance will be analyzed for what this means in the practice.

Finally, Chapter VII is a critical reflection on this study with some final conclusions and ideas for future research.

## **II. Literature Review / Theoretical Framework**

Requirements for English language teaching certification in Switzerland and Europe are very similar as all require mandatory components of language proficiency and methodology. These decisions are set by boards of education with certain reasoning as to what is important for a teacher to know and be able to do to teach generally, and to teach specific subjects. With the aim of creating hypotheses for the research questions (provided Section 2.4 of this chapter), this literature review winds in and out of English and non-English language teaching specific contexts and through examples from other subjects in the hopes of transferring what can be learned in these different contexts to English language teaching. It attempts to dissect the various factors that can influence learning from various levels of teacher subject knowledge and expertise.

Section 2.1 starts with an overview of English language teaching in Europe and more specifically in Switzerland in order to provide the context for this study. Section 2.2 addresses the issues of non-native teachers and what it means to be highly proficient enough to teach. Section 2.3 covers the ground of knowledge – both content and pedagogical – and includes a section on instructional design, which explores the match between pedagogical knowledge and meaningful activities in language classrooms. The role of teacher qualification is addressed in this section as a measure of knowledge. Studies from various disciplines are analyzed for what they can tell us about foreign language teaching. Section 2.4 consolidates this literature and offers more open questions. Finally the working hypotheses for this study are stated in Section 2.5.



## 2.1. English language teaching in Europe and Switzerland

### 2.1.1 The European context

For several decades, the general trend in Europe has been towards an increase in the number of years during which the teaching of at least one foreign language is compulsory, and a lowering of the age at which this begins (Education, Audiovisual and Culture Executive Agency, 2008). Thus, modern foreign languages are increasingly taking up more space in school curricula and are being taught by classroom teachers in public school settings to younger and younger learners, with the effect that learners leave their school careers having had foreign languages for longer periods of time and have hopefully attained higher levels of fluency than previous generations. A report from the Barcelona Council (European Council, 2002) recommends the teaching of at least two modern foreign languages both starting at the elementary school (point 44, p. 19). Most European nations, including Switzerland, currently follow these guidelines. In most countries, these changes occurred mainly after 1984. Between 1994 and 2003, the average number of years during which teaching of least one foreign language was compulsory rose from 8.4 to 9, or by around one school year (Education, Audiovisual and Culture Executive Agency, 2008). Although there is quite a range of languages being taught throughout Europe - from local languages, other national languages, border languages, to English as a *lingua franca* and even Mandarin Chinese - it is more often the case that English and another language are being introduced into the elementary school curriculum.

In the majority of European countries, foreign languages are taught in primary education by general teachers who are qualified to teach all (or almost all) subjects in the curriculum. This applies irrespective of the curricular status of the language concerned, as a compulsory subject or a core curriculum option. (Education, Audiovisual and Culture Executive Agency, 2008, p. 58)

As of 2012 this had not really changed as explained in the most recent Key Data on Teaching Languages at School in Europe (Education, Audiovisual and Culture Executive Agency, 2012) which clearly shows the complexity and variation of teacher certification requirements amongst the various nations of Europe to teach foreign languages – ranging from generalist teachers who teach every subject within a

national curriculum to specialist teachers who have further qualifications to teach the target language to teachers with no specific qualifications to teach (in) a foreign language.

In general, it seems that a lack of qualified foreign language teachers is often a matter of concern in primary education. This may be because children now learn foreign languages at an earlier age and education systems have not yet fully adapted to these changes. A number of measures have been taken across Europe to address the shortage of foreign language teachers in primary education. These actions include programmes for upgrading the qualifications of generalist teachers as well as revising the content of initial teacher education for prospective primary education teachers. (Education, Audiovisual and Culture Executive Agency, 2012, p.87)

The requirements to teach foreign languages to elementary school children range from country to country. In some countries, the provision to teach a foreign language is a four-year degree with a major in the language and a one year general teaching credential. This supports knowledge of a specific subject (knowing the language well and some teaching skills) but perhaps not subject-specific pedagogical or general pedagogical knowledge. In other words, teaching skills that are specific to a subject (such as specific phonemic instruction) or a deeper understanding of general teaching skills (such as classroom management skills) may not be taught. Other countries let generalist teachers upgrade their skills to teach a new language, thus supporting the other direction – placing a value on general teaching skills but perhaps less on subject-specific knowledge. Enever (2014) provides an overview of multiple countries, their forms of certification, and the means by which they are attained. She concludes,

Firstly, it is evident that questions of course design for generalist teachers of primary FLs [foreign languages] should be addressed at both national and European levels, including a strategic plan for dissemination via the OMC [Open Method of Coordination]. Secondly, there is an urgent need to substantially increase the provision and availability of in-service courses and workshops for teachers if quality is to be improved and sustained. (p. 241)

Furthermore, only about half of the European countries in the Key Data 2012 report have official suggestions that pre-service teachers spend time abroad in a country where the target language is the language of communication. As we will see in Section 2.1.2, it would appear that teacher education in Switzerland provides a relatively thorough model of teacher education in general and for the teaching of foreign languages as well.

### 2.1.2 The Swiss Context

**Elementary school languages.** The Swiss situation is no different than the situation in the rest of Europe. Stemming from these European trends, the Swiss guidelines (Schweizerische Konferenz der Rektorinnen und Rektoren der Pädagogischen Hochschulen) from March 2004 suggest that cantons introduce two modern foreign languages into the elementary school, thus supporting the longer-term and earlier-start guidelines set by the Barcelona Council (European Council, 2002). By 2013, all Swiss cantons had adapted their curricula to integrate two foreign languages - a first language in the third grade (8-9 year olds) and the second foreign language in the fifth grade (10-11 year olds) – the 3/5 model. This incorporation came in some cantons at the cost of handicrafts and German lessons, which had to be cut in order to fit in an additional subject. In all of the German speaking cantons of Switzerland excluding the border cantons (EDK East cantons), English is not just an elective, but rather a compulsory subject, and is taught to third graders (9-10 year olds). Zurich is the only exception to this: English lessons start one year earlier, with second graders. French language instruction, in most cases, remains a compulsory subject in the fifth grade, as it has been for many years. However, despite all the hype around the new “common core” curriculum (Lehrplan 21<sup>7</sup>) for all the eastern cantons, individual cantons are slowly bending the “rules”, as shown in Table 1 (Schweizerische Konferenz der Rektorinnen und Rektoren der Pädagogischen Hochschulen, 2013, p. 7).

---

<sup>7</sup> Lehrplan 21 can be consulted at <http://www.lehrplan21.ch/> and is an attempt to standardize and converge the individual cantonal curricular standards into one document. This has not yet been accepted by all cantons and is quite controversial.

*Table 1: Status of English/French in EDK East Cantons, 2014/2015 school year*

Canton	1 <sup>st</sup> foreign language	Grade	2 <sup>nd</sup> foreign language	Grade	Status
Argau	English	3	French	6	French to be moved up when Lerhplan 21 comes in effect
Appenzell innerrhoden	English	3	French	7	3/5 in effect, situation being re-evaluated
Appenzell Ausserrhoden	English	3	French	5	3/5 in effect
Glarus	English	3	French	5	3/5 in effect
Lucerne	English	3	French	5	3/5 in effect
Nidwalden	English	3	French	5	3/5 in effect, situation being re-evaluated
Obwalden	English	3	French	5	3/5 in effect
St. Gallen	English	3	French	5	3/5 in effect
Schaffhausen	English	3	French	5	3/5 in effect, situation being re-evaluated
Schwyz	English	3	French	5	3/5 in effect
Thurgau	English	3	French	5	3/5 in effect; French to be moved to secondary school
Uri	English	3	French	7	French to remain in 7 <sup>th</sup> , Italian remains elective per 5 <sup>th</sup>
Zug	English	3	French	5	3/5 in effect
Zurich	English	2	French	5	3/5 in effect

Most recently, as of February 2014 and August 2014, the cantons of Schaffhausen and Nidwalden have decided to renege on the official EDK decision and will be putting the decisions whether to teach one or two languages and if one, which one, on the drawing board in the upcoming months. English and French language teaching in the German part of Switzerland has become quite a hot topic over the past years as reports from national newspapers (such as the *Neue Zürcher Zeitung*) will attest. The dispute around national languages being taught before or more extensively than foreign languages is an issue with which there is little agreement.<sup>8</sup>

English is not an official language of Switzerland, and it is therefore primarily taught by non-native teachers, and teachers who themselves have often learned English in a classroom setting as well as through stays abroad. Although there are no exact figures as to who is teaching English, in the canton of Zurich, where English

<sup>8</sup> Note October 1, 2014: There is currently upheaval going on in many Swiss cantons about the role of national and foreign languages and there are decisions and suggestions affecting the upcoming school years that are occurring on a daily basis. In November 2014, some cantonal ministries will be making decisions about French or English first and whether or not to eliminate one of the two compulsory languages in the elementary schools.

starts in the second grade and has been taught since 2005, and in the canton of Schaffhausen, where English starts in the third grade and has been taught since 2008, a rough estimate<sup>9</sup> would suggest about 95% of the teachers trained to teach English are non-native speakers of English (NNES). The same estimate can be applied to the other EDK East cantons listed in Table 1. The federal ministry of education requires that teachers be fully trained classroom teachers, not subject-specialist teachers.

**Teacher certification.** Pre-service elementary school teacher training in the EDK East cantons of Switzerland is a three-year bachelor's degree which includes the testing of basic skills and teaching knowledge in mathematics, German, science and social studies, and English or French as well as three other subjects of choice (French or English, physical education, music, visual arts, woodshop, or handicrafts). Field experiences and courses in general teaching skills and foundations of education and development are also requirements for graduation.

On the level of subject-matter knowledge, Swiss teachers are required to show proof of math skills, German language proficiency and a B2 or C1 level of foreign language skills depending on the canton. In some subjects like physical education and handicrafts, subject-matter knowledge and teaching knowledge of these subjects are a combined measure and not separated clearly. In the canton of Schaffhausen, in 2009 an estimated 30%<sup>10</sup> of the pre-service teachers failed the German language exam and had to re-take it. Every year, in both Zurich and Schaffhausen, a surprisingly large number of students have to retake the exam and of these, several drop out and have to wait a year to restart their studies should they so desire. There is a similar situation for mathematics, though perhaps not as dramatic. For foreign languages, teachers in most of the eastern cantons of Switzerland are required to attain a C1 level of English (or French) as shown through the Cambridge Advanced Exam of English or other similar qualifying exam. These requirements are called "basic skills" and there is an attempt to separate general teaching knowledge and skills from subject-specific knowledge in the Swiss system

---

<sup>9</sup> Over the past 12 years, the author has trained hundreds of teachers as part of the compulsory methodology courses at Zurich and Schaffhausen Universities of Teacher Education.

<sup>10</sup> This is an estimate and has not been confirmed. It was a comment overheard in the teacher's lounge and confirmed by students who were upset about this.

(though what the minimal subject-knowledge level necessary would be is fairly open to interpretation).

**Foreign language teacher certification.** On the level of teaching knowledge, Swiss teachers are qualified in their initial studies in many ways, on a general teaching level through coursework and fieldwork and on the level of foreign language teaching skills which are primarily measured through a final oral exam but also punctually through lesson observations with feedback. Once certified teachers, a more general model of teaching-skills' evaluation is in place (no longer subject-specific) through feedback by various levels of school-based personnel (e.g., board members, principals, peers).

In order to become fully certified to teach English (or similarly for French), pre-service students and in-service teachers in most EDK East cantons who have yet to meet the certification requirements to teach English or French need to:

- complete 4-6 ECTS points of coursework in the teaching of English / French at the elementary school level;
- attend a 3-week “teacher assistantship” in a country where English / French is the language of instruction;
- show proof of a C1 level of English/French (for English, to pass the CAE exam (or other recognized exam showing proof of the C1 level) with a score of A, B or C).

If these components have not been satisfactorily completed, teachers theoretically may not teach English or French, though in many schools they either do anyway or receive reduced pay until all parts have been satisfactorily completed. Not passing the required language exam (though proof of having taken it must be provided) is less serious – a teacher has approximately five years in which to complete it though again theoretically receives a decrease in official salary. The required level of language proficiency varies from B2 to C1 in various non-EDK East cantons, or in those cantons where teachers receive a teaching degree qualifying them to teach all subjects. In the eastern part of Switzerland, pre-service teachers choose amongst several subjects and become subject-specialist teachers. These EDK decisions are simply “recommendations” and there is some freedom for the different cantonal boards of education to set their own guidelines.

There has been discussion amongst language teaching specialists from the public school sector in Switzerland as to whether or not the Swiss should invest in

creating a specific language teaching profile for teachers. This means: What do teachers in the public school system really need to teach a foreign language to 8-12 year-old children? As an answer, in part the EPOSTL (European Portfolio for Student Teachers of Languages, Newby et al., 2007) is being used in some cantons, though this tool is only used at a teacher trainer's individual initiative and is not a compulsory tool. A Swiss-wide attempt has been launched to develop a portfolio for Swiss public school foreign language teachers – the “Berufsspezifische Sprachkompetenzprofile für Lehrpersonen für Fremdsprachen” [Language Knowledge Profile for Foreign Language Teachers] (Egli Cuenat, Klee & Kuster, 2010), which is partially integrated into the Passpartout<sup>11</sup> project.

For English specifically, at the Zurich and Schaffhausen universities of Teacher Education (PHZH and PHSB respectively), it has always been policy not to combine the understanding of foreign language teaching (methodology) with language proficiency. Language proficiency standards are outsourced to the Cambridge Syndicate. So for the past twelve years, a student or teacher could not fail the methodology exam for lack of language proficiency though this exam does take place in English. This, however, has recently changed. As of 2014 elements of language are evaluated as can be seen in bold in Table 2 in the grading rubric for the methodology exam.

---

<sup>11</sup> The Passepartout project coordinates French teaching in the 3<sup>rd</sup> grade and English in the 5<sup>th</sup> grade in the Swiss cantons of Berne, Baselland, Baselstadt, Solothurn, Valais and Fribourg. See <http://www.passepartout-sprachen.ch/de.html>



*Table 2: Teaching English final exam grading rubric at the PHZH (Büchel & Sprague 2013)*

To receive a 6 [the highest grade], the candidate demonstrates:

- clear and detailed reference to and profound understanding of assigned textbook, in-class materials, additional reading and critical analysis of such;
- differentiated academic / professional discourse (**correct use of terminology**);
- an explicit reference to classroom experiences (positive or negative) in light of appropriate theories;
- the ability to elaborate through numerous relevant examples;
- very **coherent**, relevant and **articulate contributions** and **responses**;
- creativity, innovation and originality;
- a very high level of awareness and understanding of foreign language learning at different ages, stages and levels;
- significant, analytical, constructive **responses in partner interaction** indicating the ability to further develop partner responses.

In years past, these references to being coherent and articulate would not have been mentioned. However, due to the occasional situation where a future English teacher expressed him or herself so poorly in the language that content was not understood, it was decided to include an element of language. Whether this is a good decision or not might also be influenced by the findings from this study.

Hence, Switzerland is one of the few countries to place such rigorous requirements on their foreign language teachers as can be interpreted from some insight into the matter from the Key Data on Teaching Languages at School (Education, Audiovisual and Culture Executive Agency, 2012) report which describes the situation in many different European countries. To conclude, requirements for teacher education in Switzerland are fairly competitive in a European context. However, the same issues of what “basic skills” are when defining subject specific knowledge and pedagogical knowledge are rather disputed. Within Switzerland, though the EDK makes recommendations, cantonal ministries of education make decisions and set guidelines [ir]respective of these suggestions.

The proceeding chapters will follow up on ideas touched upon here by analyzing the research on the aspect of non-native teachers in the classroom, on subject-specific knowledge in various school subjects and from other countries in order to make hypotheses as to what skills might be necessary for Swiss elementary school English language teachers to have in light of the requirements set by various boards

of education. By looking outside of the Swiss borders and also outside of English language classrooms, some ideas might be able to transfer back within the Swiss English language classroom.

## 2.2 Non-Native Language Teacher Proficiency

The main population of European English language elementary school teachers are non-native speakers of English, and there is ample research on and controversy around non-native speaking teachers to devote some time to analysing the situation of their possible effects in the classroom. Although the situations in the US and Switzerland are quite different, discussions about teacher proficiency in both foreign language teaching and bilingual instruction have been held in the past and are still being held. In bilingual programs across the US, defining the necessary language proficiency of the teachers remains on an emotional level, not necessarily a scientifically founded one, which could be the same in Switzerland as well. As Grant (1997) alleges:

Decisions based on these tests [language competence exams] have far-reaching implications: they affect students' exposure to appropriate role models where language is concerned as well as teachers' careers, salaries and job security. And, although the need for bilingual teachers is very great, it is unfair to the teachers, parents and students alike not to require potential teachers to meet a certain minimum standard of language ability. The standards set not only affect the quality of Spanish used in the classroom but also the attitude towards bilingual education in the community. (p. 40)

Similar discussions have been held in the field of modern foreign language teaching in the United States where foreign language education is notorious for being of questionable quality but where guidelines set by the ACTFL (American Council on the Teaching of Foreign Languages) may be even more stringent than the European ones. The need to know what is important in foreign language teaching does undergo regular scrutiny, as Chambless (2012) points out.

The intuitive assumption of a causal connection between a teacher's oral proficiency in the TL [target language] and the teaching and learning that take place in classrooms enjoys scant support in the empirical research literature. However, an indirect connection can be made through research in second language acquisition (SLA) that indicates that the quantity and variety of TL input will indeed affect student learning. (p. 142)

On an international level, the question of teacher proficiency has been extremely relevant. If we look to Hong Kong, for example, there is a huge emphasis on “native speaking teachers” and norms which cost the government 1.5 billion dollars over five years for one such program (Coniam & Falvey, 1999b). The Swiss have probably spent just as much for a non-native norm. In both cases, have teachers received what they actually need to teach young learners? Is the amount of money and energy spent on one small piece of a teacher’s skills worth the investment in what concerns the effects on learners?

Despite their supposedly “bad rep”, the important role non-native teachers of modern foreign languages play is clearly evident when looking at the number of publications and websites directed towards them<sup>12</sup> and the advantages non-native teachers bring to the classroom is well documented (e.g. Celik, 2006; Braine, 1999; Medgyes, 1999; Llurda, 2006). Among the many advantages in the classroom, non-native teachers of English mostly speak the community language of instruction, thus they understand their learners; they are aware of use of English outside the classroom and can encourage learners to take advantage of certain settings; and thus know why English is to be taught and in which situations it will be used by their learners outside of the classroom. Furthermore, they may know when to use the students’ mother tongue in foreign language lessons to foster learning, such as to explain a grammar rule.

Further research on non-native teachers can be primarily broken down into several areas: identity (e.g. Lazaraton & Ishihara, 2005); self-awareness (e.g. Llurda & Huguet 2003); error recognition (e.g. Murray 2002); relationships between native and non-native speaking teachers (e.g. Matsuda & Matsuda 2001); and learner attitudes towards non-native teachers (e.g. Lipovsky & Mahboob 2007, Murray 2002). All of these factors may be more important than the specific language skills of the teacher, yet it is the language skills themselves that are often the most criticized.

**Summary.** As discussed in Chapter I, in Switzerland all public elementary school foreign-language teachers have to be fully certified teachers from a Swiss or other accredited teacher training college. The value of non-native teachers has been

---

<sup>12</sup> See <http://nnest.asu.edu/> (last access October 1, 2014) and various books listed in the references, to mention a few.

much reviewed and there have been attempts to define a more differentiated profile of teacher language skills for native and non-native teachers alike. The valuable and beneficial role non-native English speaking teachers (NNESTs) play should not be undermined, yet there needs to be more literature supporting their contributions as well as showing the necessary proficiency levels which are at least not a hindrance to “correct”<sup>13</sup> uptake by learners. Furthermore, contributing to a growing field of research with younger learners, not teenagers and adults, but children in an elementary school setting is essential as more early foreign language programs are becoming established throughout the world.

The following sections will continue by looking at the role of the language proficiency of the non-native English teacher in defining what “proficiency”, or “high proficiency” means and how it is locally defined in Switzerland. The role of the quality of the linguistic model and if teachers can teach what they don’t master will be discussed.

---

<sup>13</sup> The word “correct” is rather vague as it depends on which norms are set – standard norms of a specific country such as Standard American English, standards based on international corpora, standardized local norms or simply a subjective judgment.

### 2.2.1 What does “highly proficient” mean?

The uncorrected quote below provides an example of how a native English speaking lay person (non-teacher) judges the decision of having qualified Swiss (non-native English speaking) teachers as English language teachers:

It's scary isn't it? Someone I know did exactly this. She was teaching 1  
 teachers who would be teaching English to Swiss school kids. Their English 2  
 was terrible and they had no interest in improving it. They were only 3  
 attending the lessons because they were forced to. Since there are a tons of 4  
 native English speaking teachers who would love to come and teach English 5  
 to kids in the school system here you'd think that this would be a perfect 6  
 solution. But no, to avoid bringing in a foreigner they'll use natives with poor 7  
 English skills to teach English... (Mark, 2006) 8

From a linguistic point of view, who does the underlined “their” (line 2) refer to – the children or the teachers? The assumption here is that this person is referring to the teachers. The second issue is a slight oversight in the use of “a tons” (line 5) – the assumption here being that this is simply carelessness. Thirdly, one indeed does have to wonder about the choice of the word “natives” (line 8) – but assumptions here are perhaps not worth the words.

From a content point of view, this is a typical example of what Swiss teachers deal with from native English-speaking parents or professional parents who use English regularly. This attitude can also be seen in informal discussions in villages by parents of all walks of life. Teachers in CH undergo criticism from parents, from other members of staff and even from their learners due to their “imperfect” language skills.

Though the debate between native versus non-native language teachers is not an overt question in this study, it is covertly understood that the better a teacher’s language skills are, the better the teacher s/he is. The research in the field of NNESTs can be enlightening in finding out what may be important to consider in teacher training for the goal of teacher effectiveness. In commenting upon the final chapter of *Non-native Language Teachers: Perceptions, Challenges and Contributions to the Profession* (Llurda, 2006), Brock Brady (2007) states “that NNESTs often lack self-confidence in their English proficiency...Regrettably though,

the chapter is long on polemic and short on actual research” (p. 234). Llurda (2006) addresses the problems proficiency plays in that

[NNESTs] are likely to experience difficulty in:

- 1) conveying messages to their students in the target language;
- 2) addressing their questions on language use;
- 3) providing a good language model (p. 146)

....In sum, a high level of proficiency must be regarded as an essential condition for language teaching, but one which alone does not guarantee successful language teaching.” (p. 234)

Could it be that the level of language competence needed to teach, one factor playing a role in this so-called lack of self-confidence, is perhaps not as high (or “essential”) as many state systems set it as or as many teachers think it ought to be? If the level of proficiency should be “high”, then what does this mean in the Swiss setting? Furthermore, would it not be possible for teachers with a lower proficiency level to compensate for this through clear tasks, learning-conducive settings, modelling via other sources of input, and tip-giving on the level of strategies?

In their 2009 article “‘Discovering Language’ in Primary School: An Evaluation of a Language Awareness Programme,” Barton, Bragg and Serratrice addressed the issue of non-specialist foreign language teachers. Though this study relied on pupils' self-evaluations and teachers' observations and did not include any test data or measure of knowledge of language, it did indicate that a discovery language program with or without teachers' having knowledge of the language can be beneficial to the general language awareness of the learners. Learners seemed to favor the language in which their teacher had the most extensive knowledge, possibly because this meant that the teacher had greater confidence or because s/he was capable of supplementing the resources provided with additional appropriate materials. This finding could as well indicate that the self-confidence of the teacher in his or her language skills, not necessarily the actual measured skills of the teacher in this language, can be telling in how instruction is perceived.

Another relevant work on the language competence of the teacher in the classroom is by Llurda (2005) in “Non-Native TESOL Students as Seen by Practicum Supervisors.” Llurda hypothesizes that “High-level language skills are essential for NNS (non-native speaking) language teachers' successful teaching.

Pedagogical skills are also important, provided an acceptable level of the former has been achieved” (p. 132). He insinuates that language skills take precedence over pedagogical skills. In his questionnaire to practicum supervisors, he asked questions about language skills of the student teacher, teaching performance of the student teacher and the relationship between the student teacher’s respective language proficiency and teaching skills. The results of this latter point came back quite diverse – answers to a question about whether the student teachers’ language awareness, fluency, grammar, listening comprehension skills, speaking rate, and foreign accent affect the students’ English teaching ability ranged from “negligible’ to ‘very important’” (p. 138). This defines a need to refine this question to find out at what level these factors do play a role with a specific target group and at what level they no longer play a role.

Llurda further states

...a NNS cannot guarantee his/her employer or students that s/he has total command of the language. The fact that there exists variability in linguistic proficiency probably explains some of the discrimination and negative *a priori* attitudes experienced by many NNS teachers. Once the students or administrators are aware of the high proficiency level of the particular NNS teacher, such attitudes tend to disappear. (Llurda, 2005, p. 44)

This statement clearly indicates that these impressions, even on the part of administrators, place teachers with high language proficiency at a distinct advantage. However, one must remain wary of impressions as when looking at the many other factors on the side of the teacher that contribute to effective teaching of language or effective teaching in general and those factors on the part of the learner on how the teacher factors are perceived and taken up – command of the language is one quality that may actually not be of much significance, but may be a consequence of another construct such as motivation.

One rare study at the elementary school level (Butler 2004) took up this topic from the perspective of teachers themselves. It was found that elementary school English language teachers in Korea, Taiwan and Japan feel there is a large gap between their own productive (worse) and receptive skills (better) and that they do not master the language well enough to actually be teaching in it. Butler concludes that "Research is also needed to better understand what types of competencies



(regarding both knowledge and the ability to use such knowledge) elementary school EFL teachers must have” (p. 269). This article clearly addresses the discussion about the various arguments which weigh the importance of knowing language and knowing about language.

There is also a bit of disagreement as to what is important in foreign language teaching as seen in the following example:

Although principals placed evidence of "proficiency in the language to be taught" as the second most important topic for a portfolio, they ranked subject matter knowledge (Standard 2) 6th in a tightly packed list of all-important state standards. In contrast, the FL [foreign language] chairs were consistent in rating Standard 2 first among standards and evidence of language proficiency as first among desired portfolio entries. (Hammadou-Sullivan, 2004, p. 398)

Though many mention the role of high language proficiency, very few researchers mention any detrimental effect of low language proficiency. In one rare case, Sadtono (1995) wrote about the "poor" language proficiency of EFL teachers in developing and other countries. He implied that many teachers do not have the language proficiency necessary to teach. He went on to dig into the concept that the "model" must be perfect for the learners to learn. He assumed the model to be a native model, did not reflect on a non-standard model as being appropriate in a English as a *lingua franca* world, and he did not mention the fact that the teacher is not the only source of input.

**Summary.** It is precisely the definition of “high proficiency level” that needs to be examined. Referring back to the Barcelona Council’s (European Council, 2002) report stating that local elementary school teachers will more frequently be asked to teach English to school children, then what should their relevant high level of proficiency be? The teachers in the Sadtono study were going back to their home countries to teach in an EFL (English as a foreign language) setting and it is not stated at what academic level. The term “near-nativeness” is too subjective to define and it is questionable as to whether it is even appropriate for teaching in central Europe, especially to younger learners. The research in this regards is imprecise and does not address the actual levels necessary to teach specific ages in specific settings.

### 2.2.2 The Swiss definition of highly proficient

On a purely bureaucratic level, the Swiss do not state that teachers should achieve near-native proficiency; statements to this effect are the subjective interpretations and attitudes of the stakeholders. Historically, the search to define what level should be set for English language teaching in Switzerland started in 1998. According to Stotz and Meuter (in Büeler, Stebler, Stockli & Stotz, 2001, pp. 120-122) in the pilot projects of English in the elementary schools, the teachers had proficiency between the levels of B2 and C1. Because a content and language based approach to English language teaching was encouraged, thus was a C1 level desired. It was assumed that teachers would have to be ready to react to the linguistic demands of a wide range of subjects in their teaching and be able to teach sport, math, arts and crafts and other subjects in English.

Furthermore, setting this high level was justified based on the idea that learners of this age are in the midst of a critical period of learning.

Die Tatsache, dass die kritische Phase des auditiven Systems um das neunte Lebensjahr liegt, ist ein starkes Argument dafür, nur erstklassig trainiert Lehrer im frühen Fremdsprachenunterricht einzusetzen. Oder anders formuliert: die hohe auditive Sensibilität von Kindern macht es sehr wahrscheinlich, dass die jeweilige Aussprache des Lehrers kopiert und multipliziert wird.<sup>14</sup> (Peltzer-Karpf und Zangl, 1998, p. 15)

This statement takes the same tone as the Sadtono (1995) study which emphasizes the importance of the model for uptake. However, even if there is a critical period for language learning, eight or nine year olds having two or three hours of a foreign language is not the same as being in an immersive setting. It could be that there is not enough input by the teacher to make a difference. Furthermore, this statement again neglects the idea of English as a *lingua franca* whereby copying a non-native accent is not learning “incorrectly” but rather another accepted standard of English. It also neglects the fact that teachers are not the only source of learners’

---

<sup>14</sup> [The fact that the critical age for the development of the auditory system is around nine is a strong argument for hiring only first class foreign language teachers. In other terms: the highly auditive sensitivity that children have makes it extremely probable that the teacher’s pronunciation will be copied and reproduced. (Peltzer-Karpf und Zangl, 1998, p. 15)]

English, especially in this day and age where learners are confronted with advertising in English, watch movies in English and listen to English music.

Since the initial pilot projects in the late 1990s, there has been little academic discussion about the decision of a C1 though teachers have been in opposition for a while.

Das in den Leitlinien angestrebte Niveau (C1/C2) ist sehr hoch. Entweder schafft man dafür die heute bei weitem noch nicht vorhandenen Voraussetzungen, oder man passt zumindest für die Primarstufe die Niveauansprüche nach unten an. Es ist in den nächsten zehn bis zwanzig Jahren damit zu rechnen, dass die Vorstellungen bezüglich Einsatz von Fachlehrkräften bzw. Fächergruppenlehrkräften an der Primarstufe wenig Fortschritte macht. Die Schulstrukturen in unserem Lande erlauben dies nur in einem Teil der Schulen; andere werden noch lange auf fachlich und fachdidaktisch sehr breit qualifizierte-Lehrpersonen angewiesen sein. Die ganze Diskussion über Allrounder versus Fächergruppenlehrpersonen dauert zwar schon lange, krankt aber immer noch an völlig unrealistischen, naiven Vorstellungen bezüglich der praktischen Einsetzbarkeit vor Ort.<sup>15</sup> (Zemp & Strittmatter, 2007, p. 3)

Officially, one source (Passpartout) that provides a justification of the minimal suggested proficiency requirement to teach (than the EDK suggestions) explains this by stating:

Der Kanton definiert die sprachliche Grundanforderung, d.h. die Einstiegskompetenz, für das Unterrichten einer Fremdsprache auf der Primarstufe mit dem Sprachkompetenzniveau B2. Ein B2 bedeutet, dass Sie über gute Kompetenzen für die persönliche Sprachverwendung verfügen. Das Unterrichten von Fremdsprachen setzt aber höhere

---

<sup>15</sup> [The C1/C2 target level expected from the guidelines is very high. Either we keep striving towards these aims - which we are far from meeting - or we lower the expected levels for the elementary school. We cannot count on getting subject-specialist teachers at the primary school level in the next ten to twenty years, our school structure does not allow for this. This is only possible in some schools in our country; other schools will be dependent on very generally qualified teachers for a long time to come. The whole discussion about generalist teachers versus subject-specialist teachers has indeed lasted a long time and is still suffering from totally unrealistic, naïve attitudes towards the practical implications on the schools. (Zemp & Strittmatter, 2007, p. 3)]

Sprachkompetenzen voraus, als man sie für den privaten Gebrauch benötigt. Deshalb wird als mittelfristiges Ziel für Primarlehrer/-innen ein Sprachkompetenzniveau C1 für Lesen, Hören und Sprechen resp. ein B2 für Schreiben definiert.<sup>16</sup> (FAQ Zu Passepartout, 2014, p.4)

A search into the decision making process in the EDK-eastern cantons of Switzerland as to the reasoning behind the C1 decision was fruitless apart from what is cited above even though original policy makers were contacted, archives were searched and team members questioned. It seems there was no quantified justification, but rather a feeling that perhaps is well stated by the Passpartout group members (who were not part of the original group making suggestions to the ministries of education in the late 1990s).

And thus with a glimpse into the Common European Framework of Reference for Languages: Learning, Teaching, Assessment, the description of level of C2 is explicitly not a definition related to near nativeness and thus by default neither is C1.

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. (Council of Europe Council for Cultural Co-operation, 2001, p. 36)

This is encouraging as it indicates that the setting of the level in Switzerland was a realistic interpretation of what is feasible and also necessary, though again grounded in what theories remains a mystery. Though the general description of C2 is not in reference to native-speaker norms, the mention of contact with native speakers in the specific CEFR descriptors is not a good reflection of the needs of the teachers in central Europe, as communication in English mainly takes place between non-native speakers of English as the research on English as a *lingua franca* will attest (see Jenkins, 2002 or Buechel, 2012). Thus the Swiss decision of

---

<sup>16</sup>[The canton has set the minimal foreign language proficiency requirements, meaning the starting out proficiency requirements, to teach a foreign language at the elementary school to a B2 level. A B2 level means having good skills for personal communication. However, the teaching of foreign languages requires higher levels of language proficiency than is needed for this type of use. Therefore a mid-term goal for elementary school teachers is a C1 in reading, listening and speaking and a B2 level for writing. (FAQ Zu Passepartout, 2014, p.4)]

C1, for whatever reasons it was selected as the level, is not defined by a reference to the native speaker though many people's understanding of striving to be better IS in reference to native speaker norms. It was set based on the idea that teachers should be flexible in using the language to deal with a large range of topics and that learners need a good linguistic model from the beginning.

**Working definition of necessary language proficiency for teachers.** To set a working definition of the level of proficiency necessary to teach in Swiss elementary schools, Table 3 provides an overview of some of the teaching-relevant descriptors lifted directly from the EAQUALS (Evaluation & Accreditation of Quality in Language Services, 2008) / ALTE (Association of Language Testers in Europe) Portfolio Descriptor Revision Project and a rationale as to why a certain CEFR level in a certain skill may suffice.

What is not stated in Table 3 is a given and that is a mastery of the local language of instruction. This is essential as this serves the purpose of allowing teachers to identify issues learners may have in the target language and clearly see what learners need the target language for, as the research on non-native teachers will attest. Thus in terms of English language proficiency, floating between a B1 and a B2 proficiency level in the target language may be all that teachers need, because though they need to be able to speak clearly and perhaps repair their speech, understanding their learners' spoken utterances may not be as demanding. Generally, a B2 level also represents a command of the language appropriate for usage in central Europe.

<i>Table 3: Working definition of language proficiency needed to teach elementary school English and rationale</i>			
Skill	Level	Justifying CEFR Descriptor*	Rationale (why not higher or lower)
Listening	B1+	I can understand straightforward information about everyday, study- or work-related topics, identifying both general messages and specific details, provided people speak clearly in a familiar accent.	B2 refers to complex speech, which is most likely not necessary with young learners. C1 refers to academic and extended speech.
Reading	B2	I can read with a large degree of independence, using dictionaries and other reference sources....	B1+ might be enough, but the reference in B2 to reading complex instructions can be useful for teacher's notes found in textbooks.
Information exchange	B1+	I can give practical instructions on how to do something...	The reference to "practical" in B1+ is what is needed in the classroom. B2 refers to "complex information" which may not be necessary.
Spoken production	B2	I can give clear, detailed descriptions on a wide range of subjects related to my fields of interest.	Generally, complexity increases with B2+ and C1, which not necessary when working with beginners.
Written production	B2	I can write at length about topical issues, even though complex concepts may be oversimplified, and can correct many of my mistakes in the process.	This is not clear – C1 refers to correcting "most" of the mistakes, but what teachers actually need to write is very little, so what is written might well be correct. B1+ neglects the level of modelling.
Strategies	B2+	Several references are useful: intervene appropriately in discussion, using a variety of expressions ; help the development of a discussion by giving feedback; overcome gaps in vocabulary; monitor speech and writing to correct slips and mistakes	B2 is too formulaic and teachers do need to have some elements of acting spontaneously. C1, on the other hand, requires more flexibility than is necessary for a teacher's needs.
Accuracy	B2 B1+	Accuracy: I can communicate with reasonable accuracy and can correct mistakes if they have led to misunderstandings. I can communicate with reasonable accuracy in familiar contexts, though with noticeable influences from my mother tongue.	B1+ does not mention the ability to correct mistakes, only that they are permitted. Unclear, though, is why in B2 the mention of influences from mother tongue has been removed. Can't one have "reasonable accuracy, correct mistakes AND have mother tongue influence"?
Range:	B1+	I have a sufficient range of language to describe unusual and predictable situations and to express my thoughts on abstract or cultural as well as everyday topics.	B2 mentions avoiding repetition, though repetition actually good for beginning learners.
Socio-linguistic	B1+	I can use uncomplicated language to interact in a wide range of situations in a neutral way.	B2 refers to native speakers and one would not want to amuse or irritate the native speaker.

\* The term 'justifying' is used as the minimal level where a level higher is not necessary for classroom English and is one point (among many) within the level that is relevant to teachers.

Of course, more might be expected from a teacher as a professional learner who reads and talks about teaching in English, but discussing teaching on the meta level outside the classroom is done in a German setting, therefore what is needed inside the classroom must be kept separate from what is simply nice to have or need from professionals functioning in an English speaking environment. With the very open references to what high proficiency is, there are not grounds enough for setting the definition to a C1 for this construct. Therefore, a working definition of a “high enough level of proficiency” is left at a general B2 level, slightly above what is realistically necessary as defined by CEFR descriptors. Research on accuracy, and the importance of the model and whether learners take on the mistakes of the teacher will be discussed in Section 2.2.3.

### **2.2.3 The quality of the linguistic model or the teaching skills?**

In this context, what level of competence might give the security and freedom needed to teach a foreign language but at the same time, not hinder “correct” uptake by the learners? Berry (1990) notes the value of a high level of competence as it increases a teacher’s confidence, makes using the target language in the classroom more flexible and allows for a broader use of methodology.

De Lima (2001) states that “in most parts of the world, the main emphasis in ELT [English language teaching] is on methodology and the language level of the future teacher. Language competence has, indeed, been rated the most essential characteristic of a good teacher” (p. 143). Unfortunately, studies described in this article do not look at already trained elementary school teachers who have a background in methodology and are fully certified educators. These articles do question, however, the balance of teaching courses and language competence courses, what do teachers need more – language competence or teaching methods? Indeed, looking at the studies provides mixed messages. On the one hand, a high level of competence is necessary, but how to define this is unknown. On the other hand, other factors like flexibility in methodology may play more of a role than just language competence itself.

Andrews (2003) addresses idea of the quality of the model in describing the interplay between learner characteristics, teacher teaching skills (pedagogical knowledge) and teacher language skills (content knowledge) where he states:

...learning may still take place outside the classroom, depending on the extent to which the learner has the opportunity and motivation to become involved in any L2 [second or foreign language] immersion. For many L2 learners, however, the classroom and any related activities taking place outside the classroom setting present their major opportunities for exposure to L2 input. Although they may encounter L2 input direct from sources such as the textbook ... and other students ..., much of the input learners are exposed to involves the teacher. The teacher may be the producer of such input: with the specific intention to induce learning, as in, for example, the presentation of new language; or less deliberately, through any communicative use the teacher makes of L2 in the classroom, such as for classroom management. The teacher may also 'shape' the input from the other major sources. In making use of the textbook, for instance, the teacher might modify (however slightly) the textbook's presentation or practice of a grammar point, or draw learners' attention to the occurrence and significance of a particular grammatical structure within a reading comprehension text. When encountering language produced by the learners, orally or in writing, the teacher has a range of options for handling that output, but very often teacher feedback will provide an additional source of input for learning (for the class or for the individual learner) as the student's original output is modified by the teacher. (p. 90)

So if we assume that much of the knowledge that is negotiated by the learner is initiated and modified by the teacher, then it must be important that the teacher's language skills be balanced and well-developed (again the definition of what "well" is here is not well defined). It is insinuated that the learning mediated through the teacher might be "more" than the implicit learning taking place outside the classroom (even in the case of English being a global language in which learners are frequently exposed) thus putting in question the role of teacher quality of the model. In Switzerland, this may not be the case as children's exposure to English is enormous in both print and auditory media. Furthermore, it could also be that observable characteristics of a teacher's own language learning strategies could well be modelled by the class, not simply the teacher's language features.



In a forum on *The Electronic Journal for English as a Second Language*, Walker (2006) states that

NS (native speaking) teachers can be excellent models. However, a good model is not by definition a good instructor. In this respect, the NNS teacher who has successfully mastered the main features of English pronunciation is probably better equipped to instruct learners than the native speaker, who, in terms of techniques, may be restricted to being a model for students to copy but not know how to teach a specific sound.

As more recent theories of learning point may suggest, if languages are not only learned behavioristically, through imitation, but in addition through pattern finding from all the sources of input, then this puts into question how good one specific model actually has to be, as there are many models from which to analyze patterns. This statement also leaves another point open: what does it mean to have “successfully mastered the main features of English pronunciation” to be a good model? Much work has been done on features of pronunciation in English in, for example, the *Lingua Franca Core* (Jenkins 2002) which may well dispute more traditional definitions of “good” pronunciation.

Árva and Medgyes (2000) found that non-native teachers, as compared to their native-speaking peers, had a harder time teaching pronunciation, vocabulary and colloquialisms and were much less flexible in finding appropriate alternatives and reformulating their speech. What effect this has on the learners is not mentioned, but what is mentioned is that non-native teachers had more of a goal-oriented approach towards teaching; whilst the native speakers were “okay, what shall we do today”, the non-native speakers had concrete plans about what they wanted their learners to learn. Furthermore, non-native speakers had an easier time teaching grammar. Thus this is perhaps an example of how self-judgments and others’ perceptions may be correct, but are not necessarily important in what is necessary to teach. The quality of teaching is measured in the achievement of aims and perhaps it is this point that is the real advantage of the non-native-speaker, not in the modelling, but in how aims are set.

In a small study by Kase and Jensen (2013), it was disclosed that learners make progress in the pronunciation of certain phonemes merely due to the model of having native-speaking teachers. This again puts in question the validity of assuming that the role of the native teacher as a model is important – learners can learn is what is

explicitly focused on and taught, not only what is necessarily indirectly around them, especially in the short term. Of course the importance of the implicit has been shown in learning in general (e.g. Reber, 1993) as well as in foreign language learning (e.g. Cetin & Flamand, 2013). However, the Kase and Jensen study suggests grounds for hypothesizing that a teacher's language competence is not the determining factor in what English learners learn, but rather where the teacher focuses instruction is what is learned.

This is also the case in a study from Switzerland (von Ow, Husfeldt, & Bader-Lehmann, 2012) which has provided some insight into factors on the level of the individual and teacher on learner performance. The teachers in this study come from the same population as those from this current study. Though the findings were primarily non-significant, they did find that teachers who focused more on speaking in assessing learners had classes who performed better (in speaking) than those classes with teachers who did not focus instruction on speaking.

Je häufiger eine Lehrperson die Mündlichkeit bei der Benotung berücksichtigt, desto besser sind die mittleren Leistungen der Schülerinnen und Schüler der Klassen. Die Bewertung der Mündlichleistungen ist dabei als Ausdruck dafür anzusehen, wie bedeutsam der Aspekt der Mündlichkeit für die Lehrpersonen ist. Positive Korrelationen über  $r = 0.1$  ergeben sich weiter zwischen der Erfahrung als Primarlehrperson und dem mittleren Leseverstehen der Klasse ( $r = 0.11$ ) und zwischen dem Vokabelnernen und den mittleren Leistungen der Klassen im Sprechen ( $r = 0.22$ ). Wenn also eine Lehrperson längere Erfahrung als Primarlehrperson hat, dann sind tendenziell die mittleren Leistungswerte ihrer Klasse im Leseverstehen etwas höher. Wenn eine Lehrperson mehr Wert auf das Vokabelnernen legt, dann sind in der Regel die mittleren Leistungen ihrer Klasse im Sprechen etwas höher. <sup>17</sup> (von Ow, Husfeldt, & Bader-Lehmann, 2012, p. 3)

---

<sup>17</sup> [The more a teacher takes speaking skills into consideration in grading, the better the class average score was. The assessment of speaking performance is an indication of how important speaking skills are for the teacher. Positive correlations above  $r=0.1$  were also found between the experience as an elementary school teacher and the average reading scores of the class ( $r=0.11$ ) as well as between the vocabulary learning and the average speaking performance of the class ( $r=0.22$ ). The longer the teacher has been an elementary school teacher, the higher the class' average reading comprehension scores were. If a teacher places an emphasis on vocabulary learning then the average performance of the class in speaking is somewhat higher. (von Ow, Husfeldt, & Bader-Lehmann, 2012, p. 3)]

From this study we can take that idea that where teachers believe something is important, then it will be measurable in the class performance but how these skills (here speaking or vocabulary) are taught, or which general teaching styles and methods are used, are not mentioned in this research. And if general teaching knowledge is measured by how long a teacher has been in the field, then perhaps other studies might show that more experienced teachers “produce” better readers.

**Summary.** So it could simply be that “time on task” is the relevant factor; and even though explicitly focusing on language that is non-standard could well possibly lead to the acquisition of non-standard features, this in and of itself is not something that hinders communication in the European setting, though it may add fuel to the fire about the quality of instruction. Moreover, though the mere linguistic features may not be taken up by learners, teacher non-mastery of the language might manifest in other teaching behaviors that are perhaps detrimental to learning such as a lack of teacher confidence or using “too much” of the local language.

It is not possible to have a simple yes or no answer to the question of the quality of the linguistic input because “poor” quality can certainly be compensated for through textbooks and audio materials or strategies and “good” quality does not always directly lead to uptake. Or we can say that the standard of “good” and “poor” quality has to be redefined to be more fitting to the needs of central Europeans using English as a *lingua franca*. Let us now turn to the case of teachers teaching what they perhaps do not know or have mastered, which may shed some light on compensation strategies, as well as on the effect of the model.

#### **2.2.4 Can we teach what we don't know?**

Could it be that the term ‘teaching’ involves getting one’s students to learn and that what they learn does not need to be in the subject-matter knowledge of the teacher? Teachers as learner models, not as linguistic models, would go more in the direction of a modern ‘didactic’ and is what is often promoted in teacher training – the teacher as facilitator of learning. However, one might not want to take Zumba lessons from a teacher who cannot properly model the steps.

Murphey (2003) states clearly from his experiences:

Many of the students commented that they felt more relaxed trying new things afterwards because they saw I was trying to learn something new. I was making mistakes, and yet I persisted. In actuality, they were modelling

something much more important than a skill; they were modelling an attitude. For them, my incompetence ... made me a near peer role model. As they were teaching me and their friends, they said that they concentrated more on their own form and began to notice more clearly some of the things they were doing. Several said that they had learned more from teaching and co-learning than they had from simply being 'directed'. (p. 1)

Being active in learning a subject in more depth but not necessarily becoming an expert in that subject might be the right combination. The textbooks can guide teacher learning, too, as long as a teacher is willing and self-critical enough to analyze his or her behaviors and take certain risks. Furthermore, again, it is perhaps this risk-taking or verbalizing or modelling of strategies that is noticed by the learners and which makes certain teachers effective in the classroom despite their language skills.

Huston's 2009 book, *Teaching What You Don't Know*, is devoted to helping but also reassuring novice instructors that one does not have to master a subject in order to teach it. She emphasizes the advantages of not mastering one's subject-matter, or being a content-novice, in that time allocated to a class for certain tasks is often better estimated by novices than experts. Furthermore, content-novice teachers often are more aware of aims and seeing if learners actually reach them than content-expert teachers who may well cover all the content and more, but may not be aware of whether or not their learners have reached the goals. These points are very similar to those made about non-native teachers in general.

So what can be learned by this research? There are distinct advantages of non-native teachers and there are also concerns that are not fully grounded. There is also agreement that stronger language skills can be related to more self-confidence in the classroom and respect by various school partners. Non-native teachers do not have a negative effect on acquisition by the mere fact of being non-natives and can emphasize language features based on standard models that are to be learned explicitly. Furthermore, they may be more capable of teaching certain skills such as grammar. By having learnt the target language themselves, they have seen a process that those having learned the language without any hurdles have not.

In conclusion, there is also a definite lack of research on many aspects of the role of non-native teachers. Primarily, the definition of "highly proficient" is disputed and often subject to political needs and pressures. Secondly, there is very little

research on the role of imperfect modelling. Perhaps in other subjects there is the concept of teachers facilitating learning without being experts, letting children discover various phenomena without the teacher him or herself having mastered the concepts, but in foreign language teaching, the teacher is often viewed as the expert and not so much the facilitator and learner miscues are often attributed to the faulty model of the teacher. Finally, there is no research linking the impact of non-native teachers' language abilities to learner performance, which is a near impossible task, but worth an attempt. Thus, Section 2.3 will take a step back from the language teacher and look more generally at teachers' skills as defined through qualifications and various types of knowledge in the hopes that what we can learn can be transferred and refined with non-native teachers in mind.

### 2.3 Different types of teachers, different types of knowledge

Section 2.2 looked into the role of non-native teachers in the classroom with a focus on attempting to define ‘highly proficient’ for a Swiss context, the role of the teacher as a learner and imperfect modeling. Another way of looking at this is through the eyes of teacher content knowledge (or how well do teachers need to know the subjects they teach and which aspects of the subject do they need to know?) and teacher pedagogical knowledge (or how well do teachers need to know how to teach and does this differ for different subjects?). For non-native speakers teaching English there is a range of levels of knowledge of and about language and understandings of what is entailed in needing adequate levels to teach. This will be elaborated upon in Section 2.3 through a general discussion of teacher content and pedagogical knowledge and manifestations of such, with examples from various school subjects. The section weaves in and out of the context of language and foreign language teaching as some of the findings from the general literature can be projected onto foreign language teaching.

**Clarification of terms.** First of all, there is much discussion in the literature about what knowledge base a teacher must have to be a good teacher. Again, references to “good” or “effective” teaching are elusive and what this means is very difficult to define. Often a definition is neglected. Therefore, most of the research here will attempt to reference studies that link good teaching to something demonstrated through student outcomes, or performance, to teacher attributes recognized as effective, but not as subjective as a learner’s judgment of liking or disliking a teacher or teaching style or as a long term effect, as in older students’ reminiscences of good teachers.

Secondly, here English language knowledge and knowledge of any given school subject matter will be referred to as “content knowledge” and occasionally, due to other researchers’ use of the term, as “subject-matter knowledge” though the research (see Kleickmann et al., 2012) does tend to use this term to imply a deep enough knowledge of the subject to be able to explain reasons behind or for the content point. Concerning English, this might be too judgmental; precisely what we want to find out would be what is enough subject-matter knowledge to be able to teach, with or without a deep understanding of the linguistic systems behind the

language. For simplification, then, “content knowledge” will be used interchangeably with “subject-matter knowledge”.

Thirdly, the concept of “knowledge about the subject” as especially when it pertains to language teaching as language awareness, will be mentioned though it is not concretely measured in this study. Andrews (2003) makes a distinction about language knowledge. He states:

This is not to say that teachers’ language knowledge in these contexts has not been seen as a cause for concern, but such concerns have generally tended to focus on the language proficiency of teachers (i.e. their knowledge of language) rather than their subject-matter knowledge (or knowledge about language). (p. 84)

Here, subject-matter knowledge equates knowledge about the subject which is precisely what this discussion will attempt to keep separate. Although the distinction between knowledge *of* or *about* the subject would be fascinating to research, it is not practical given the requirements set in many countries to be able to teach a foreign language in the elementary school. In many central European countries, pre-service and in-service teachers must demonstrate knowledge of English through a standardized test (Enever, 2014) and knowledge of their mother tongue through an internal university test. These measures, here, represent subject-matter knowledge. Testing a teacher’s knowledge *about* language would go back to their mother tongue and the other languages they know, and which may be a key factor in general subject knowledge but that is beyond the scope of this study. So though Andrews’ description defines “subject-matter knowledge” as knowledge *about* language, for all practical purposes in this study, subject-matter knowledge is simply proof of skills in a language as a representation of knowledge *of* this language and further mention of knowledge about language will use the preposition ‘about’ or the term ‘language awareness’.

Fourthly, neither subject matter knowledge nor content knowledge insinuates any ability to teach the subject, though it is unclear if subject matter knowledge is a prerequisite as Llorca (2005) alludes. The ability to create meaningful educational settings (or teach) can be referred to as “pedagogical content knowledge” as first described by Shulman in 1987. Pedagogical skills are hinted at in this study through the discussion on a teacher’s choice of methods in teaching reading, writing, speaking and listening and amount of time spent using German in the classroom. A

choice of methods does not necessarily imply a deeper or less deep knowledge of teaching. This is a judgment call – if a teacher is more traditional in the sense that s/he creates settings where grammar is quite explicitly taught and neglects speaking and listening (hinting at a more communicative approach), then some might say this demonstrates a lack of modern pedagogical content knowledge. This study will not delve into aspects of pedagogical content knowledge related to classroom management or curricular aim setting, but rather focus on basic breakdowns of language skill teaching.

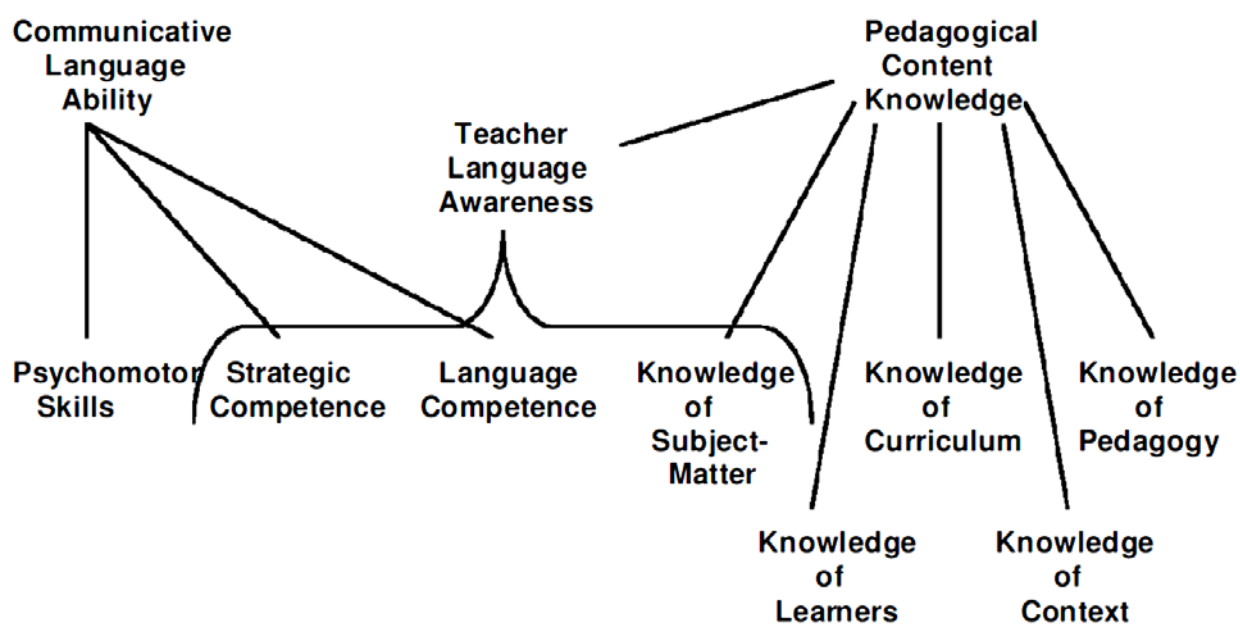
Thus, in the Section, 2.3.1, some general models of language teacher knowledge, which are combinations of content knowledge and pedagogical content knowledge, among other points, will be reviewed. Then, in Section 2.3.2, content knowledge in general and from other subject areas will be discussed. This includes a bit on teacher qualification, which is at times inseparable from content-knowledge in the literature, as it shows proof of deeper knowledge in a specific field. Section 2.3.2 opens up to other subjects but ends by examining content knowledge and language teaching more specifically. Section 2.3.3 then reviews pedagogical content knowledge at a language-related angle, leads back into foreign language instruction and concludes with some discussion of instructional design in language teaching as an expression of pedagogical content knowledge.



### 2.3.1 General models of teacher knowledge from foreign language teaching

Andrews (2001, p. 10) (Figure 1) provides a model of foreign language teacher language awareness, language proficiency, and pedagogical content knowledge. He does not distinguish between non-native speaking or native speaking teachers and he implies that this model is also useful for local language arts instruction - an indication that much of what is useful for the specific case of foreign language teaching may be useful for language teaching in general and vice versa.

Figure 1: Andrews' 2001 model of teacher knowledge (p.10)



Although Andrews has a newer model (2003), the reasoning behind the changes are slightly unclear, thus the 2001 model is more comprehensible in the context at hand. Most universities of teacher education in Europe have some quality control of the elements in the Andrews' model for teacher licensure. Swiss elementary school teachers, for example, know the context and they know their learners because they are most often the classroom teachers. They know the curriculum and they are tested in their teaching skills, at least on a theoretical level. However, for what is behind "Teacher Language Awareness", this might be disputed. In the Andrews' model, the level of teacher language awareness is a fairly all-encompassing umbrella term for other factors and is at times understood as grammar knowledge; thus it is a term that remains slightly unclear in the definition. Swiss teachers are required to know the foreign language they teach; whether this is "language

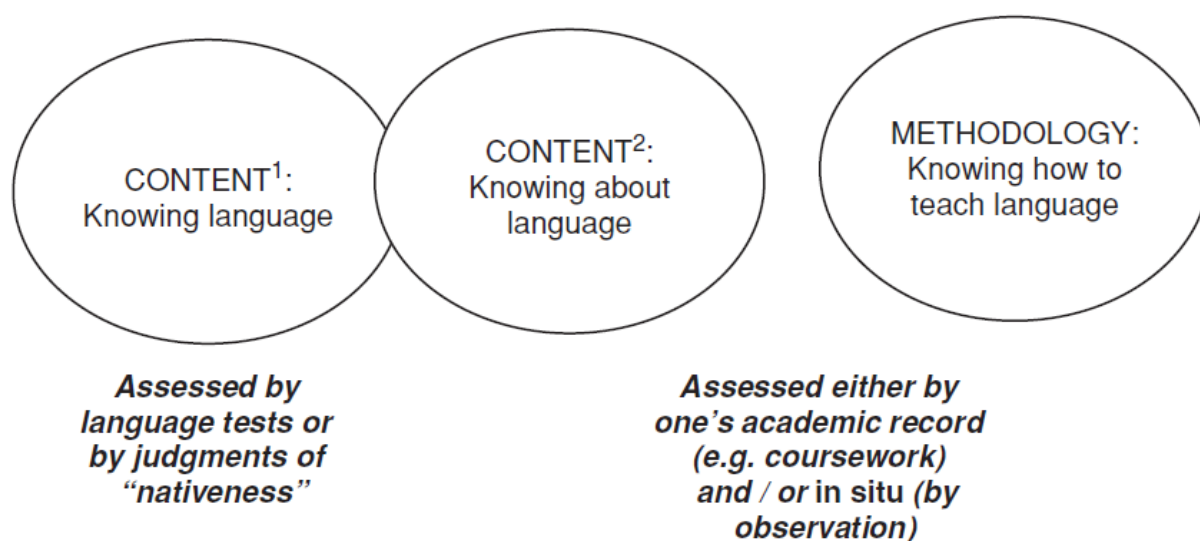
competence” or “knowledge of subject matter” in the Andrews’ model is open to interpretation. Strategic competence in regards to English language teaching specifically is not currently measured for licensure.

In this current study, the factors that get the closest at representing teacher language awareness are target language proficiency and a teacher’s feelings of learning as indications of subject matter knowledge. Pedagogical content knowledge refers to the teachers’ balance of skills teaching in the classroom. These are small parts of a “knowledge base” (Andrews, 2001) of a teacher, but they are ones that are perhaps less subjective than others.

Freeman, Orzulak and Morissey (2009) also present two models. Though simpler, they are perhaps just as rich as the Andrews model. Freeman, Orzulak and Morissey discuss the difficulties in assessing foreign language teachers and thus the difficulties in separating subject knowledge and content knowledge as

...methodology is delivered in language, if the language of delivery is the language that the students are learning, then methodology becomes content and vice versa. This is the interrelationship between we have called content<sup>1</sup> (or medium of instruction) and content<sup>2</sup> (or subject matter). (p. 85)

Figure 2: Freeman, Orzulak and Morissey (2009, p. 83) – Elaborated frame



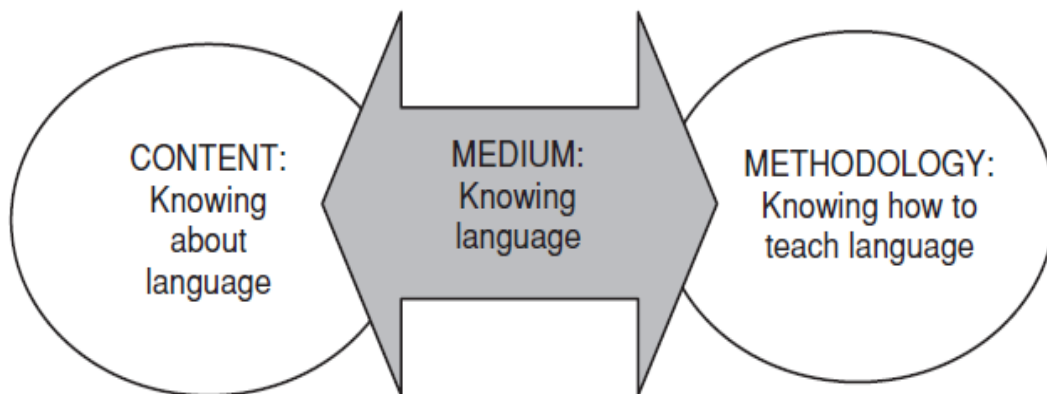
This is shown in Figure 2 and is relatively clearly the model used at many universities of teacher education where the foreign language is tested through a standardized international exam, where methodology in foreign language teaching is

measured through general pedagogy and language specific teaching skills via observations or exams and knowledge about language is fairly neglected.

Freeman, Orzulak and Morissey (2009) go on to produce a second model, an emergent model (Figure 3), in stating that:

This emergent framework also offers a new and useful lens for conceptualizing assessment in second language teacher education. In this view, knowing a language is a medium that interacts with both the content of knowing about the language and with methodology, or knowing how to teach it. And methodology is a dynamic process of interacting with what students know and do. Since the relationships among these three domains is neither sequential nor cumulative, they cannot be logically separated for the purposes of assessment, as is done currently, and for that matter in teacher education. Rather, assessment of language knowledge for / in teaching is likely to become an increasingly messy and emergent process, particularly as the stakes of such judgments are increasing. (pp. 86-87)

*Figure 3: Freeman, Orzulak and Morissey (2009, p. 86) – Emergent frame*



**Figure 3** Emergent view – Language knowledge for / in teaching

This model also reflects the newest changes in qualification rubrics (described in Table 2) at the Zurich University of Teacher Education, where language is recognized as playing an intermediary role. The role of content – knowing about language - is something that could theoretically be measured in a learner's mother tongue or an invented language, though it is not an official construct for certification. It is probable that that the range of “knowing about language” varies quite dramatically amongst European teachers because while they all teach the language of their linguistic area (they are all language teachers in this sense), some are generalist teachers and teach everything and others are relatively specialized in that they teach a range of selected subjects (some are more apt towards music, others sport, others arts, some two foreign languages). In regions where teachers are specialist teachers, one can hope or even assume that a foreign language teacher's knowledge about language in general is high.

These models provide a basic understanding of the knowledge behind foreign language teaching, but each of the elements within the specific models can be interpreted with many different constructs behind them. The following sections will delve into some of these constructs by looking at what effect teacher knowledge and qualification have on learner performance, as policies are set for the amount of subject-specific knowledge or skills teachers must have to teach English and other subjects in the elementary school. By looking at this generally, we might find some language-specific variables, which will be included in Section 2.3.2.3.

### 2.3.2 Teacher qualification and content knowledge

Teacher qualification is often seen as a means of measuring teacher content knowledge. But how much and what types of knowledge (or qualification) have an impact on learner performance? From another angle, in a time of accountability, especially in the United States, teachers are always under scrutiny for how well their learners perform on standardized tests in comparison to other teachers. Provocatively put, good teachers are those whose students perform well on standardized tests or, with a more complicated algorithm, are those whose students have made progress as measured by exams. Teachers are compared to teachers, districts to districts, states to states but isolating the factors which cause differences is essentially impossible. There are myriad studies (e.g. Andrew, Cobb, & Giampietro., 2005; Clotfelter, Ladd, & Vigdor, 2007; Darling-Hammond, Holtzman, Gatlin, & Heilig, 2005) that describe attempts to measure or define teacher effectiveness via tests. There are simple models that simply correlate and compare standardized test scores by teacher by grade and there are more complex measures such as the Tennessee Value Added Assessment System (see Kupermintz, 2003 for a description) that factor out individual school district / child variables such as socioeconomic background and student aptitude, among others.

In Switzerland, there are few standardized measures for which to compare class performance and teacher skills or knowledge across and within grade levels, especially at the elementary school level and in English, though von Ow, Husfeldt, and Bader-Lehmann (2012) do go in this direction. Though it is positive in some ways that the Swiss are not constantly comparing schools and classes – the Swiss value the freedom of the teacher, the creativity to teach learners to learn and not to the test, and trust the educational system – this attitude, however, does not provide much data about effective teaching from an achievement point of view.

Thus teacher content knowledge is something extremely tricky to measure, and here we want “teacher content knowledge” to relate to “teacher effectiveness”, though it is not “effectiveness” on a whole-school, whole curriculum level, but rather “effectiveness” in the [English language] classroom for the sake of defining what characteristics of a teacher’s language competence and contact with the target subject (here English) outside the classroom encourage performance. These findings, then, will hopefully be worth developing in teacher training with the aim of

knowing what factors will benefit class performance – indications of tendencies that make language teaching more effective.

### **2.3.2.1 Teacher qualification**

Huang and Moon (2009) quite bluntly state: "...a highly qualified teacher was not necessarily a highly effective teacher. Certified and uncertified teachers as well as teachers with master's degrees or bachelor's degrees performed no differently from each other in terms of raising student achievement" (p. 231). In times of teacher-shortages, many seemingly "un[der]qualified" individuals are called upon to take up a teaching career. There are studies (as in Huang & Moon, 2009) that show class performance (based on standardized test results) does not change with a teacher's years of experience or degree of certification. Studies by Jacob and Lefgren (2004b) for example, indicate that money for teacher education (thus more on-the-job training) does not improve class performance. Other studies (e.g. Hanushek, 1986) show that other individual factors of the children such as socioeconomic status play a more important role as predictors of performance than money spent on teacher qualification<sup>18</sup>. Arguments such as those by Darling-Hammond, Berry and Thoreson (2001) also plant the seed that teacher certification should become more differentiated and provide a more in-depth look about what types of knowledge are important for certification in that.

A responsible research and policy agenda that builds on the evidence currently available about teacher education and certification should aim to illuminate more fully the specific aspects of teachers' knowledge and skills that make a difference for student learning and the ways in which the features of different teacher education models-how they organize the acquisition of content and teaching knowledge and build knowledge about practice as it is applied - are related to different teaching outcomes. (Darling-Hammond, Berry and Thoreson, 2001, p. 71)

---

<sup>18</sup> This is stated rather simply here, and Hanushek's 1986 work has been considerably discussed and refined and the findings have been put in question due to the methodology, though to certain extent the original conclusions have only been refined and made more precise to the effect that only the funding of certain programs (and thus the teacher education related to it) leads to positive results (e.g. Hanushek 1997).

In their ground-breaking but highly controversial work “Evaluating the Effect of Teacher Degree Level on Educational Performance”, Goldhaber and Brewer (1996) stirred up the field with their findings. This study was one of the first to have the data to link individual students to classes and teachers and also have data on teachers. This study has been highly criticized, namely due to the small sample size, when the data were gathered and that the findings are often oversimplified. However, few other studies have a more differentiated take on this question which can lead us to a better understanding of the question of how teacher qualifications can influence learner achievement.

Many of these studies put into question the effectiveness of teacher certification and undoubtedly, certification does not equate expertise or knowledge but it does give an indication of motivation or interest in a particular field. A more differentiated approach to certification would help educational policies to become more focused on the learners and learning.

Goldhaber and Brewer (1996) found that the higher a teacher’s qualifications (holding a BA or MA) in math or science were, then the better the class performed in those same subjects. For language arts and history, this was not the case – teacher qualification only played a role in science and mathematics. As to be expected, family factors (e.g. educational level of the parents, socioeconomic status) were important predictor variables across all the groups, but once these were accounted for, the level of subject-specific education the teacher remained important for mathematics and science teaching. In their general models, not accounting for subject-specific interplays, they find:

The years of teaching experience variable is not statistically significant in any subject area, nor is it statistically significant whether the teacher has an MA degree. This implies that teachers with an MA degree are no more (or less) effective than those without advanced degrees, clearly a counterintuitive finding. The results for teacher certification are similar in that we find the coefficient on teacher certification to be statistically insignificant (except in English, where teacher certification is significant and negative). (Goldhaber and Brewer, 1996, p. 205)

Why would teacher certification in English have a negative relation to learners’ scores? Are teachers with certification here perhaps more linguistically apt or interested in articulating about their subject and thus perhaps end up confusing their

learners by their wordiness, at least to a certain extent? In this case, why would it not be the same finding for teachers with language-related MA degrees?

In a similar type of study from Mexico, Santibañez (2006) found a positive relationship between teacher test scores on both content and methodology tests and student achievement, though this effect was larger in the secondary school than the elementary school. Interestingly, they found no differences in learner achievement from teachers with two year or teachers with four year degrees. Their research findings put teacher policy decisions into question as subject-matter knowledge and pedagogical content knowledge seemed to be better predictors of learner achievement than degrees or years of experience. For older and younger learners, these attributes can play different roles. In many parts of Europe, there are separate cohorts for primary and secondary school education; in the US, very often the programs are kindergarten through high school though licensure exams are different.

In Palardy and Rumberger's 2008 article entitled "Teacher Effectiveness in First Grade: The Importance of Background Qualifications, Attitudes, and Instructional Practices for Student Learning;" it is shown that teacher qualification does indeed influence reading achievement though not math achievement in first graders. However, the type of qualification was either having or not having a teaching degree, which does not give a detailed picture of the teachers themselves and their other experiences, though it could indicate some importance of state-regulated pedagogical and content knowledge. If "higher and lower qualifications", as some of these studies suggest, were broken down into the type of qualification per se – in a specific subject of teaching, in a subject related to teaching skills, or in something unrelated – there would be a better picture of what level of qualification or type of qualification can perhaps be considered as an acceptable replacement for fully-trained teachers in these times of shortages.

Clotfelter, Ladd and Vigdor (2007) carried out a study looking at the effect of teacher credentials (years' experience, type of degree and licensure) and learner factors (such as parent education levels) on learning performance in math and English. They found that was only in math where elements of credentials positively varied with learner achievement. In English, they were not significant at all, which is in line with other findings (such as Goldhaber and Brewer, 2000).



Some studies in teacher and school effectiveness (as defined by learner performance) use multilevel analyses to look at teacher background qualifications, attitudes and beliefs and instructional practices, thus a combination of factors. Palardy and Rumberger (2008) showed that only full certification was associated with reading achievement (and not with math achievement) and that only specific practices (e.g. explicit phonics instruction) was associated with reading achievement. In a similar study based on the same methodology out of Belgium, Boonen, van Damme, and Onghena (2013) found no effect of teachers' background on reading or spelling skills of first graders, though for math this was not the case. Conversely, teacher instructional practices had a positive effect on reading and spelling achievement of learners but not on mathematics achievement.

**Summary:** These studies compared credentials and often years of experience, and the findings vary. Some say credentials are more important. Many more say experience is the key, so it is difficult to draw any conclusions. For this research, all teachers have similar qualifications from a certification point of view, but they all have varying degrees of subject-specific knowledge which can be seen as a reflection of qualification.

In the case of English language teaching here in Europe, one way of looking at “qualifications” is from the formal side – taking a course and receiving a specialized certification. On the other hand, “qualification” can also be seen as an ongoing process. So if we see that teachers with poor language skills (as measured by the CAE) who have regular contact with English actually have learners who perform well, then we can ask ourselves if formal qualification as measured by an exam, or informal qualification as measured by contact (as an indication of seeking out opportunities for improvement) is a better predictor of learner performance.

Even if many studies are inconclusive about the role of qualifications, universities of teacher education are investing in the people who are teaching our children and are providing a certain qualification or level of qualification to them. Knowing if a level of teacher qualification in a certain subject interacts with learner performance is an initial step in untangling the variables and creating better statistical models for uncovering even more. What these studies can tell us in relation to English language teaching in Switzerland is that teachers with high levels of English language skills (non-native speakers having willingly taken an exam at the C2 level, which is above the minimal certification) or interest in an EFL teaching course abroad have sought

out additional qualifications in the subject with the aim of making themselves expert teachers. Therefore, if there is a certain amount of faith in teacher training, we should still be doing our best to set the right focus in teacher education and improving and encouraging qualifications. And another way of looking at this is through the eyes of content knowledge.

### **2.3.2.2 Teacher content knowledge**

In his groundbreaking work from 1987, Shulman was one of the first to mention “content knowledge” as being different from “pedagogical content knowledge” and thus coined the terms that have been used frequently thereafter. Shulman, however, did not describe one as being the basis for the other. Hattie (2003, 2009) more recently has been searching for excellence among teachers, and he has, if one can quantify such a thing, defined teachers as having 30% of the variance in student achievement in terms of the influence of teacher-decided elements in the classroom such as instructional quality, remediation and feedback and teacher style. He does not mention content knowledge though he does add that:

Our argument is that content knowledge is necessary for both experienced and expert teachers, and is thus not a key distinguishing feature. We are not underestimating the importance of content knowledge – it must be present – but it is more pedagogical content knowledge that is important: that is, the way knowledge is used in teaching situations. (Hattie, 2003, p. 10)

So Hattie (2003), as well as others in the literature on non-native speakers, take a certain level of content knowledge as a given and not a factor which determines learner achievement. Thus if it is a “given”, it is still worthwhile finding out to what extent content knowledge, even if it is a minor factor in student achievement, does play a role as it is a factor that can be and does influence selection policies in teacher education through cutoffs, benchmarks and minimal standards.

Teacher effectiveness studies, such as that by Hattie, and others as found through the *Journal of Personnel Evaluation in Education*, focus mainly on teaching techniques, assessment forms, methods – pedagogical knowledge – and not so much on subject matter / content knowledge. It could be viewed, then, that content knowledge is not nearly as important as general knowledge of teaching and the

willingness of the teacher to develop. However, it could also be said that these studies assume that factors other than content knowledge ARE more important, but at the same time there is perhaps a minimal given level of content knowledge which does have an influence in the classroom, though - as seen from the discussion of “high proficiency” – it is hard to describe or quantify. Perhaps content knowledge can even compensate for a lack of some types of pedagogical knowledge as it manifests itself in a certain security in the subject or enthusiasm for teaching the subject (which has a direct effect on how the learners perceive and cooperate with the teacher). Content knowledge could indicate motivation which can lead to better pedagogical skills (though perhaps not necessarily pedagogical knowledge).

**Examples from non-language subjects.** As the research on foreign language teaching is not extensive in the discussion about content knowledge, and there are few concrete measures of what “highly proficient” means, let us thus start by turning to examples from other fields. There has been much debate over the effect teacher knowledge of teaching has on learner achievement, however the question of content knowledge has not been overly researched in fields other than mathematics. There is a certain agreement that the easiest way to improve the mathematics performance of learners is to improve that of the teachers (Hill, Rowan and Ball, 2005). Whether this is true for other subjects, and even in mathematics, is worth investigation, though it must be looked at critically as some interpretations are very short term and identifying the underlying constructs of what it is that make teachers different in terms of manifestations of content knowledge (how they show it, how schools measure it) may be near impossible. It could be that content knowledge is simply an indication of motivation or it could mean that teachers with high levels of content knowledge in a subject simply drill skills into learners with disregard to more open instructional settings. In any case, from these studies, some transfer can be made to the foreign language classroom in trying to narrow down content knowledge into various subjects and definitions.

In *Teacher Math Test Scores: Classroom Teaching Practices and Student Achievement in Kenya* (Oketch, Mutisya, Ngware & Sagwe 2010), teachers took both a pedagogical knowledge test (not administered to the students) and the same mathematics knowledge test as their sixth grade students. Though nothing concrete could be pulled from the pedagogical knowledge test as there was no difference in scores for those teachers from poorly and highly performing schools, “There is a

positive and relatively strong relationship (correlation of 49.52%) between the teacher's [mathematics knowledge] score and grade six school means scores among the bottom schools..." (p. 15). These findings indicate the direct nature of subject-matter knowledge – the higher the teacher's is, the higher the student's is, at least here in the case of schools with low-performing learners.

Carnoy and Arends' "Explaining Mathematics Achievement Gains in Botswana and South Africa" study (2012) looked at teacher subject knowledge and teacher experience and found that teacher subject knowledge played a role in learner achievement but that other aspects of teaching (quality of teaching, instructional forms) were better predictors and that years of teaching experience was not significant. On the other hand, Buddin and Zamarro (2009) showed that neither teacher test scores (mixed subject-specific knowledge and pedagogical knowledge) nor teacher degree (BA or MA) were related to learner performance, but rather experience, as defined by the number of years of teaching, had a positive relation, which speaks for retaining teachers despite a lack of subject-matter or pedagogical knowledge.

So to focus more into the world of language, some of these studies also include results on teachers' verbal abilities or reading skills which can help to point this research in the right direction. The literature, despite being often focused on mother tongue, or language arts instruction, can help identify factors that may be language specific as it seems that much on the findings about qualifications shows a clear difference in pertinent conditions for effective teaching for more technological subjects.

### **2.3.2.3 Language and foreign language content knowledge**

In turning to more language related subjects, regarding reading, Phelps (2009) found that just being a good reader was not enough to be a good reading teacher – that teaching reading entails a specific knowledge of reading and decoding subskills that do not necessarily belong to even good readers. This would be an argument against direct interaction between teacher and learner reading skills and would support a hypothesis that a teacher's general reading performance is not enough to teach reading well. Phelps states "These results indicate that reading teachers can hold a specialized knowledge of reading – knowledge that differs significantly and

substantially from common adult reading and verbal ability” (p. 147). This triggers another idea – that pedagogical knowledge is more than just the content knowledge of how to teach, but requires a specialized form of general content or subject-matter knowledge, as mentioned by Freeman, Orzulak and Morissey (2009).

Metzler and Woessmann (2010) used a Peruvian data set to find effects of teacher performance on a math and language test. They found a much larger effect for math than for language, though teachers with better performance on a reading skills test did have learners who also performed better on the reading skills test than learners with a teacher with lower level reading skills. However, these differences were not significant. They state: “Robustness analyses indicate that the result is unaffected by considering between-subject differences in teaching hours, teaching methods, and student motivation, and by restricting the model to students whose main language is Spanish” (Metzler and Woessmann, 2010, p. 23).

In their 2003 article, “Teacher Characteristics and Student Achievement Gains: A Review,” Wayne and Youngs went meticulously back through some older US studies (no longer accessible) for evidence. Their data controlled for student socioeconomic status, teacher prior achievement (degree held and from which university), and data collected on standardized test scores of a specific teacher’s learners. In one such study, they found “Among the 627 sixth graders in their elementary school sample, students learned less when their teachers scored higher on the NTE (National Teachers Examination) Common Examinations” (p. 98). Similar findings were found for word tests – “6th-grade students learned less when their teachers had higher scores on the word test” (p.99). This and the Goldhaber and Brewer (2000) study indicate an inverse nature of high level language skills on learner performance.

In another instance, although studies by Andrew, Cobb and Giampietro (2005) could not show conclusively that there is a positive relationship between a teacher’s verbal ability and the performance of his/her learners. The findings tend to support that direction and clearly indicate the negative impact of a teacher’s poor verbal ability on their learners. Their definition of verbal ability is stated as such:

The teacher:

- clearly and cogently presents information;
- gives clear explanations;
- helps students put their ideas into words;

- helps students improve their communication skills;
  - helps students understand the meaning of written language;
  - provides apt analogies to assist learning;
  - communicates well with parents both in speech (be “well spoken”) and in writing; and
  - communicates effectively with administrators.
- (Andrew, Cobb and Giampietro, 2005, p. 344)

Their definition is a dynamic mix of pedagogical ability and content knowledge and this ability is certainly perceived differently by each teacher’s unique group of learners and workmates. It is a description of verbal ability that can well also be used for foreign language teaching. This research suggests that teachers who have strong verbal ability have beneficial skills to the classroom and that perhaps there is a threshold for what is simply not “good enough”.

Though the Andrew, Cobb and Giampietro study provides insight into the nature and importance of verbal ability in teaching, it is connected to teaching performance as measured by the teacher’s completion of learning goals and not on the learner’s performances. In a study by Aloe and Becker (2009), the question of the importance of a teacher’s verbal ability on school outcome was explored by going back through some of the initial studies making claims that this verbal ability was indeed a determining factor in learner achievement. They found that "The results from correlational studies and from multiple regression studies are very similar, and the overall conclusion is that the relation of teachers’ verbal ability to school outcomes is extremely weak at best" (p. 621). They suggest that a teacher’s qualifications are more important and admit that this is something that needs extensive further development – which types of qualifications (more subject-knowledge generally or domain-specifically?) for which consequences on teaching? Other studies (e.g. Ehrenberg and Brewer 1994; Ferguson and Ladd, 1996) support the claim that the verbal ability of the teacher has a positive association with student achievement. Carroll (1975) as well found a positive relationship in that: "a test of French language teachers’ speaking skill was found to have significant correlation to students’ achievement in speaking and listening" (in Darling-Hammond, L. (1999), p. 7).

One subfield of teacher content knowledge, teacher language awareness, is a cross between pedagogical and content knowledge, but again, its effect on learners

has not been researched. Teacher language awareness implies knowledge of the target language well enough to have the ability to know what to select to teach, as Andrews (2001, p. 81) suggests:

The major pre-lesson task in which TLA [Teacher Language Awareness] plays a part involves analysing the grammatical area from the learner and learning perspective. TLA affects the teacher's ability to identify the key features of the grammar area for learning and to make them salient within the prepared input. It also affects her ability to specify the most appropriate learning objectives, and to select materials and tasks which are most likely to serve those objectives, ensuring that they are appropriate in terms of the learners' age and previous learning, and that they serve the desired learning outcomes.

So this is a form of content knowledge that non-native speaking teachers of English (as discussed in Section 2.2) often have at their advantage, at least in terms of identifying grammar points and perhaps breaking them down in a way learners understand. This language awareness is a necessary ability or knowledge in selecting what to teach. This, however, is not necessarily measured by the qualifying exams in Switzerland as these exams measure language use, not language awareness. The dilemma is again here where and how important is teacher language awareness generally as compared to teacher language competence.

**Summary.** The amount of subject-knowledge a teacher should have and which specific subject knowledge (within every subject are an endless number of subtopics) is a policy decision made by boards of education and universities of teacher education in setting standards for certification. Whether these decisions are founded in theory as to why they have been chosen remains unclear. That certain skills such as reading or verbal ability on the side of the teacher have been found to both positively and negatively be associated with learner performance is perplexing and puts into question much of what we believe are selling points in good teaching.

If reading tends to figure often into the literature as correlating highly with learner reading scores, then perhaps it is the love of reading or the successful use of a good range of reading test-taking strategies that are important. Generally, if more general language competence measures have a negative influence on learner scores, then it could be a certain "show-off-ishness" or "pedantry" that is the issue at stake, not necessarily verbal ability or language skills of the teacher. If general teacher subject

knowledge is positively linked to learner performance, perhaps it is curiosity or awareness that are the factors at hand. Thus, it could well be that teacher subject knowledge is one factor that is actually a reflection of a teacher quality that has yet to be identified or is unobservable.

Through the examples in Section 2.3, it becomes clear that subject-matter / content knowledge is at times, and in some studies, inseparable from elements of pedagogical content knowledge. Kleickmann et al. (2012) discuss the possibility that subject specific content knowledge must precede pedagogical content knowledge though they argue that a high level of content knowledge does not necessarily lead to pedagogical content knowledge.

Pedagogical content knowledge could well rely on content knowledge because in order to select appropriate elements upon which to focus instruction, a teacher needs to have a sufficient amount of language awareness or awareness about language (though the accuracy of the selected aims may still be put into question). In this regards, a differentiated take on teacher subject-matter knowledge and pedagogical knowledge indicates that a teacher's reading skills and verbal ability may play a role in learner performance but wherein and which specific subskills is unclear. There is little or no literature referring to teachers' knowledge of grammar skills or skills in listening or writing and their direct effects on learner performance though researchers such as Vandergrift and Goh (2011) do go into detail about effective instruction in these regards. The more general studies from the examples in mathematics [and science] have similar conclusions – the higher the teacher's pedagogical knowledge, the better his or her class performs, thus we will now turn to some discussion around pedagogical content knowledge.



### 2.3.3: Pedagogical content knowledge and instructional practices

Is it possible to have pedagogical content knowledge without subject-matter knowledge? Is pedagogical knowledge really something measurable or does it fluctuate too much with trends – nowadays a focus on speaking in foreign language lessons is of utmost importance, yesteryear ‘twas writing? Wouldn’t, then, content knowledge or pure subject-matter knowledge be a more objective measure of knowledge than attaching the “pedagogical” to it?

Knowing the subject matter well and knowing how to teach well certainly sounds ideal for many cases in the foreign language classroom. But this statement could well be disputed in other school subjects that emphasize and allow for more open instructional settings, where learners have a large choice of topics and ways of acquiring information and skills, such as literature or science projects, activities which are less focused on basic skills or specific content-knowledge. But can it be assumed that just knowing the subject matter (here foreign languages) well is not necessarily enough to teach the subject? Or that not knowing the subject well but knowing how to teach well can be enough to teach the subject? However, it could also be that content knowledge is not the deciding factor, so therefore it won’t be significant to learner performance.

Even though some studies described in Section 2.3.1 imply that knowledge of the language and knowledge *about* the language are essential for language teaching – perhaps a teacher’s philosophical take plays a role; teachers with a more “holistic” approaches to teaching need less knowledge of elements such as grammar as identified by, e.g. Thornbury (1999), than teachers who have more analytical approaches to teaching language and break language down into very small parts. In any case, findings about the amount of content knowledge necessary to teach and an indication of a style or philosophy of teaching that enhances student learning would help to refine the policy of which measures of English language competence and teaching skills must be demonstrated to teach effectively. Section 2.3.3.1 presents pedagogical knowledge in a more general nature and Section 2.3.3.2 looks into instructional design in the classroom as an operationalization of pedagogical content knowledge.

### **2.3.3.1 Pedagogical Content Knowledge**

Wright, Horn and Sanders (1997) suggest that teachers have more influence on student achievement than the heterogeneity or the size of a group and that effective teachers are effective regardless of many learner-related variables. Their study, however, does not determine which factors make teachers more or less effective but only that certain teachers within school districts have students who consistently score higher on standardized achievement tests than other teachers regardless of class size and heterogeneity within the group. Certainly pedagogical knowledge is of utmost importance, and content knowledge can be because it represents a certain enthusiasm for the subject which can be positive in age-appropriate doses. However, how does pedagogical content knowledge manifest in age-appropriate doses? How do teachers break down their subject-specific “expertise” into pedagogically sound principles for a specific class of learners?

Furthermore, pedagogical content knowledge can also be a reflection of the times – in the 21<sup>st</sup> century, pedagogical content knowledge implies a slightly more global and constructivist approach than the more controlled and direct approaches of earlier times. Moreover, pedagogical content knowledge can be similar to content knowledge in general, but situational, as in the examples in Andrew, Cobb and Giampetro’s (2005) definition of verbal ability could possibly be interpreted. For example, a teacher’s ability “to help students put their ideas to words” (p. 344) can be seen as a very teacher-centered behavior (“Repeat after me”) or a more constructivist approach (“Oh, yes, say that again with this word or give me an example”) and how this is handled or whether it is assimilated depends on the class dynamic or individual’s reaction to the intervention.

Kourieos (2014) states that in terms of subject-matter knowledge and pedagogical content knowledge in language teaching “both are necessary in order to avoid fragmented language teaching” (p. 1). Shulman’s 1987 and Grossman and Shulman’s 1994 research concerning content knowledge have also been the basis for models of foreign language teaching in pedagogical content knowledge. Richards (1998) adapted this model for the foreign language classroom and defined pedagogical content knowledge as containing six parts. More recently, Moradkhani, Akbari, Ghafar Samar and Kiany (2013) defined pedagogical content knowledge not

as separate elements (though at times connected) but as elements in order of importance. Table 4 provides an overview of these two definitions.

*Table 4: Definitions of pedagogical content knowledge for foreign language teachers*

Richards (1998)	Moradkhani, Akbari, Ghafar Samar and Kiany (2013)
<ul style="list-style-type: none"> <li>• theories of teaching</li> <li>• teaching skills</li> <li>• <b>communication skills and language proficiency</b></li> <li>• <b>subject matter knowledge</b></li> <li>• pedagogical reasoning skills and decision making</li> <li>• contextual knowledge</li> </ul>	<ol style="list-style-type: none"> <li>1. <b>Knowledge of language and related disciplines</b></li> <li>2. Knowledge of ELT theories, skills, and techniques</li> <li>3. Knowledge of context and social relations</li> <li>4. Knowledge of class, time, and learning management</li> <li>5. Knowledge of research and professional development</li> <li>6. Knowledge of practicum</li> <li>7. Knowledge of teachers and assessment</li> <li>8. Knowledge of reflective and critical thinking</li> </ol>

These breakdowns are perhaps useful for a general understanding of pedagogical content knowledge in foreign language teaching, but there is quite a lot of room for interpretation about what is meant by the specific points. And both contain content knowledge as part of pedagogical content knowledge (in bold in Table 4) and this content knowledge ranks first on the Moradkhani, Akbari, Ghafar Samar and Kiany suggestions. They imply that knowledge of language does not necessarily have to be the target language; it could be perhaps of language in general. Conversely, Richards makes mention the role of the proficiency in the target language and subject matter knowledge. These breakdowns are complex and yet vague and leave open the question of what are the measurable effects on learner performance, as measuring each of these points listed or any combinations thereof might be quite difficult. Though these models were made specifically for the case of foreign language teachers, there is nothing overly different than what might be included in models for general teaching.

So the discussion on pedagogical knowledge is not concrete enough – what is important in teaching is a dynamic interplay of the learners' and teachers' backgrounds, interests, abilities and beliefs. Pedagogical knowledge is a term that changes its definition with every single setting; perhaps it would be better coined as “appropriate pedagogical reactions”. Therefore, perhaps some research on

instructional practices will narrow down the field a bit in order to see what general settings are effective in the foreign language classroom.

### ***2.3.3.2 Instructional practices and their effect on achievement***

In their study using hierarchical models in determining teacher predictors of learner performance in reading, Muñoz, Prather, and Stronge (2011) state: “The most important school-related contributing factor to student achievement is the quality of teaching” (p. 268). Pedagogical content knowledge manifests itself in classroom instructional design decision-making and the skills of the teachers to provide quality teaching. It could be said that instructional practices in the English language classroom are a combination of Moradkhani, Akbari, Ghafar Samar and Kiany’s (2013) ideas behind the “knowledge of ELT [English language teaching] theories, skills, and techniques” and “knowledge of class, time, and learning management”.

The research on the actual effects of combinations of pedagogical decisions on the performance of learners through achievement tests or measures of progress is scant. As in previous discussions of definitions of pedagogical knowledge, it is difficult to separate language-teaching-specific practices and general practices. In this regards, Seidel and Shavelson (2007) state that:

...domain-specific learning activities in general have the largest effects on student achievement. Teachers who provide domain-specific learning activities for students, such as mathematics problem solving, using reading or writing strategies – all of which are the most proximal to student learning processes – tend to have the greatest effects on student achievement. ...There is no single instructional practice, however, that can explain all differences in instructional effectiveness. (p. 480)

These ideas support the concept of process-oriented teaching and time on task which would also be supported by Kase and Jensen (2013) in their work with English language learners. Although ideally being taught in one domain should hopefully help students transfer their skills or knowledge to another domain, as in the case of explicit reading strategy instruction in English being applied automatically in a French reading task, this may not necessarily be the case with younger learners as they are still developing such skills.

Furthermore, the teaching skills that need to be demonstrated for certification vary from one country to the next, thus “Given the varied quality of pre- and in-service provision for primary FL teachers currently available in the ELLiE [Early Language Learning in Europe] countries, it is inevitable that classroom practices may not always meet learner needs” (Enever, 2014, p. 237). Thus the complexity of pedagogical content knowledge is clear – what might be good may not be what teachers are trained in, the definitions are open to interpretation and what teachers know about what is effective may contradict how they actually teach.

**Transfers to language teaching from general research.** From the research on content knowledge in general, one transfer that could be made from the Kenyan study in mathematics education mentioned in Section 2.3.2.2 (Oketch, Mutisya, Ngware & Sagwe, 2010) is their analysis of three main types of pedagogical behaviors– the amount of individual work, whole class work and recitation work (“verbal/nonverbal question and answers between the pupils and their teacher, all pupils in the class reading or reporting and individuals reading or reporting ...” p. 16). Here the authors found that the more often the usage of the recitation method, the better the performance was on the math skills test. This would be slightly counter-intuitive to often-encouraged inquiry-based models of instruction such the Initiation/Response/Follow-up communication patterns<sup>19</sup> which have often been promoted by proponents of student-centered learning (see Black and William 2009). On the other hand, it shows that even in mathematics education, in comparison to less interactive forms of teaching (individual and whole class work), more interaction amongst the participants leads to higher learning.

In the Belgian Boonen, van Damme, and Onghena (2013) study, which combined qualification and instructional practice, no effect of teachers' backgrounds on the reading or spelling skills of first graders was found, though for math this was not the case: there were positive effect of teacher experience and negative effects of in-service training on math performance. On the other hand, teacher instructional

---

<sup>19</sup> Communication patterns such as the Initiation/Response/Evaluation (first discussed by Mehan, 1979) are often visualized with a ball. The teacher asks a question and throws the ball to the student. The student answers and throws the ball back to the teacher and communication thus ends. Other, perhaps more challenging interactional structures would be that the students throw the ball to one another, asking questions and expanding ideas, before the ball returns to the teacher, known as Initiation/Response/Follow-Up patterns.

practices had larger effect on reading and spelling achievement of learners than on mathematics achievement.

With regard to spelling achievement, we found the largest effect size in teacher attitudes and practices (combined effect size = 0.27), attributable to three teaching modalities (homogeneous reading groups, individual activities, and peer teaching) and one language-specific activity (story projects). ...The results here suggest that instructional practices have the strongest association with student achievement, except for mathematics. As for reading and spelling, this is an interesting finding since this aspect of teachers is more easily altered throughout teachers' careers via training courses or mentoring than teachers' background. (Boonen, van Damme, and Onghena, 2013, pp.19 - 20)

So although they find in their study that all teacher variables (background, instructional practices and teacher attitude) are generally more important to math teaching than language teaching (as learners get much more exposure to literacy-based activities than math activities outside the classroom), within the field of what is "more or less effective", a combination of language teaching activities conducive to learning become apparent.

Contrary to what might be the expected outcome for language is their finding that storybook projects were negatively associated with learner performance. The more frequently they took place in the classroom, the poorer the learners performed. The other practices (homogeneous reading groups, individual activities and peer teaching) were positively associated with performance. Peer teaching here in the language classroom, whether it is explicit instructional design or just a setting where learners feel comfortable helping one another with content, could indicate a communicative setting, again with room for interaction.

Could it be that teacher qualification and background in math is generally more important to learner performance than teacher qualification in other subjects? Is learner performance more sensitive, then, to instructional practices than teacher qualification in the languages than in math? However, "students learn math mainly at school, and thus math achievement may be more directly affected by teachers, whereas language and especially reading are acquired both in and outside school, and thus the effect of the teacher is smaller" (Nye, Konstantopoulos & Hedges, 2004). This statement may actually speak in favor of the idea that effective language

teaching takes a lot more instructional design than effective mathematics teaching because teachers have to deal with a much larger ranges of abilities due to types of exposures (written or oral, for example) than in mathematics. So effective teaching in languages takes perhaps different considerations and more specialized and class-individualized instructional designs.

The first researchers of effectiveness studies to explicitly address the point about the differences in mathematics teaching and the transfer to foreign language teaching are Hlas and Hlas (2012) in their article “A Review of High-leverage Teaching Practices [HLTP]: Making Connections between Mathematics and Foreign Languages.” Similarly to the discussion on how to define “high levels of proficiency” in Section 2.2.1, they address the issue of how to define “good” in “good teaching”

The relationship between an HLTP and student impact and how this impact is determined becomes a rather thorny issue. Researchers and teacher educator reformers have tended to determine impact based on empirical research; however, the type of data that count as evidence varies considerably. (Hlas and Hlas, 2012, p. 78)

They narrow down their definition slightly by focusing in on studies which relate (in math) to learner performance on some forms of testing and measurable achievement games. Hlas and Hlas then provide examples of models from mathematics teaching, such as ways of leading discussions and scaffolding through questioning techniques, which might be applicable to teaching foreign languages. Though the examples provided are rather teacher guided, they are not overly teacher-centered<sup>20</sup> and permit learners to discover, problem-solve and negotiate meaning together: “teaching through problem solving is a possible HLTP that involves providing students with a problem, rather than a lecture on solutions, as the basis for exploring and learning new content” (Hlas and Hlas, 2012, pp. 87-88). Though some (e.g. Schwerdt and Wuppermann 2011) might encourage the use of lecture-style activities, much more research (e.g. synthesized in Harris and Ó Duibhir, 2011) and with younger learners tends towards encouraging more interactional settings. Thus far, we see the importance of process-based teaching,

---

<sup>20</sup> Teacher-guided here means that the teacher has the role of facilitator, defining steps but learners work rather autonomously as in a task-based approach to language teaching. Teacher centered indicates that every step is in dialogue with the teacher as in a presentation-practice-approach to teaching.

time on task such as engagement in individual activities, interactive settings, homogeneous reading groups, and peer teaching.

**Foreign language teaching.** In the foreign language classroom the findings thus far might indicate that level groups, guided activities (not too open, but not overly controlled) with opportunities for pair work and thus peer correction might be advantageous. It is possible that, in foreign language teaching, it might also be that particular “methods” are more conducive to learning than others as they contain or are understood as promoting some of the elements found above. Is it simply a balanced-skills approach (controlled activities in each of the four skills), a discovery approach (letting learners use examples of language to find rules), a “task-based” approach (where meaningful interactional settings precede explicit instruction or discovery) or a more traditional grammar-focused approach that is enough to decide what is effective or not? Most likely not, as Edelenbos, Johnstone and Kubanek (2006) state

None of the research which has been surveyed should be considered as demonstrating universal proofs which lead directly to particular desirable practices. This is not the fault of the researchers; it has much more to do with the nature of research in our highly diverse, complex, fast-changing and contested field which cannot be as controlled and scientific as is research in various other domains of human investigation. (p.147)

And though a peek into methods may be one way of looking at effectiveness studies, there may be a fairly large discrepancy between what methods learners want, what teachers think is good for learners and also what is shown to be effective in effectiveness studies as Brown (2009) points out. Here, students tended to want more grammar-translation styles of teaching whilst teachers favored a more communicative approach. These findings cannot perhaps be transferred to elementary school students though one should not underestimate the role of what parents expect of teachers which may have an impact on the teaching.

A classic example is a personal situation where a parent asked during a fourth-grade parent evening about where the vocabulary/grammar translation lists were. The teacher went into a long spiel about how we no longer need these things in a more communicative approach to teaching, that language testing concerns reading, writing, speaking and listening and not vocabulary testing, and that the learners are working in learning notebooks. But alas, a few weeks later, these lists were being



sent home. What effect this instructional decision, based on parental expectations, had on the learners (do primary children in this class now think that they absolutely need wordlists with translations to learn a language?) and on the instruction in the classroom remains unclear. The Brown study also makes us think that the mismatch between what learners (or parents) want and what teachers are doing is like a pendulum – now that more communicative practices are in place, learners perhaps expect these, but with the pressure from parents, teachers revert to habits which may or may not be beneficial.

Methods also change from elementary school and secondary school foreign language instruction though if this is for the better or worse is unclear. Llorca and Huguet (2003) state that

Primary school teachers, instead [as compared to secondary school teachers], tend to have a more communicatively oriented teaching philosophy, but suffer from a greater insecurity with regard to their own language skills, and appear strongly attached to the myth of the NS as the ideal teacher. (p. 230)

This citation is supported through comments made by secondary teachers in the canton of Zurich who were given the task of observing elementary school English lessons. Between 2010 and 2012 I taught approximately 100 secondary teachers in an in-service training program to prepare them for having learners who had English per the second grade. Looking through their comments, the focus was clearly on what the learners did not know in terms of grammar and vocabulary. Very, very few comments related to what the learners “can do” in English. Here are just a few examples:

- How many words should they know?
- If the teacher can't get it right, then the kids certainly can't either.
- Don't the pupils learn wrong structures if their texts aren't corrected?
- There are so many mistakes (“so” vs. “such”; “he do” instead of “he does”)
- Why don't they use the 3<sup>rd</sup> person singular correctly?
- Where are kids just practicing (e.g. “simple past”)?

This demonstrates that although the elementary school teachers were perhaps focusing on what the children can do through various reading, writing, speaking, listening exercises in a communicative fashion, the secondary teachers primarily

saw the mistakes and the language in parts as compared to the whole. What's more, primary teachers are in a dilemma about aspects of methodology as perhaps what they feel is important is not emphasized because they are afraid it won't be accepted by secondary teachers, thus they often revert to traditional models of teaching which may or may not be good for elementary school learners.<sup>21</sup>

Hence the findings from this study could shed some light on whether a bit more of this "secondary teacher analytical attitude" is necessary for good learner performance in the skills at the elementary school level or if a more holistic approach is not really detrimental to learner performance. It may or may not be that teachers who are very good at grammar (as defined perhaps by scores on the Use of English part of the required exam) have learners who performed better than the others. On the other hand, if children's performance is the same across groups of teachers with higher and lower language skills, then this insecurity mentioned in Llurda and Huguet (2003) is simply not necessary.

Effective teaching does not only have to be measured by learner achievement, it can be measured by how effective learner's feel it is. In her dissertation, Shono (2004) attempted to make some matches between what teachers deem as efficient instruction and what learners deem as such. One finding that could provide some insight for this study is that:

...students and teachers maintained that in order for learners to acquire English as a target language effectively, they must be given many opportunities to use and practice the language. The teacher participants also agreed that ESL learners in most classrooms do not receive enough speaking time in class and are thus being deprived of necessary, valuable practice and experiences due to time or class size constraints. (p. 166)

So teachers who are aware that learners feel that speaking is important and also believe the same may have learners who have higher levels of achievement. Though the Shono research pertains more to adult learners, anecdotally it could well belong to elementary school learners who do, in classrooms, want to participate and speak.

---

<sup>21</sup> On March 19, 2014 during a four hour course to primary teachers become qualified to teach English, one teacher came to me with the statement that "I give vocabulary translation tests and grammar translation tests because that is what secondary teachers expect me to do".

In another study from Switzerland by Kreis, Williner, and Maeder (2014) differences in activity types and frequencies (e.g. listening to CDs, singing songs) between primary and secondary students were studied. What was found in regards to achievement was

Wer in der Primarschule besser schreibt, gibt häufiger an, dass im Englischunterricht *Grammatik geübt* wird ( $r_s = .27$ ,  $p = .000$ ). Je häufiger die Schülerinnen und Schuler der Sek G angaben, dass im Englischunterricht *Dialoge / Umfragen / Theaterszenen / Rollenspiele* geprobt werden, desto höhere Leistungen erzielten sie im Schreiben ( $r_s = .25$ ,  $p = .043$ ) und Sprechen ( $r_s = .39$ ,  $p = .031$ ). Zudem gaben Sek G Schülerinnen und -Schuler, die kompetenter schreiben können, häufiger an, dass im Unterricht Wortergelernt (vocabulary) werden ( $r_s = .23$ ,  $p = .050$ ). Bei den Schülerinnen und Schülern der Sek E finden sich keine signifikanten Zusammenhänge zwischen der Auftretenshäufigkeit der Aktivitäten und den erzielten Leistungen.<sup>22</sup> (Kreis, Williner, and Maeder, 2014, p. 37)

This indicates the importance of grammar-based activities on writing, but also later, in the secondary, the role of interactive sorts of activities, though there is no indication here of how these dialogues and surveys were actually carried out (in more or less controlled settings). While some research suggests that teachers should encourage more open settings "... teachers' difficulty in moving away from the traditional teacher-fronted position towards a more facilitative role during some phases of the lesson" (Enever, 2014, p. 240), it could be that individual activities and more controlled grammar-based activities also have their place or that open activities such as plays may be taught in a very controlled manner. Of course, if what is being measured is performance in a language, then direct instruction in the constructs being tested will be conducive to performance, but perhaps what learners learn from

---

<sup>22</sup> [Students who write better in elementary school indicate more frequently that grammar is practiced ( $r_s = .27$ ,  $p = .000$ ) in their English lessons. The more frequently secondary school G students find that dialogues, surveys, plays and role plays occur in their English lessons, the better they performed in writing ( $r_s = .25$ ,  $p = .043$ ) and speaking ( $r_s = .39$ ,  $p = .031$ ). In this regards, secondary school G students who can write well admit that vocabulary instruction often takes place in their lessons ( $r_s = .23$ ,  $p = .050$ ). For secondary school E students, there were no significant correlations between the frequency of certain activities and the students' performance. (Kreis, Williner, and Maeder, 2014, p. 37)]

more open settings helps them in the longer-term or the jumps in performance occur at different rates.

**Local language use in the foreign language classroom.** Another way of looking at instructional design is via the amount of input learners receive in the target language which can be seen in a few different lights. Bruhlmann (2012) provides a concise overview of behaviors on both the part of the learners and the teachers, and in the following discussion, the focus will remain on the teacher's use. The American Council on the Teaching of Foreign Languages posted a position statement about the use of the target language in 2010 which encourages teachers to "use the target language as exclusively as possible (90% plus) at all levels of instruction during instructional time and, when feasible, beyond the classroom" (ACTFL, 2010). In the EDK curriculum it is stated that "Der Fremdsprachenunterricht findet grundsätzlich in der Zielsprache statt"<sup>23</sup> (Erziehungsdirektoren-Konferenz Ostschweiz., 2010, p. 7). However, there are no references as to why these recommendations are in place and how the amount 90% came about or what 'grundsätzlich' [basically] means.

In evaluating the literature, first topic to review are the reasons the local language is used. Teachers use the local language for multiple reasons, sometimes pedagogically sound, other times not. As Enever (2014) states:

... there was evidence of teacher anxiety in relation to language competency, a concern that may well also relate to fear of losing control of the class. Whilst the use of L1 has an important role in the first stages of young children's introduction to learning a FL, these teachers tended to overuse it and sometimes failed to strategically plan purposeful use. Their lack of expertise in structuring interaction tasks, in ways that could maximize FL production in both controlled and free practice events, was particularly evident. (p. 240)

And what is pedagogically sound may be a good decision for the local language<sup>24</sup> and a bad decision for the target language (it would have been nice to have heard that structure in the target language and many awareness activities take place

---

<sup>23</sup> [Foreign language teaching should basically take place in the target language.]

<sup>24</sup> Such as hearing a seldom used structure in German like the use of the future in English ("I will do this tomorrow" / "Morgen mache ich es" / "Morgen werde ich es machen")

implicitly, not everything must be explicit) or a good decision for both<sup>25</sup> (e.g. a deeper understanding of how the “conditional” form works in two languages). Commonly, however, teachers are not even aware of how much of the local language they are using and in which situations they are using it in (see Macaro 2001). Their decisions for using the local language are often slightly haphazard.

Secondly, there is the effect of the use of the local language in the classroom. It can be that learners are put more at-ease (Anton & DiCamilla 1998). It could also be that learners become motivated to know how to say things in the target language when they feel like their mother tongue is “safe”. However, the opposite could also be true; they can lose motivation to learn the target language and don’t take risks (see Nation 2003). Furthermore, the effect of the learners using the local language may be quite different than the effect of the teachers using it. When teachers use it, there is less input in the target language and personal teacher observations in Switzerland lead the author to believe that teachers are not systematic about their use of German or that they have strong principles about why they use it. This is thus perhaps an ineffective usage of the local language for learning purposes and less input; in other words, this limits learners’ exposure. Butzkamm (2003) would clearly disagree with this statement, though he might concede that *how* it is being used in the classroom makes the difference, not *that* it is being used.

Research (again see Macaro 2001) may speak in favor of allowing natural codeswitching from the side of the learners, especially younger ones, but explicit mother tongue use (as compared to haphazard) use from the side of the teachers. And it could be that selected use (by the teacher) of the local language for grammar explanations is not detrimental to the performance of learners, but for other reasons (giving instructions, disciplining, and so on) it is as these may make other activities more efficient, but provide less input. Although the opposite was found with beginning adult learners in the Viakinnou-Brinson, Herron, Cole, and Haight (2012) study (learners performed better when grammar instruction was in the target language instead of the local language), other studies indicate the value of linguistic comparisons (e.g. He, 2012; Butzkamm, 2003; and Butzkamm and Caldwell, 2009) for grammar-related discussions. Kreis, Williner, and Maeder (2014) found:

---

<sup>25</sup> Such as learning how role playing in German uses the second conditional “[Nehmen wir an,] “ich wäre jetzt der Grossvater” but in English uses the future “I’ll be the grandpa” or recognizing prefixes, such as that ‘vor’ in German is ‘pre’ in English).

Interessant ist dabei, dass die Ausprägungen der Skala ‚Unterrichtssprache‘ und die Ergebnisse der Primarklassen zum Leseverständnis ( $r_s=.22$ ,  $p=.001$ ), Hörverständnis ( $r_s=.24$ ,  $p=.001$ ), Schreiben ( $r_s=.27$ ,  $p=.000$ ) und Sprechen ( $r_s=.29$ ,  $p=.015$ ) signifikant zusammenhängen. Je häufiger damit gemäss den Einschätzungen der Primarschülerinnen und –schüler im Englischunterricht überwiegend Englisch gesprochen wird, desto höhere Kompetenzniveaus werden erreicht.<sup>26</sup> (p. 38)

In conclusion, the less input one receives in a target language, in general, the less the potential for learning the target language. Secondly, the decision for teachers to use the local language may be for reasons related to specific instructional goals, or simply due to habits. In the case of instructional goals, using the mother tongue to support metalinguistic awareness can be beneficial, and may help performance in both the mother tongue and the target language. In the case of “haphazardness”, it may have negative consequences on performance in the target language.

**Local instructional practices.** Switzerland has a take on instructional design in the foreign language classroom that is different from other countries as teachers are obliged to use certain textbook series’ and are not allowed to freely choose others, though they certainly choose their own supplemental materials. As an example, in 2010, a letter went out to Zurich local Board of Education members stating:

In Absprache mit den Präsidenten des Verbandes Zürcher Schulpräsidentinnen und Schulpräsidenten (VZS), des Verbandes der Schulleiterinnen und Schulleiter des Kantons Zürich (VSLZH), und der Lehrpersonenkonferenz der Volksschulen des Kantons Zürich (LKV) möchten wir Sie bitten darauf zu achten, dass an Ihren Schulen die obligatorischen Lehrmittel verwendet werden. Aufgrund zahlreicher Rückmeldungen ist ersichtlich, dass mancherorts Lehrpersonen in ihrem Unterricht nicht die obligatorischen Lehrmittel einsetzen. Dies geschieht

---

<sup>26</sup> [An interesting correlation is that the relationship between the language of instruction scale and the performance in elementary school classes in reading comprehension ( $r_s=.22$ ,  $p=.001$ ), listening comprehension ( $r_s=.24$ ,  $p=.001$ ), speaking ( $r_s=.27$ ,  $p=.000$ ) and writing ( $r_s=.29$ ,  $p=.015$ ) are significantly correlated. The more often the learners indicate that English is being used in their lessons, the higher their performance was. (Kreis, Williner, and Maeder, 2014, p. 38)]

teils in Unkenntnis oder falscher Interpretation der Vorgaben, aber auch vielfach in der Absicht, ein anderes, vermeintlich besseres Lehrmittel einsetzen zu wollen...<sup>27</sup> (Wendelspiess, 2010)

What this means plays out on several levels. First of all, this has the effect of irritating teachers. Whether this takes their joy away from teaching English cannot be said, but it complicates the idea that instructional design is a reflection of the needs of the learners if a textbook is inappropriate and teachers do not adapt it. Secondly, some materials used in lower primary school are poorly developed in terms on listening and writing. So teachers using this series would have to compensate if they believe a balance of skills, or well-developed skills' work, is a good approach. This takes time and energy. Of course teachers are allowed a certain "Methodology Freedom" [Methodenfreiheit], so they can use the provided materials in any way they deem fit. How much they actually adapt the materials based on research indicating what is effective is questionable.

The use of songs and rhymes is encouraged in the Swiss materials and these elements have been shown to be effective as Jarvis (2013) concludes in her research comparing young learners learning through flash card activities and those learning the same language through song and rhyme (and performing better).

Often in the Swiss-German materials, a CLIL (Content and Integrated Language Learning) approach is suggested (in the First Choice, Explorers and parts of the Young World series). However, though it has been quite the hype over the last years, Bruton (2010) states "the research conducted here does not show conclusively that CLIL is either positive or negative for FL development, in this case mainly English" (p. 5). And the most commonly used textbooks in Swiss-German schools are topic-based, not content-based although they do market the CLIL approach.

In the Explorers and Young World series, a "task-based" approach is provided. The understanding of "task-based" in both series is interpreted quite differently and

---

<sup>27</sup> [In accordance with the heads of the Zurich School Presidents Union, the Zurich Cantonal Principals' Union and the Canton of Zurich Public School Teacher Representatives, we kindly ask you to ensure that the compulsory textbooks are being used in your school. Due to numerous responses, it has become clear that in some areas teachers are not using the compulsory books. This can happen out of ignorance due to being unaware or incorrectly interpreting the guidelines, but also because of wanting to use another, seemingly better textbook...(Wendelspiess, 2010)]

as in the research on CLIL, the term “task-based” can be understood to be many different processes with many different consequences on learner performance; thus it is not worth the effort in trying to narrow down the definition to one understanding. Rather, by looking at more general teacher behaviors, we have a slightly more precise understanding of what is effective than by trying to define a series of steps and looking at them as a “whole”.

Instructional practices are complex because there are also characteristics of the teacher that cannot be accounted for, as in the example of the teacher’s accent. It could be that a teacher has a seemingly good understanding of effective instructional design for his or her specific group of learners but has an accent that the learners do not react well to, be it a strong local, non-native or a strong native accent. Thus there are perhaps some factors that get in the way or others that help instructional design, though design can compensate for these things as well.

So there are general teaching practices, philosophies and practices to which teachers adhere and which reflect their understanding of what is good teaching. Within this there is the teacher’s belief that this is “the way to do it” and then the match or mismatch to the learners’ conviction of the teacher’s way and their own ideas of what might be relevant. On top of this is the influence of the parents, especially with younger learners. And then there are all the teacher individual factors which influence how his or her “pedagogical decisions” are received – in foreign languages this may well be accent, willingness to share experiences, willingness to admit that s/he does not know something or attitude towards the target language, to name a few.

Within these larger teaching decisions and the individual teacher’s personal characteristics, there are micro decisions about how to go about an activity – are activities just done or are they guided step by step? In the research on listening, for example, metacognitive models as proposed by Vandergrift and Goh (2011) are rarely seen in primary classrooms, though teachers who have the knowledge of certain such of language teaching and are able to structure lessons around such knowledge may be more effective than those lacking the knowledge or the ability to implement meaningful models.

The discussion on the use of the local language can influence effectiveness – some teachers may use it efficiently and others not. The use of the local language in teaching can theoretically be used at the cost of performance in the target language



but to the advantage of knowledge about language, and the local language, in general. But are teachers who use the local language in their instruction actually doing so in ways that support this development?

That said, generally the less the use of the local language, the more involvement of the learners be it controlled or open, perhaps the better the performance of the learners in English. The more time spent on specific language skills' development – time on task – the more time for learning and developing reading, writing, speaking and listening skills.

Every instructional decision can be sold as effective – effective use of storybooks, effective use of the mother tongue, effective metacognitive instruction – but shown to be ineffective when looking at the details in implementation or interactions with certain teacher characteristics. And there is a lot more in instructional design that could be looked at such as forms of corrective feedback or approaches to grammar teaching. Thus, it is difficult to draw conclusions about the right combinations of classroom situations that properly foster learning, and fostering learning is not the same as performing on a test, as will be the evaluation tool used in this study. There are too many variables on the part of the teacher and on the part of the learners which influence effectiveness.

Section 2.3 looked at teacher knowledge from many different angles. What measures successful teaching is difficult to define – is it what is represented by a qualification (motivation in the field) or what is specifically done in the classroom? Is it pedagogical knowledge or content knowledge? Is it general knowledge or language specific knowledge? Is it the effect on the learners in their own content knowledge or their future liking and motivation in a certain field? Generally, from some of this research, it could be that what is commonly encouraged in today's universities of teacher education may need to be looked at quite differently in the future.

## 2.4 Consolidation of the literature and open questions

Despite the numerous suggestions (Llurda, 2006, among others) that the language proficiency of the non-native teacher is of utmost importance, there have been few attempts to define this level and in my experience in observations of teachers, this is not the key issue in teaching effectiveness. In terms of specific language skills, if Phelps (2009) questions the difference between the “lay” reader and the “teacher” reader, then why are teachers required to provide a “lay” proof of language skills, and, in my opinion, not required by boards of education to show more in-depth proof of their teaching-necessary language skills (which are as of yet ill defined)? Although teacher knowledge does play a role, it is perhaps knowledge about language (Andrews 2006) or pedagogical content knowledge in language (as compared to math) that may be more essential. And even more importantly perhaps, as Murphey (2003) might agree, is the teacher’s feeling of learning. Based on my observations, teachers who admit to not knowing everything and who voluntarily say they’ll “find out” are more credible, even amongst young learners. Finally, time on task leads to performance in the same task as is indicated by research with non-native teachers (Kase and Jensen, 2013) and also in general (Boonen, van Damme, and Onghena, 2013).

It appears, however, there is a lack of research coming from the elementary school setting and from the foreign language setting. The mass of literature from the US has to do with the fact that in the US there are standardized test data from both teachers and learners which can be used. Although this may be problematic, it has been useful in formulating hypotheses in this study to have profited and been given some guidelines and tendencies. One further point missing in this literature is that there was almost no mention of informal qualification, such as the role of math or nature related hobbies or foreign language learning outside of the classroom as an indication of qualification and sustained teacher learning. Furthermore, in some studies, such as Darling-Hammond, Holtzman, Gatlin and Heilig (2005), there have been quite controversial discussions about the methods and models used to attribute factors such as teacher knowledge to learner outcomes. The mere number of variables, the varying degrees of controlling for other factors such as socioeconomic status, spending per child with a district, etc., have provided many a provocative concept but have also made analyses extremely difficult.

The research exploring teacher subject matter knowledge is full of disparity but can lead to certain questions about language teaching. If verbal skills are not as important as previously estimated, then which skills are? If qualifications such as holding a degree in a specific field are more important than subject matter knowledge in that field then in terms of the specific case of foreign language teaching, which sorts of qualifications are necessary? If general teaching skills are more important than pure language skills, then why are language skills a large part of initial teacher education and also the point that seems to cause the most discern for non-native speakers of English? If some studies (Hill, Rowan, & Ball, 2001; Oketch, Mutisya, Ngware & Sagwe, 2010) indicate that a teacher's math knowledge is more indicative of learner outcomes in math than a teacher's language knowledge in the local language, then can it be assumed that these results would be the same for teacher language ability in foreign language teaching?

Finally, from the Murphey (2003) research, we can agree that teachers who see themselves as learning with students contributes positively to learning, and this represents an attitude (being on an upward trend) that is important. So although no conclusions can be drawn about the definition of "highly proficient", enough other literature suggests the advantages of the non-native teacher and the relative unimportance of the verbal model and emphasizes the instructional setting and atmosphere.

The present study looks at various elements of teacher proficiency, contact with English and time on skills in the classroom and their associations with learner performance. The role of language proficiency (or lack thereof) and a teacher's feeling of proficiency since it was measured can be an indication of subject matter knowledge; contact with English outside the classroom and time abroad can represent indications of proficiency and qualification.

In short, there are four main questions that will be treated in this study and one mixed question.

**Teacher language competence:** Do teachers with higher levels of English have learners whose competence in reading, writing and listening are also higher than those learners in classrooms with teachers having lower levels of language competence?

**Teacher feeling of language learning:** Do teachers with the feeling that they are on an upward learning curve have learners who perform better in reading, writing

and listening than those learners in classrooms where the teacher feels to be on a downward or stagnated learning curve?

**Teacher contact with English outside of the classroom:** Do learners perform better if their teacher has had an extended stay in an English speaking country or regular contact outside the classroom with the English language?

**Teaching balance of time on skills:** Does the estimated amount of time a teacher spends teaching reading, writing, speaking and listening in English interact directly with the learners' performance in each of these same skills? Does the estimated amount of time a teacher speaks German in class detract from learners' performance in reading, writing and listening?

**Combination (multivariable) question:** Which combinations of teacher-level variables (language proficiency, feeling of improvement, contact outside of school, time teaching certain skills) are associated with learner performance on a test of reading, writing and listening?

## 2.5 Working hypotheses

These hypotheses are based on the review of the literature as well as on my field observations having taken place between 2002-2014.

### 2.5.1 Teacher proficiency level

**Question:** Do teachers with higher levels of English have learners whose competence in reading, writing and listening are also higher than those learners in classrooms with teachers having lower levels of language competence?

**Hypotheses:** The Cambridge exams (FCE, CAE, and CPE) taken by the teachers all contain five parts: Use of English; Speaking; Writing; Listening and Reading. Each of these parts will react differently with learner scores and subscores. A positive correlation between the Use of English and the Reading parts of the Cambridge exams with learner performance is expected.

*Reading.* Though the CAE does measure reading and listening not just for general information but for inferring, comparing, synthesizing, and more, the results of these specific skills tested are not available in the results the teachers receive. Since it is impossible to get a more detailed picture of reading subskills, the hypothesis (assumed from Phelps 2009) might have to be that reading skills do not interact with learner proficiency in reading as here we only have the total reading score and not those scores of reading subskills. The Phelps study might be grounds enough for excluding reading scores as a significant variable; however, the CAE reading test does have some components of grammar knowledge in it, and some of the reading subtests require high level analytical skills, which might play a role that is implied in teacher knowledge about language (Andrews 2003). Thus the working hypothesis here is that reading will be positively associated with learner performance on the total test scores and subtest scores.

*Use of English.* Teacher language awareness is not overtly measured in the Cambridge Exams – the speaking, writing and listening subtests generally measure knowledge of language. However, the Use of English and reading are slightly less holistic and more about language analyses thus these bits are most likely more important or interact more highly with learner performance on all subtests. The Use of English test also shows how comfortable teachers are at breaking down language into smaller parts, thus this may be reflected in the association, and a skill at which

non-native speaking teachers are adept. For that reason, a positive correlation between the Use of English scores and learner performance on all subtests can be expected.

*Speaking.* The effect of verbal ability on learner performance is often disputed in the literature (Aloe & Becker 2009 or Ehrenberg & Brewer 1995). The working hypothesis here is that a teacher's speaking skills will not have any association with the learners' performance as the role of input from the teacher is perhaps not as important as the role of input learners get in English, even at this young age, from music and other sources as Andrews 2003 implies. Although Chambless (2012) hints that oral proficiency of the teacher may be important because as "it is intuitive to assume that teachers who have not attained a certain proficiency level will be hard pressed to provide a linguistically rich instructional environment that will enable student learning to progress beyond the basics" (p. 142), in this study only the basics are being tested.

*Listening.* A teacher's performance on the listening subtests is not expected to interact with learner performance. Unlike the reading subtest, the listening subtest does not measure the same analytical language skills, thus the listening skills' construct is very general. As listening is a receptive skill and not transmitted to the learners in this context, it will not play a role in learner performance on an English test.

*Writing:* A teacher's performance on the writing subtests is not expected to interact positively or negatively with learner performance. Though writing is a productive skill, teachers in young learners classrooms have to use their writing skills very minimally – board work, wordlists and perhaps some worksheets – thus the writing that the learners are getting from the teacher is significantly less than the writing learners are getting from their textbooks or print media in their surroundings.

### 2.5.2 Teacher learning curve

**Question 2:** Do teachers with the feeling that they are on an upward learning curve have learners who perform better in reading, writing and listening than those learners in classrooms where the teacher feels to be on a downward or stagnated learning curve?

**Hypothesis.** From the Murphey (2013) and Huston (2009) research, we can agree that teachers who see themselves as learning with students is an attribute

positive to teaching, and this represents an attitude (being on an upward trend) that is important. Being aware of one's learning and taking a learning path with one's learners is perhaps one essential key to affect in the classroom and encourages an atmosphere of learning and discovery. Teachers who have not learnt much or who have stagnated may let these feelings show through or perhaps show a lack of curiosity in the language itself. Therefore, it is expected that teachers' general feelings of improvement will correlate positively with learner skills on all the subtests.

### 2.5.3 Teacher exposure to English

**Question 3:** Do learners perform better if their teacher has had an extended stay in an English speaking country or regular contact outside the classroom with the English language?

**Hypotheses:** For some of the same reasons in Question 2, that this contact keeps them on a positive learning curve, teacher regular exposure will interact positively with learner performance. Specific exposure is expected to be associated with specific performance – teachers who read a lot will have learners who perform well on reading, as this probably represents one of the constructs of finding reading important with the teachers emphasizing it in teaching or spending more time on specific tasks. Generally, regular contact represents a certain maintenance, motivation and perhaps enjoyment of the language.

Teacher exposure in terms of length of stay abroad is not expected to have any effect on learner performance, because if this stay was perhaps too far in the past, it is not as relevant as what teachers do regularly, and is quite individual. There are teachers who spent years abroad in native speaking countries or those who spent years abroad in non-native countries where English was the language of communication. There is simply too much variation in individual histories to include the hypothesis that this factor will be significant.

### 2.5.4 Teacher time teaching language skills profile

**Question 4a:** Does the estimated amount of time a teacher spends teaching reading, writing, speaking and listening in English interact directly with the learners' performance in each of these same skills?

**Hypothesis:** Here it is hypothesized that the more time a teacher spends teaching a specific skill the better the effect on learner performance in that same

skill, though not necessarily in other skills. The research by Seidel and Shavelson (2007), Kase and Jensen (2013) as well as von Ow, Husfeldt and Bader-Lehmann (2012) all point in this direction: the more a skill is practiced, the better one becomes at it (explicit focus). There will be a direct relation between time on task and performance in that same skill (reading, writing, and listening).

**Question 4b:** Does the estimated amount of time a teacher speaks German in class detract from learners' performance in reading, writing and listening?

**Hypothesis:** It is expected that the more German used in English lessons, the worse the learners will perform on all subtests. Similarly to the hypothesis in question 4a, the more a teacher uses German in class, the poorer the learners will perform as this represents less time on task in English. Furthermore, it is questionable as to whether not Swiss teachers use German in the English language classroom haphazardly or with real educational purposes and since here only English is measured, and general metalinguistic awareness or knowledge of German is not, nothing can be said about the positive effect of the local language use.

### 2.5.5 Combined / multivariable model

**Question:** Which combinations of variables on the side of the teacher, controlling for the years of English the learner has had, contribute the most to learner performance?

**Hypothesis:** When all the variables related to teacher proficiency, teacher background (years teacher or number of languages spoken), teacher learning curve, teacher exposure, teacher classroom skills profile on the side of the teacher and the number of years of instruction the learners have had are thrown into the regression, it is expected that learners having had two years of English, that the teachers' classroom skills profile (as a measure of instructional design) and use of German as well as teacher regular contact with English will interact strongly with learner performance – all positively except in the case of time spent in German. Other factors, such as the language level of the teachers or their time spent abroad are too questionable in the literature to be strong predictors of learner performance and in their interplay with the other variables, are not expected to be significant.



### **III. Development and preliminary results**

The background information on the Swiss system, standards in English language teacher training and qualification, the role of various types of knowledge and the influence of teachers by nature of qualification, instructional methods and language skills leads us to the study design. In order to be able to make any inferences about the association between teachers' test scores and contact with English on learners' performance, and to be able to collect the necessary information with the least amount of work for the teachers, a survey and a test were developed to provide this information.

In 2010, a pilot test was designed to test learners' performance in reading and listening. These results were then analyzed and the test was revised and improved. The newly developed test was finished in 2011 and used to gather data at the end of the 2011 and 2012 school years. The survey for the teachers was developed in 2010 and used with the final data collection in 2011 and 2012.

In Chapter III, the development of the instruments will be described. Section 3.1 starts with a description of the instrument developed in the preliminary pilot study to measure learners' English language skills. Afterwards, the findings from the learner test are presented. Section 3.2 describes the development of and findings from the teacher survey. Later, Chapter IV provides the methodology behind the inferential part of the study and provides the results in relation to the research questions, the associations found between the learner performance on the developed test and information gathered from the teacher survey. Analyses mentioned in Chapter III were run with SPSS (Statistical Package for the Social Sciences version 20) unless otherwise stated.

### 3.1 Learner subtests

#### 3.1.1 Reading and listening test development

**Development.** The first step in determining what to test was to look at the descriptors in the official EDK curriculum for English in order to see where children should be by the end of the third or fourth grades, after one or two years of English language instruction in Swiss elementary schools. The Common European Framework of Reference (CEFR) descriptors serve as the basis for the performance standards in the official curriculum and were thus used for the context of the subtests. As the EDK curriculum only provides two to five descriptors to represent each achievement level (A1, A2, etc.) the Lingualevel (Lenz, 2007) database was accessed to find more descriptors to represent these levels. Despite the limitations of using the CEFR which Hulstijn (2007) summarizes in his paper “The Shaky Ground Beneath the CEFR,” as teachers and many learners are familiar with the language of the CEFR and with the descriptors themselves, and as these descriptors are the basis for the EDK curricular standards, they were seen as a simple way to get the learners to self-assess their skills and with which to link to the subtests.

As the descriptors in the curriculum demonstrate the minimal standard, and learners are minimally to be at A1.1 by the end of the third grade, higher-level descriptors at the A1.2 and A2.1 levels were used in this study to elicit a higher level of discrimination and wider range of item difficulty. Therefore, A1.2 and A2.1 Lingualevel tests were then selected which were contextually appropriate for eight to ten-year-old children, where the formats varied (not only multiple-choice) and where different individual descriptors found in the official curriculum and matched through the Lingualevel database were evident. Further subtests were created by the author under the same criteria but used to cover different descriptors and test formats. The pilot test was thus a combination of listening and reading subtests from Lingualevel and those created by the author. This original pilot test comprised 5 listening subtests (42 items) and 4 reading subtests (57 items). Five of the sub-tests were taken from Lingualevel and four of the tests were created by the author.

Within the specific reading and listening subtests, questions for gist, detail and inference were used, or as Song (2008) would state, “implicitly” and “explicitly” stated information. Though detail questions (understanding details explicitly stated in the text) were most commonly used, for most of the subtests there were either one or

two inference (understanding implied meaning) and gist (understanding the general content) questions. The three main genres used for the text types were narrative, expository and descriptive as they are most relevant to these levels and to the audience. Hortatory and procedural texts were not used as they did not fit the nature of the tests – these are better used in other assessment settings.

**Pilot testing.** The pilot test was administered in July 2010 by a total of eight third-grade teachers from the Canton of Schaffhausen. The teachers were both from the city of Schaffhausen and the countryside. In total, 105 children (60 boys and 45 girls) participated in the pilot test. A test-retest situation was not possible due to the large number of items on the test (100) and because the teachers were not willing. Teachers were instructed to carry out the test at the end of the school year in several sittings, to not complete the whole test with their learners in one session. Furthermore, teachers could choose the order in which they completed each section of the test as they deemed fit for their class. It was made clear that the test was difficult but that nothing would be learnt if it were too easy (see Appendix 8.3 for the instructions to the teachers in the final project, a similar page was provided to teachers in the pilot phase).

The official test used in July 2011 and 2012 was created through analyses of the pilot test. First of all, item-discrimination was determined through an inter-item-analysis of the 99 test items and 105 subjects. Items with a negative point-biserial coefficient or no variance were eliminated. This eliminated three Lingualevel subtests (most likely due to format, not language) and several items on the author's own tests. The Lingualevel subtests to be eliminated had one or two items per test with a point-biserial coefficient of  $>.25$ , but with only one or two items, it was not worth revising the subtest. Secondly, item difficulty was measured through calculating the mean for all learners and those items with a mean lower than  $.3$  and higher than  $.7$  were eliminated.

This original pilot test testing reading and listening was then shortened to a core of 30 items for reading, 30 items for listening and a writing test (open-ended prompt) was included. Table 5 and Table 6 provide short overviews of the final reading and listening tests and the full test is located in Appendix 8.2.

Table 5: Reading test overview

CEFR descriptors from lingualevel.ch	Subtest formats	Subtest genre	Subskill types	Time (min)
<b>Reading Subtest 1</b>				
A1.2 I can roughly understand short texts in picture stories if the pictures help me to guess many things.	Cloze summary Open answer Yes/No	Narrative: A short story about one boy's experience on St. Patrick's Day	Inference Summary Textual comprehension	10
<b>Reading Subtest 2</b>				
A2.1 I can understand the main points in short articles in magazines for children and teenagers, provided I know the subject relatively well.	Matching titles to excerpts Matching which excerpt answers a specific question.	Descriptive/ Expository: Excerpts in the style of „Time Magazine“ for children ( <a href="http://www.timeforkids.com/">www.timeforkids.com/</a> )	Textual comprehension Summary	10
<b>Reading Subtest 3</b>				
A2.1 I can understand much of the information provided in a short description of a person (for example a celebrity).	Open answer	Descriptive: A boy presents himself	Textual comprehension Summary	5
<b>Applicable to all reading tests</b>				
A1.2 I can read a simple, very short text, sentence by sentence and understand it; in so doing I pick out the information which I find clearest and read it, if necessary, several times.				

A range of formats was used in the reading tests: gap-filling for various types of words (nouns and verbs); some open answers where spelling was not corrected unless the written word simply did not make any sense; and matching text to title. For open answers, children were permitted to respond in either English or German. For the majority of the items, detail comprehension questions were asked (textual comprehension) though per subtest, there were either some gist questions (summary) or some inference questions. For example, in the first reading subtest, the question “How does Pat feel at the end of the day?” was asked in which the answer was not directly stated in the text but the expected answer went in the directions of “relieved”.

Table 6: Listening test overview

CEFR descriptors from lingualevel.ch	Subtest format	Subtest genre	Subskill types	Time (min)
<b>Listening Subtest 1</b>				
<b>A 2.1</b> I can understand the main points in short, simple stories and reports, but it is important that people talk slowly and clearly and that I know something about the subject they are talking about.	Multiple choice	Narrative: Two people discuss their winter vacation	Inference Textual comprehension	5
<b>Listening Subtest 2</b>				
<b>A 1.2</b> I can understand some words and expressions, when people are talking about me, my family, school, my free time or my surroundings but only if they are talking slowly and clearly.	Cloze summary	Narrative: A 2nd grade girl describes a project week about Africa.	Inference Summary Textual comprehension	10
<b>A 2.1</b> I can follow presentations which are quite straightforward, well-structured and illustrated by pictures or transparencies and which concern familiar subjects (for example, music, sports and other hobbies) but the speaker must speak slowly and clearly.				
<b>Listening Subtest 3</b>				
<b>A 1.2</b> Provided that people talk slowly and clearly, I can understand short conversations on subjects I am familiar with (for example school, family, hobbies).	Correct/ Partially correct/ Incorrect	Descriptive: Different class members are described.	Textual comprehension	5
<b>Listening Subtest 4</b>				
<b>A 1.</b> I can understand what people are saying about the colour and size of cars, houses, etc.; I can also understand who these things belong to.	Matching description to picture	Descriptive: Cars, houses and bikes described	Textual comprehension	5
<b>A 2.1</b> I can follow presentations which are quite straightforward, well-structured and illustrated by pictures or transparencies and which concern familiar subjects (for example, music, sports and other hobbies) but the speaker must speak slowly and clearly.				

The style of the listening test was similar to that of the reading test in that mainly narrative and descriptive text genres were used and similar subskill types were used. Two Lingualevel tests were adapted and used in Listening Subtest 1 and Listening

Subtest 3. All other recordings were professionally recorded in a studio in order to guarantee quality of the sound files, as poor quality recordings can influence the performance of the participants. Both the listening and reading tests were of equal length though there were more listening subtests. For the convenience of the teachers, the recordings could be played straight through, with each text twice.

The writing test was not included in the pilot study. The new test, with the additional writing test, was administered to third and fourth graders at the end of the 2011 and 2012 school years. The decision to open the test up to fourth graders having had the same teacher in the third grade was made as data collection was extremely difficult and it was challenging to get teachers to participate.

**Validity of the finalized version:** From the development in the pilot testing, the test used for the official round of data collection was developed. The reliability and validity of the test are ensured through the number of test formats used, the range of CEFR descriptors provided, the age-appropriateness of the content, the range of genres provided and the item type. Affective factors were left up to the teachers and not regulated, though teachers were encouraged not to do all the tests in one day and to give the children a break between the parts should they have chosen to do more than one part in one sitting. Furthermore, as this was a proficiency test, teachers were asked not to pre-teach the language found on the test or help with content during the test. These conditions were explicitly mentioned in the pilot testing and the official data collection period.

The three subtests (listening comprehension, reading comprehension and written production) comprised a total of seven different test formats. In what concerns the test format, works by various researchers (In'nami and Koizumi, 2009 or Vandergrift 2007) were referenced. Based on these studies, it was decided that more closed formats like multiple choice and short answer should be used as learners tend to perform better on them than on open-ended formats and they are less complicated to score. For the sake of diversity, a range of test formats was used to avoid better performance due to task familiarity – the author wanted to ensure performance on the test was not due to familiarity with the format, though in situations where testing is fairly regular, this can be an advantage. That said, many formats were quite familiar to the children due to the fact that textbook exercises use similar ones.

The subtests all came from contexts in which eight to ten year-old children are familiar. This can be justified through the fact that many topics are represented in the

learners' textbooks, a nine-year old child was asked to read through the test and provide feedback, and the author has experience with children of this age. In one listening subtest, for example, the author's daughter (second grade at the time) described a project week in schools which most schools have in this region though the topic of Africa may not be a common one.

**Objectivity.** In order to see if the test was objective, two measures were performed. First, the learners were asked if they liked reading and listening in English. A clear 50% said they liked (and didn't like) reading English and 84% ticked that they did NOT like listening to English. Liking reading correlated positively with reading performance ( $r = 0.16$ ,  $p < 0.01$ ) whereas liking listening did not correlate significantly with performance on the listening test. Thus learners who like reading performed well on the reading test though the same could not be said for the listening test.

Secondly, in order to see if the reading and listening tests were objective in the sense that learners could predict their performance, they were asked to say if they "can" or "can't" do certain tasks related to reading and listening – a simplified wording of the corresponding CEFR descriptors which were the basis for developing the learner tests (see Table 7). There were three to four descriptors for reading and the same for listening. The average score for each skill (sum of the reading and sum of the listening descriptors divided by the number of distractors) was calculated in terms of 1 being "yes, I can do most of these things" to 0 being "I cannot do these things", thus each child received a score between 0 and 1 indicating a tendency of performance in that skill.

*Table 7: Children's self-assessment of listening skills (example)*

Tick whether you can or cannot do the following things in English:	Yes	No
I can understand short conversations about things I know something about (school, family, houses, cars...) when people speak slowly and clearly.		
I can follow presentations where the speaker speaks slowly and clearly, where there are pictures or transparencies, and which concern familiar subjects (for example, music, sports and other hobbies).		
...		

A Spearman correlation was conducted to see how self-assessment related to performance on the specific reading or listening test. There was a slight positive correlation between learners predictions of the abilities in reading and their actual

performance on the reading test,  $r_s = 0.12$ ,  $p < 0.05$ . There was a positive correlation between learners predictions of the abilities in listening and their actual performance on the listening test,  $r_s = 0.20$ ,  $p < 0.01$ . These results suggest that indeed, learners could predict their skills, especially in listening, though not strongly.

### **3.1.2 Writing test development**

For the writing test, a simple open prompt requiring the form of an informal letter was used: "Imagine that you have a new pen pal in the USA. In your first letter, you want to tell him about your life. What would you write?" This prompt corresponds to the CEFR level of A1.2: "I can write down some personal information concerning myself (age, address or hobbies) on a list or a letter in which I introduce myself." The official aims for the end of third grade are not mentioned in the regional curriculum as writing is assumed to be a skill that is taught, but difficult to measure with beginners. Table 8 provides the rubric created to score the writing samples. This rubric is based on a combination of the American Council on the Teaching of Foreign Languages writing proficiency guidelines, on elements of writing that are described in the Swiss elementary school English Explorers series textbooks (Achermann & Sprague, 2006) and from Lingualevel (Lenz, 2007).



Table 8: Writing – scoring rubric

<b>Points</b>	<b>Communicative aspect/ (readability)</b>	<b>Range and fulfillment (range)</b>	<b>Focus on form (form)</b>	<b>Total words (length)</b>	<b>Sentence length (complexity)</b>	<b>Spelling (correct up to A2)</b>
<b>3</b>	<ul style="list-style-type: none"> <li>At least one question asked of pal.</li> <li>Thoughts connected and exemplified (I have one sister. Her name is Sarah).</li> <li>Appropriate letter ending (bye, love, etc).</li> </ul>	<ul style="list-style-type: none"> <li>Wide range of relevant structures and vocabulary within the task set (4 or more topics and 3+ starters). E.g. family; hobbies; school; house with “I have”; “My favorite”; “It is”...</li> </ul>	<ul style="list-style-type: none"> <li>Awareness of capitalization in English.</li> <li>Minor slips in accuracy (word order/tense / choice of preposition).</li> </ul>	66 +	6.5+ words / sentence	1-3 mistakes
<b>2</b>	<ul style="list-style-type: none"> <li>Thoughts connected but not always exemplified.</li> <li>Inappropriate letter ending but attempted.</li> </ul>	<ul style="list-style-type: none"> <li>More than adequate range of relevant structures and vocabulary within the task set (at least 3 topics and 2 starters).</li> </ul>	<ul style="list-style-type: none"> <li>Slips in awareness of capitalization in English.</li> <li>Some constructions incorrect.</li> </ul>	37-65	5-6.5 words / sentence	4-6 mistakes
<b>1</b>	<ul style="list-style-type: none"> <li>Requires some effort by the reader to follow thoughts (disconnected) and the fact that this should be a letter.</li> <li>Thoughts listed, not connected.</li> <li>No ending.</li> </ul>	<ul style="list-style-type: none"> <li>Not very wide range of relevant structures and vocabulary (2 topics, 1-2 starters). Same structure repeated.</li> </ul>	<ul style="list-style-type: none"> <li>Erratic / No awareness of capitalization in English.</li> <li>No awareness of English grammar.</li> </ul>	37 or fewer	4.8 or fewer words / sentence	7-10 mistakes
<b>0</b>	<ul style="list-style-type: none"> <li>Too little to evaluate, not a letter</li> </ul>	<ul style="list-style-type: none"> <li>Too little to evaluate; irrelevant contribution</li> </ul>	<ul style="list-style-type: none"> <li>Unreadable without knowledge of German.</li> </ul>	Words on paper show presence	Words on paper show presence	More than 10 mistakes

**Communicative aspect:** The category “communicative aspect” refers to elements not exclusively related to English language skills but on appropriate letter forms and ability to exemplify. This measure is also part of the official qualifying exams that the teachers took.

**Range and fulfilment:** This refers to the number of different structures used. For examples, learners who used “My name is...”, “I like...”, “I have...” and “My favorite” were accredited three points. Those who stuck to one or two of these structures were given one or two points.

**Focus on form:** This criterion refers to the awareness of primarily capitalization in English but also to accuracy of grammar structures. Those learners who over-capitalized nouns scored lower than those who showed an awareness of capitalization rules. Furthermore, here grammar plays a slight role in cases where, for example, a child writes "I have 2 sisters. She has..." would receive a lower score than a child who writes this correctly.

**Total words:** The evaluation of the total words was based on the Cambridge Young Learners Combined Vocabulary Lists for Starters, Movers, and Flyers tests (2007) (up to the A2 level). In cross-checking the learners’ texts according to these lists, very few learners used words not on the lists and with the number of spelling mistakes, it was decided to count the total quantity of words recognized as English, for example: ‘cat’ or ‘kat’ did not play a role, but if a child wrote “geh” instead of “go” this word was deducted from the total number of words. For each learner, the number of words was thus calculated and then four groups of learners were created though very few learners received a “0”.

**Spelling:** Spelling was included in a separate category and learners who used two or more words that were not on the Cambridge Young Learners List but were considered “high level” words (example: leopard) and spelled correctly were given an extra point. The number of incorrect words from these lists was counted for each writing sample and scores were assigned by having four groups (0-4) of more or less equal numbers of learners.

**Sentence length:** Finally, sentence length was calculated as a measure of text sophistication and complexity. Thus, the average total number of words per sentence for each learner was calculated and the learners were scored based on four categories (0-4) with very few learners receiving no points.

**Reliability of the writing categories.** The total number of points possible on this test was 19, and only one child who got “extra credit” on the spelling got a score of 3 on all the other criteria. The subcategories of “readability”, “range” and “focus on form” are more subjective measures than “spelling”, “total words” and “sentence length”. For this reason, two independent raters, both acquainted with the author, were asked to score three writing subskills: the communicative aspect of the text (readability); the range of vocabulary and structures (range) and the focus on form (form) as described in Table 8. Rater A was a student at the Zurich University of Teacher Education at the time and the Rater B is an experienced educator; both raters are non-native speakers of English but both have experience with children of this age in educational settings and recently passed standardized tests of English language skills (Rater A with a CAE B [approximately C1 level], Rater B with a CPE B [approximately C2 level]). Neither of the raters had any contact with any of the children involved in the study – impressions were made purely based on writing performance.

The Intraclass Correlation Coefficient (see McGraw & Wong, 1996) was used as a measure of reliability for the more subjective parts of the writing test. A two-way mixed analysis was used because each subject is assessed by each rater, but the raters are not random: they were specifically selected for the task. The absolute agreement setting was used because we want to see how identical the raters’ scores were. The average measures score is presented here. The closer to 1 the coefficient is, the better the agreement between the two raters. Table 9 provides these coefficients.

*Table 9: Interrater reliability – learners writing test scoring*

	Readability	Range	Focus on Form	Sum of raters scores on these three variables
Cronbach’s alpha	0.70	0.76	0.67	0.85
ICC(3,1)	0.68	0.70	0.63	0.84

There was the highest agreement amongst the raters for range, but they did not agree as much on the readability or the levels of accuracy of the extracts. In scrutinizing the data, it appears that the younger, and less fluent, rater was slightly harsher in scoring. It was decided to use the average of both raters’ scores for each of the three subcategories to contribute to the total score. Here, the ICC and

Cronbach's alpha are in the acceptable range of 0.7 - 0.8 (Larson-Hall 2010) and thus these ratings can be considered as reliable measures.

Thus the final total scores for the writing test was the average of the raters' scores for the three variables of readability, range and focus on form and the score based total words, sentence length and spelling. Figures 4 and 5 provide examples of high and low scoring learners. Figure 4 is an example of two extracts from relatively high scoring learners.

Figure 4: Writing sample from high scoring learners

Learner A

Dear Mark, Hallo have you a Playstation in the home? When you have one can you write a paper with yes on it when you don't have one can you write a paper with No on it okay Mark!

Mark My Name it [REDACTED] I'm living Switzerland I'm come's actually from Italy, I have '2' Brothers and 3 Sisters we are 3 Kids and 2 Parents, [REDACTED] Say's Good Bye I'm with glassis

Learner B

I'm [REDACTED]  
 I'm nine years old  
 I have one brother, one cat, one dog and one mother. My cat's name is Helen.  
 I have blue eyes and light brown hair. My best friend is Angela.  
 I live in [REDACTED] Have you a cat? R

In the top example (Learner A), the child's communicative ability was clearly highly rated due to the fact that questions were asked. Both children showed usage of different sentence starters, not just repeating the same ones all the time, with ample content. Awareness of capitalization rules in English is demonstrated, though not always consistently applied. Understanding German is not necessary to understand these extracts.

Figure 5 provides an example of a sample from two low scoring learners.

Figure 5: Writing sample from low scoring learners

Learner C

Dear Mark, mei heime is [redacted].  
 I haf a broder and a  
 cat. I haf a teich in  
 da Schwizer Land.

Learner D

Dear Mark, Was geht ab nichts. Was tust du in  
 Amerik. Mi hoby is Footbool gud  
 Samed. Waht is dein Hoby auch  
 Footbool. Is big Amerika yes. Waht  
 machst du in Amerika. Games bei  
 Dad spilen. In Gokarts fahren.

It is quite obvious in Figure 5 that one would have to be a German speaker to know what the learners were trying to express. Learner C's text was short, with only one structure used, "I have". Although with Learner D, quite a number of topics were covered and the text was rather long, there was no awareness of capitalization rules, of letter forms, and of spelling.

From these examples, we see how the tests were scored and what was asked of the learners. In Section 3.1.3, the results of the tests will be presented.

### 3.1.3 Learner test results

**Participants:** The final version of the learner test was carried out thanks to 18 classes and 279 learners who were provided color copies of the test. The condition for taking part in this study was that the same teacher had taught English in the third and fourth grades.

Of the learners, 148 were boys and 131 were girls. Two learners were omitted from the study as they noted that they speak English at home and their test results supported that statement. One hundred and eighteen learners were in their first year of learning English, one hundred and fifty-seven in their second year of learning English and four in their third year of learning English (the assumption here is that they repeated a class). Sixty-eight percent of the learners mentioned having one mother tongue (German), 29% mentioned two languages (being bilingual – namely in Italian, Albanian or Portuguese among a few other instances of other languages) and 3% were trilingual and listed 3 languages spoken at home.

**Scoring:** General scoring for each item on the reading and writing tests was "right or wrong" and answers left blank were counted as "wrong" if other items on that specific subtest had been completed. One part of a class (8 learners) did not participate in the reading test. These missing scores were imputed based on the writing and listening scores of all the learners. Another class of eighteen learners did not participate in the writing test. These missing scores were imputed based on the reading and listening scores of all the learners having completed the entire test. A copy of the final version of the test is found in Appendix 8.2 as well as online at <http://www.edacross.org/index.php/dissertation>.

**Reliability.** In the final version of the test, upon closer examination of the individual items, two items on the reading subtest had a point biserial of less than 0.15 and were dropped to make the test stronger and to have the same number of

items on the listening test (see the crossed out items in Appendix 8.2). Item difficulty for the reading test ranged from 0.2 to 0.8, with the majority of the data being in the 0.4 - 0.6 range. Cronbach's alpha for the 30 items on the reading comprehension subtest was 0.83.

Item difficulty for the listening test ranged from 0.3 to 0.9, with the majority of the data being in the 0.4 - 0.6 range. Cronbach's alpha for the 30 items on the listening comprehension subtest was 0.78. Thus, both the reading and listening comprehension subtests were one point each, each section having 30 points and demonstrate a respectable internal consistency. The results for the writing test were reported in Section 3.1.2.

Once the item analyses were completed for each subtest, an exploratory data analysis was conducted to determine if the total and subtest score distributions were normally distributed. Results for the Kolmogorov-Smirnov test for normality indicated that the distribution of the test scores on the total test did not deviate significantly from a normal distribution ( $D(279) = 0.04, p = 0.20$ ). Table 10 provides detailed descriptive statistics for the total subtest scores and total test score with a breakdown by number of years of English instruction the children had.

Table 10: Descriptive statistics of the learner subtests and total scores

	<i>N</i>	<i>Minimum</i>	<i>Maximum</i>	<i>Mean</i>	<i>Std. Deviation</i>	<i>Variance</i>	<i>Median</i>
<b>Listening scores (3<sup>rd</sup> and 4<sup>th</sup> graders)</b>	<b>279</b>	<b>5</b>	<b>29</b>	<b>15.54</b>	<b>5.18</b>	<b>26.84</b>	<b>15.00</b>
3 <sup>rd</sup> graders	118	5	29	15.34	5.45	29.78	14.00
4 <sup>th</sup> graders	161	5	28	15.70	4.98	24.08	15.00
<b>Reading scores (3<sup>rd</sup> and 4<sup>th</sup> graders)</b>	<b>279</b>	<b>3</b>	<b>31</b>	<b>16.14</b>	<b>5.98</b>	<b>35.74</b>	<b>16.00</b>
3 <sup>rd</sup> graders	118	4	30	15.53	5.80	33.66	16.00
4 <sup>th</sup> graders	161	3	31	16.58	6.08	37.18	16.00
<b>Writing scores (3<sup>rd</sup> and 4<sup>th</sup> graders)</b>	<b>279</b>	<b>4</b>	<b>19</b>	<b>12.45</b>	<b>3.14</b>	<b>9.87</b>	<b>12.50</b>
3 <sup>rd</sup> graders	118	4	18	11.38	3.05	9.31	11.50
4 <sup>th</sup> graders	161	6	19	13.23	2.98	8.89	13.46
<b>Total scores (3<sup>rd</sup> and 4<sup>th</sup> graders)</b>	<b>279</b>	<b>20</b>	<b>71</b>	<b>44.13</b>	<b>11.12</b>	<b>123.63</b>	<b>44.50</b>
3 <sup>rd</sup> graders	118	24	69	42.25	10.91	119.05	43.00
4 <sup>th</sup> graders	161	20	71	45.50	11.10	123.24	44.50

The information in Table 10 shows us that, as to be expected, mean scores increased from one year to the next between the third and the fourth graders, though never dramatically, and the range of performances was much more diverse for fourth graders than for third graders. That the means and medians are fairly close indicates that the scores are fairly well divided around the mean, an indication of normalcy. It would appear that learners in the fourth grade improve in reading and writing more than in listening as the difference in average scores for each of these skills is higher from third to fourth grades, though that could well be simply in this population.

In summary, the reading and listening tests provide solid grounds for describing learner performance after one or two years of English instruction. Reliability analyses show a high internal consistency of the individual items. Teachers were encouraged to create a suitable test setting and to take their time with the tests. The test types were developed to reflect the world of third and fourth grade Swiss children on the level of the topics chosen and the setting of testing reading, writing and listening as is done for report card grades in Swiss elementary schools. Although some of the subtests received criticism from participating teachers for being too difficult, none were criticized as being contextually inappropriate. The tests attempted to reflect authenticity in providing contexts learners of this age group are comfortable with, and through a variety of test formats. Therefore, the scores can be interpreted as a reliable measure of learner performance.



## 3.2 Teacher survey

### 3.2.1 Teacher survey overview

The teacher survey (Appendix 8.1) was created and administered in 2011 and 2012 in order to gather information on teachers' language skills, their exposure to English outside of the classroom and their ranking of time spent on specific skills-based activities in their teaching. This survey is a general data collection instrument and the term "survey" here is not used in the sense of gathering opinions and attitudes, but rather teacher scores and estimated exposure to English and time spent teaching various skills.

**Teacher language skills.** In part 1 of the survey, teachers were asked to provide their official test results on the Cambridge Advanced Exam (or other qualifying exam) as well as an estimation of whether their ability in each subskill (Table 11) had improved, stayed the same, or worsened since they took the respective test. The Cambridge exam official reports that the candidates receive provide total scores for each of the categories listed in Table 11.

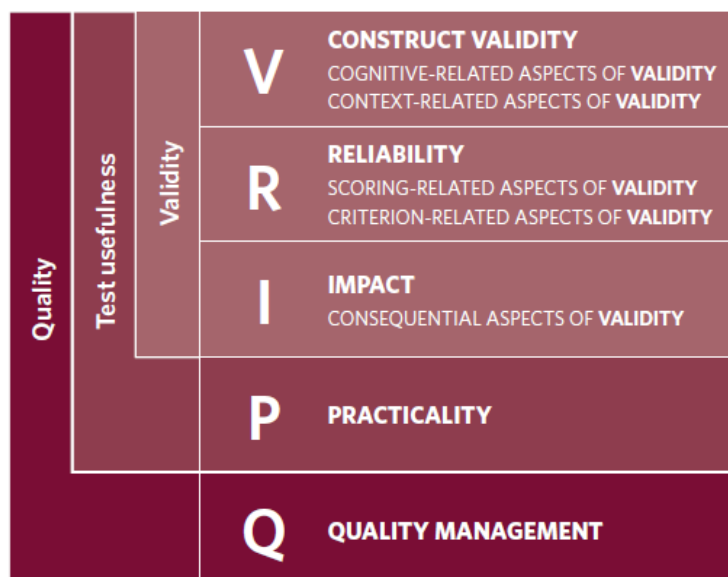
*Table 11: Cambridge Exams Overview*

Reading	Shows candidate can deal confidently with different types of text, such as fiction, newspapers and magazines.
Writing	Candidate creates two different pieces of writing, such as articles, essays, letters, proposals, reports and reviews.
Use of English	Tests use of English with different types of exercise that show how well candidate can control grammar and vocabulary.
Listening	Tests ability to follow and understand a range of spoken materials, such as interviews, radio broadcasts, presentations, talks and everyday conversations.
Speaking	Tests ability to communicate effectively in face-to-face situations.

<http://www.cambridgeenglish.org/exams-and-qualifications/advanced/whats-in-the-exam/>

Some of the test constructs parallel to, albeit at a different level, what the learners were asked to do. The Cambridge Syndicate guarantees the quality of their exams as described in *Principles of Good Practice: Quality Management and Validation in Language Assessment* (Cambridge, 2013) through the use of the VRIPQ (validity, reliability, impact, practicality, quality management) structure, as seen in Figure 6.

Figure 6: VRIPQ model (Cambridge, 2013, p. 10)



The diagram above illustrates our approach to achieving fitness for purpose, and shows how developments in language testing theory and in quality management have been taken on board.

**Teacher contact with English.** Part 2 of the survey asked teachers to provide an overview of their contact with English outside the classroom. They were asked to state their longest stay abroad in a country where English was the primary language of communication, the vacation weeks per year they stay in a country where English is used, and how often they read, write, speak and listen to English (never, weekly or monthly).

**Teacher instructional practices.** Finally, part 3 of the survey gathered information on instructional practices as well as on the amount of German used in one forty-five minute English lesson. Instructional practices in this context are defined as priorities in teaching based on a general estimation of occurrences of certain activities in the classroom as well as knowing the frequency of these activities in the sense that the activity occurred in every lesson (see Section 3.2.2.3, Table 16). From this information, impressions of teaching behaviors and thus teachers feelings about what is important (strategy building, teacher-control, learner-centeredness) can be suggested. An equal number of activities for each language skill was presented. The original survey was tested on three teacher acquaintances for short feedback about the readability, content and structure and was used with minor changes to formulations and format.

### 3.2.2 Teacher survey findings

In the following sections, the data collected via the teacher survey will be presented first generally, then section by section. Results for the entire teacher survey (in table form) are presented at the end of the survey in Appendix 8.1.

**Participants.** Of the teachers participating in the survey, the number of years of teaching experience ranges from three to twenty-eight. All teachers speak German as their mother tongue and one teacher grew up bilingually with English as well. Though only 50% of the teachers consider themselves bilingual, numerous other languages were listed as other foreign languages spoken (including French, Japanese, Finnish, Italian, Spanish, Portuguese and Russian). All teachers ticked that they enjoyed teaching English though one left the comment that she did not enjoy using the compulsory textbooks.<sup>28</sup> The teachers in this study had been teaching, on average, for 16 years ranging from one to twenty-five years.

#### 3.2.2.1 Survey section 1 results – Teacher language skills

Section one of this survey required teachers to provide information about their test scores on the qualifying exam to teach English and an estimation of their current skill level as compared to when the qualifying exam was taken. The exams taken were:

- the FCE (Cambridge First Certification in English) by one teacher;
- the CAE (Cambridge Certificate in Advanced English) by nine teachers;
- the CPE (Certificate of Proficiency in English) by four teachers;
- the IELTS: International English Language Testing System by one teacher; and
- an internal PH Zurich examination by two teachers.

Teacher test scores were converted to the scale shown in Table 12. Cambridge Examinations offers a converter between several exams<sup>29</sup> and this was used for the teachers' individual subtest (reading, writing, speaking, listening, Use of English)

---

<sup>28</sup> The *First Choice* and *Explorers* series textbooks which teachers are expected to use in the canton of Schaffhausen have undergone a lot of criticism. Decisions about which textbooks to use are a top-down process where the boards of education decide which books teachers should use.

<sup>29</sup> Exam result converter: <http://www.cambridgeenglish.org/exams-and-qualifications/advanced/results/exam-results-converter/>

scores (A [best], B, C, or D [worst]) as well as the score on the entire test. The internal PH Zurich exam is made up of copied bits and pieces of the CAE exam and administered by lecturers who also work with the Cambridge system. As the IELTS (International English Language Testing System) exam does not test the Use of English, for the one teacher with no scores for this section, a score was estimated based on the scores from the other subtests.

Table 12: Conversion chart and teacher scores

Level	CEFR Level	Code	FCE score	CAE score	CPE score	IELTS score	PH Zurich Internal exam score	% of learners with teachers obtaining this total score
Low	B2	1	FCE 3 (B)				1	7.90%
		2	FCE 4 (A)	CAE 1 (D)			2	
↓	C1	3		CAE 2 (C)		6.5 -7.0	3	29.00%
		4		CAE 3 (B)		7.5	4	37.30 %
↓	C2	5		CAE 4 (A)	CPE 1 (D)		5	
		6			CPE 2 (C)			7.90 %
↔high		7			CPE 3 (B)			14.30 %
		8			CPE 4 (A)			
		9	English mother tongue					3.60 %

Total exam scores from the teachers in this study are shown in Table 12 and range from one mother tongue teacher (who grew up in Japan to Swiss-English parents and was schooled in English) to two teachers with a FCE B level, which is below the official qualifying level to teach English in Zurich and Schaffhausen elementary schools. In short, of the teachers involved in this study, there is a relatively wide range of measured skills. The frequency of the scores are presented in percentages as what is important in the rest of this study is the number of students having had this teacher – some teachers had classes of fifteen learners, whilst others had classes of twenty-three. This is thus a reflection of which percentage of learners had a teacher with a certain score.

Table 13 provides an overview of teachers' subtest scores on the respective tests.

Table 13: Teacher subtest scores

Scores (low to high) →	1 FCE B	2 CAE D / FCE A	3 CAE C	4 CAE B	5 CPE D / CAE A	6 CPE C	7 CPE B	8 CPE A	9 native speaker
Subtest	Percentage of learners with a teacher obtaining this score								
Reading	7.90%	13.30%	14.70%	38.40%	--	--	7.90%	14.30%	3.60%
Writing	14.00%	7.50%	20.80%	24.40%	7.50%	--	7.90%	14.30%	3.60%
Use of English	8.60%	5.40%	19.40%	31.50%	6.50%	0.40%	0.70%	23.30%	4.30%
Listening	7.90%	6.10%	15.40%	44.80%	--	7.90%	--	14.30%	3.60%
Speaking	7.90%	7.50%	14.30%	30.80%	13.60%	--	--	22.20%	3.60%

Again, the range is fairly well distributed with learners having teachers across the spectrum of levels of skills. Teacher writing performance does not seem to be very strong. Scores on the Use of English were relatively high and well distributed. The majority of teachers had relatively good speaking skills compared to writing skills which might be expected due to the necessary communicative ability of teachers.

Teachers were also asked if they thought their skills at gotten worse, stayed the same, or gotten better in each of the subtest categories since having taken the qualifying exam (Table 14).

Table 14: Teachers' feeling of skill change since exam

Feeling of improvement	Gotten worse	Stayed the same	Improved
Subtest	percentage of learners having this teacher		
Reading	7%	60%	33%
Writing	32%	48%	20%
Use of English	49%	43%	8%
Listening	7%	66%	27%
Speaking	29%	39%	32%

Although the majority felt their skills had stayed the same, many felt that their "Use of English" skills had gotten worse. Reading and listening skills appear as well to have improved since the time of the qualifying exam, perhaps an indication of daily usage in the classroom.

### 3.2.2.2 Survey section 2 results – Teacher contact with English

The second section of the survey asked for information about teachers' stays abroad in countries where they used English (whether or not it was the official language of the country or not), about their regular contact with English outside of the classroom through an estimation of the frequency their use of reading, writing, listening and speaking skills (Table 15).

*Table 15: Teacher contact with English outside of the classroom*

	Range	Mean	SD
Years abroad in a country where EN was language of communication	1-19	5.90	6.90
Weeks per year spent on holiday in a place where EN used	0-5	1.30	1.70
Frequency in EN outside of school time (1=never; 2=monthly; 3=daily)			
Reading	1-3	2.60	0.60
Writing	1-3	2.00	0.78
Speaking	1-3	2.30	0.68
Listening	1-3	2.56	0.60

The length of stays abroad ranged from not quite a year to nineteen years and due to this there is quite a large standard deviation. The teachers also ranged quite drastically in the number of weeks they spend using their English on holidays – from nothing at all to a full five weeks a year (almost their entire vacation).

Regular contact with English outside the classroom is fairly common for all skills, but especially in terms of reading and listening; most have contact with English through reading for fun, through searching on the internet and with friends and neighbors, as was mentioned by data collected in the open field with the statement: "Please describe your regular contact with the English language (for example, your partner is English-speaking, you regularly read English literature, you take English lessons)". The largest variation was for writing, also with the lowest mean, where many teachers had very little contact via writing. It cannot be judged as to whether this is representative of the average teacher in the region, though contact with teachers in in-service workshops in the field would support this contention.

### 3.2.2.3 Survey section 3 results – Instructional practices

In the third section of the survey, teachers were asked about their teaching habits in order to find out how much emphasis they place on specific reading, writing, listening and speaking skills in the classroom. Firstly, they were asked if an explicit

focus on the activities listed in Table 16 occurred in every lesson and to rank the frequency of occurrence. The choice of activities reflects common elements of teaching such as a focus on strategies, group work and individual work, and whole class teaching.

*Table 16: Class activities*

Activity done with the class	Number of classes where it takes place in every lesson
speaking / singing in chorus	14
speaking / listening to English in pairs/groups	9
speaking in pairs / to the teacher individually	8
developing speaking strategies (practice in front of a mirror, etc...)	1
presenting to the whole class	4
listening to the teacher	18
listening to a CD or other media	10
developing listening strategies	0
reading words	16
reading short texts / stories	6
developing reading strategies	2
writing words	14
writing short texts (e.g. sentence completion)	3
developing writing strategies	1

We see in Table 16 that teachers tend to prefer rather controlled activities – listening to the teacher, speaking in chorus, reading and writing words. Furthermore, the results of this simple tally clearly show that the teachers involved do not place an explicit focus on strategy building and much class time is devoted to the word level in a fairly controlled fashion. Presenting to the class and more open activities that require more learner output were not favored.

Teachers were also instructed to rank the activities mentioned in Table 16 as to their importance in their teaching and to provide the same rank number if activities had more or less equal frequencies: “Number the activities in the order of how frequently they occur in your lessons (1 for the most often).” Unfortunately, the ranking data were ambiguous. From a total of fourteen items/activities, a maximum of one to fourteen would have been possible but as many teachers place equal

weight on certain activities or certain activities don't take place at all in their classrooms, the range never varied to that extent.

Thus, as the ranking data were so difficult to interpret, the information about whether a type of activity occurred in every lesson was the main source of information for the proceeding step. An independent expert (Rater B who scored the writing tests) was asked to see if any main categories (clusters) of teacher types or styles of teaching could be made from these data. The expert found three main categories to be apparent through the information that certain activities occurred in every lesson. These categories were then cross-checked with the rankings to ensure if a teacher said "reading, writing, speaking and listening" occurred in every lesson, then the frequency of these activities was also high. The following categories were suggested and accepted:

- teachers who had a relatively balanced focus on teaching all the four skills;
- teachers who were very literacy oriented (children spent most of the time in class reading and writing); and
- teachers who did not focus on writing.

These three categories allow for the following thoughts. First of all, teachers with a balanced skills approach are perhaps aware of the report card criteria in the upper elementary where all four skills are equally reflected in report card grades. They may also be aware of the difficulties of writing in English, thus promote an early start. Furthermore, they may be aware of secondary teachers' wishes that writing be more of a focus in the elementary school. Secondly, teachers who were more traditional perhaps are concerned with classroom management or do not feel their own skills are good enough to focus on speaking-oriented settings. Finally, teachers who neglect writing perhaps have the tendency to have "communicative classrooms" without understanding that writing can also be a part of those.

Teachers were asked questions about the textbooks they use as well. The majority of them taught with the Young World materials from the Klett Verlag (8 of 18). Five teachers use First Choice, three use Explorers and one each use Primary Colours and Super Bus.<sup>30</sup> Eleven teachers do not follow the teacher's notes while

---

<sup>30</sup> First Choice and Explorers are officially recognized textbooks in the EDK-East cantons from the Lehrmittelverlag des Kantons Zürich. The Klett Verlag books are also approved. Primary Colours (Publisher) and Super Bus (Macmillan) have not been approved.



five tend to stick to them. All the teachers surveyed use additional materials, but especially those working with Young World.

Finally, teachers were asked about their time spent in German per 45 minute lesson of English. One teacher refrained from answering but the range varied from two to thirty minutes per forty-five minute lesson. More explicitly, teachers were asked about their decision-making process for switching to German in English lessons. The most common reason for German was giving complex instructions, followed by:

- repeating in German for the whole class when there are some children who did not understand;
- repeating in German to individual children when they do not understand;
- explaining grammar points in German.

These data provide insight into the populations involved in this study. The participating teachers have quite a range of abilities in the English language and experiences abroad which will be advantageous in answering the research questions. Their habits in the classroom are not as diverse, but three main tendencies can be found and the amount of German in the classroom differs quite dramatically.

From the learner test, we learned that third and fourth graders are capable of rather complex tasks in English. Their listening and reading performance indicates that they are able to work up to an A2.1 level of skill. Writing skills between third and fourth graders was quite varied and ranged between an A1.1 level and an A2.1 level. The data from the tests appear to be reliable.

#### **IV. Associations between teacher factors and learner performance**

Chapter III described the process of developing the measurement instruments and the descriptive findings from the learner test and the teacher survey. Chapter IV will now provide findings in relation to the research questions by looking at the associations between the teacher survey and the learner test. Section 4.1 describes the methods used and the results are presented in Sections 4.2 and 4.3. Section 4.4 is a consolidation of the findings.

## 4.1 Regression Methods

The aim of these analyses is to see which teacher factors, controlling for the grade level of the learners, could perhaps be the best predictors of learner performance at the end of the first years of English instruction. In order to do this, the analyses were performed using the statistical software R (version 2.15.3). Examples of the codes used can be found in Appendix 8.4 and the full codes are available online at (<http://www.edacross.org/index.php/dissertation>).

First of all, each research question was investigated separately (Section 4.2). Afterwards, multivariate mixed models (multilevel models) were run for each of the learner language proficiency subtest scores using a selection of the teacher variables and the learner grade level (Section 4.3).

For each of the separate language subtests that the learners took, the following linear mixed model equation was calculated:

$$\text{Learner Score}_{ij} = \beta_0 + \beta_1 \cdot \mathbf{\text{Teacher Score}}_i + \beta_2 \cdot \text{Grade level}_i + b_i + \varepsilon_{ij}$$

where

- $i$  and  $j$  are indices for classes ( $i=1, \dots, 18$ ) and for the learners ( $j=1, \dots, n_j$ );
- $\beta_0$  is the mean learner score when the teacher's score is zero and the grade level corresponds to the 3rd grade;
- teacher score (in bold) is the independent variable and was replaced by other variables (teacher contact with English, teaching type) for each research question;
- $\beta_1$  is the linear effect of teacher language competence (or teacher contact or teacher teaching type) on learner language competence;
- $\beta_2$  is the difference between the mean learner score in the grade level category „4th grade“ and the reference category „3rd grade“;
- $\varepsilon_{ij}$  denotes the error term of the model and is also assumed to have a standard Gaussian distribution (mean zero);
- $b_i$  is the teacher-specific random intercept assumed to have a standard Gaussian distribution (mean zero).

If there had only been one single observation per teacher (such as the class average score on a test), it would have been possible to calculate the correlation coefficient between the teacher scores and learner scores and thereby the research question could have been answered through a simple regression model. However,

due to the complex data structure (each teacher has a group of individual learners with varying scores and each learner has one teacher with many different attributes, such as level of English), a linear mixed model had to be used in order to obtain a valid coefficient (parameter  $\beta_1$  in the model). Furthermore, a random effect is included because it is assumed that there is some sort of relationship between the teacher and learner variables.

The association between teacher factors and learner language competence is significant when the coefficient  $\beta_1$  is significantly different from zero, thus when the corresponding p-value is lower than the significance level  $\alpha = 0.05$ . The research question is confirmed when the coefficient is significantly positive. This approach has been applied to most of the research questions in this study, unless otherwise noted.

## 4.2 Results for individual predictors

### 4.2.1 Teacher proficiency level

*Do teachers with higher levels of English have learners whose competence in reading, writing and listening are also higher than those learners in classrooms with teachers having lower levels of language competence?*

This research question relates to a teacher's total language competence measure (total score on the qualifying exam) as well as to the five separate subtests of reading, writing, listening, use of English and speaking. Table 17 provides an overview of teacher scores on their official qualifying exams and learner total scores on the administered test.

*Table 17: Learner total scores grouped by teacher total scores*

Grade	Learner total scores				Learner total scores distribution	
	Teacher score (CEFR level)	Mean	Med	Sd	Min	Max
3rd	CAE C (C1)	43.42	44.25	11.21	23.58	66.50
	CAE B (C1)	44.38	45.00	10.60	24.00	68.50
	CPE C (C2)	32.33	33.25	6.78	23.50	45.50
	NS*	37.50	35.09	9.22	24.80	53.60
	Total	42.24	43.00	10.96	23.50	68.50
4th	FCE B (B2)	50.90	52.50	10.27	25.50	65.00
	CAE C (C1)	46.54	45.00	10.32	25.00	69.50
	CAE B (C1)	42.01	40.50	12.72	20.00	71.00
	CPE C (C2)	47.52	45.00	10.23	34.67	70.50
	CPE A (C2)	44.95	44.75	9.27	23.50	67.50
	Total	45.49	44.50	11.07	20.00	71.00
3 <sup>rd</sup> and 4 <sup>th</sup> grades	FCE B (B2)	50.90	52.50	10.27	25.50	65.00
	CAE C (C1)	45.07	44.50	10.80	23.58	69.50
	CAE B (C1)	43.31	43.50	11.60	20.00	71.00
	CPE C (C2)	39.24	38.25	11.35	23.50	70.50
	CPE A (C2)	44.95	44.75	9.27	23.50	67.50
	NS*	37.50	35.09	9.22	24.80	53.60
Total	44.13	44.50	11.12	20.00	71.00	

\*NS refers to "native speaker of English".

Table 17 indicates that the teachers with the higher level proficiency (native speakers or those with a CPE exam) were not necessarily those with the best

performing classes. On the contrary, it would appear that the lower the teachers' scores, the better the class performed. Fourth graders did perform slightly better than third graders, though not by much.

Table 18 provides an overview of the mean, median and standard deviations for the learner listening, reading and writing subtests and is broken down again by teacher score on the qualifying exam.

*Table 18: Learner listening, reading and writing scores grouped by teacher total scores*

Grade	Teacher score (CEFR level)	Learner listening scores			Learner reading scores			Learner writing scores		
		Mean	Med	Sd	Mean	Med	Sd	Mean	Med	Sd
3rd	CAE C (C1)	16.39	16.50	5.39	15.67	16.00	5.38	11.35	10.50	3.31
	CAE B (C1)	16.63	16.00	5.15	16.54	17.00	6.03	11.21	11.50	2.97
	CPE C (C2)	9.67	10.50	3.31	11.33	12.50	4.14	11.33	11.00	3.08
	NS*	10.90	11.00	2.96	14.19	14.10	6.30	12.41	12.46	2.79
	Total	15.35	14.00	5.48	15.53	16.00	5.83	11.37	11.50	3.06
4th	FCE B (B2)	17.36	19.00	4.25	20.05	21.00	5.77	13.49	13.25	3.05
	CAE C (C1)	16.50	16.00	4.59	16.32	16.00	6.18	13.57	14.00	2.87
	CAE B (C1)	14.38	14.00	5.40	14.62	14.00	6.59	13.01	13.00	3.47
	CPE C (C2)	17.00	15.00	4.88	18.30	17.00	4.97	12.22	11.75	2.88
	CPE A (C2)	14.90	14.00	4.88	16.83	17.50	4.85	13.22	13.25	2.45
	Total	15.68	15.00	4.97	16.58	16.00	6.07	13.23	13.46	2.97
3rd and 4th grades	FCE B (B2)	17.36	19.00	4.25	20.05	21.00	5.77	13.49	13.25	3.05
	CAE C (C1)	16.53	16.00	4.95	16.02	16.00	5.79	15.52	12.50	3.26
	CAE B (C1)	15.62	15.00	5.36	15.67	16.00	6.33	12.02	12.00	3.32
	CPE C (C2)	13.00	12.50	5.47	14.50	15.00	5.67	11.74	11.25	2.95
	CPE A (C2)	14.90	14.00	4.88	16.83	17.50	4.85	13.22	13.25	2.45
	NS*	10.90	11.00	2.96	14.19	14.10	6.30	12.41	12.46	2.79
Total	15.54	15.00	5.18	16.14	16.00	5.98	12.45	12.50	3.14	

\*NS refers to "native speaker of English".

It appears here that the third and fourth grade teachers with the highest levels of measured language proficiency clearly had learners who performed more poorly on the listening test than learners of other teachers. This is not the case for writing, however, where there is not such a large difference in scores between the teachers and the distribution seems a bit more random. For reading, fourth graders performed better than third graders and there is some variation from teachers scoring higher and lower, but generally, the teachers with the highest scores on their measure did

not have learners in their classes with the highest reading scores. As the standard deviations for the reading test vary to such a degree, it would seem that in this sample of teachers and learners there was quite a range of performances.

Table 19 shows a summary of all linear mixed models calculated in order to investigate this first question.

*Table 19: Linear mixed models for teacher and learner scores*

Learner score	Teacher score	$\beta_0^*$	$\beta_1^{**}$	$sd(\beta_1)$	$p$	$\beta_2^{***}$	$p$	$sd(b_i)$
Total	Total	45.44	-0.83	0.67	0.23	4.06	0.17	5.25
Total	Reading	44.50	-0.66	0.61	0.30	4.29	0.15	5.31
Total	Writing	45.06	-0.76	0.58	0.21	4.09	0.23	5.44
Total	Listening	45.40	-0.85	0.66	0.22	4.26	0.21	5.20
Total	UE****	46.97	-1.10	0.57	0.07	4.01	0.15	4.87
Total	Speaking	47.65	<b>-1.25</b>	<b>0.55</b>	<b>0.04</b>	4.30	0.11	4.65
Listening	Total	17.99	-0.63	0.38	0.12	0.73	0.65	3.15
Listening	Reading	17.31	-0.50	0.35	0.18	0.90	0.59	3.21
Listening	Writing	17.42	-0.50	0.33	0.16	0.74	0.65	3.18
Listening	Listening	17.90	-0.64	0.38	0.11	0.88	0.59	3.13
Listening	UE	18.89	<b>-0.77</b>	<b>0.31</b>	<b>0.03</b>	0.72	0.62	2.84
Listening	Speaking	19.23	<b>-0.85</b>	<b>0.30</b>	<b>0.01</b>	0.92	0.52	2.70
Reading	Total	16.42	-0.25	0.29	0.41	1.35	0.29	2.11
Reading	Reading	16.14	-0.20	0.27	0.48	1.43	0.27	2.12
Reading	Writing	16.34	-0.23	0.26	0.38	1.36	0.29	2.10
Reading	Listening	16.40	-0.25	0.30	0.40	1.41	0.27	2.10
Reading	UE	16.90	-0.33	0.26	0.22	1.33	0.28	2.02
Reading	Speaking	17.12	-0.38	0.25	0.16	1.42	0.25	1.98
Writing	Total	11.24	0.02	0.13	0.89	<b>1.90</b>	<b>0.00</b>	0.91
Writing	Reading	11.23	0.02	0.12	0.86	<b>1.89</b>	<b>0.00</b>	0.91
Writing	Writing	11.45	-0.03	0.12	0.80	<b>1.90</b>	<b>0.00</b>	0.91
Writing	Listening	11.23	0.02	0.13	0.88	<b>1.89</b>	<b>0.00</b>	0.91
Writing	UE	11.36	-0.01	0.12	0.95	<b>1.90</b>	<b>0.00</b>	0.91
Writing	Speaking	11.47	-0.03	0.12	0.80	<b>1.91</b>	<b>0.00</b>	0.91

\* Mean learner score when the teacher score is zero and learners in the third grade / \*\* Linear effect of teachers' measured tests scores on learner's test performance (points) / \*\*\*Difference between mean learner score in 4th grade and reference category "3rd grade" / \*\*\*\*UE stands for "Use of English"

The standard deviations  $sd(b_i)$  of the random intercepts are large. This emphasizes the finding that there is a large (unobserved) heterogeneity between

the teachers in this data (as will also be seen in the boxplots in Figure 7 onwards). It is not possible to confirm the hypothesis that teachers' language scores interact positively with learners' scores. On the contrary, the teachers' Use of English (UE) scores and the teachers' speaking scores were found to have a significantly negative association with learners' listening scores and the teachers' speaking scores were also significantly negatively associated with learners' total scores. It appears that though insignificant, the better teachers in this study performed on the language tests, the poorer their learners performed on theirs, as there is, with the exception of learners' writing scores, a negative association between teacher and learner scores. Furthermore, it is clear in this sample that there is no significant effect of grade level in reading and listening though there is on the writing test.

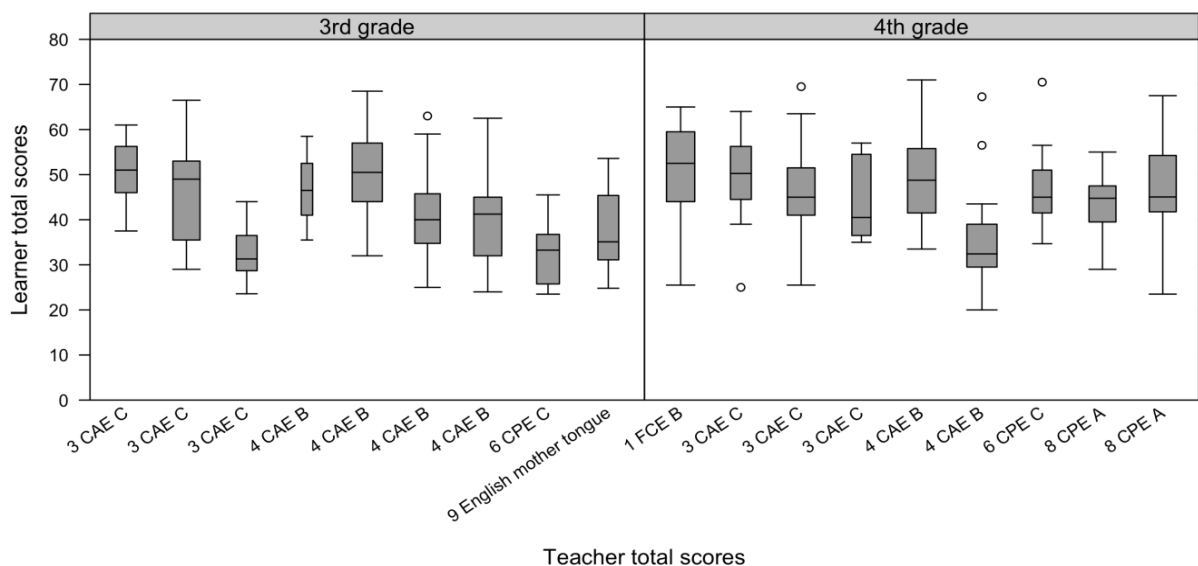
The hypotheses were that there would be positive correlations between the teachers' performance on the Use of English and the Reading parts of the Cambridge exams and learner performance, and thus through Table 19 it appears that these hypotheses must be rejected.

In the following, boxplot diagrams will be shown for the teacher total scores (increasing from left to right for each grade) and the learner total scores as well as for each of the significant combinations between teacher listening scores and learner score. These boxplots help to visualize the distribution of learner total scores for each teacher and each box stands for the learners of one teacher. The widths of the boxes are scaled relative to the group sizes (narrow box = small group size). The dots represent outliers in the respective group.

**Teacher total scores, learner total scores.** As can be seen in Figure 7, there is no clear indication that the mean or median learner total scores increase with increasing teacher total scores as hypothesized – on the contrary, they seem to decrease. Furthermore, one can see the large heterogeneity between teachers. The linear mixed model estimated the coefficient for the association to be  $\beta_1 = -0.83$  ( $p=0.23$ ).



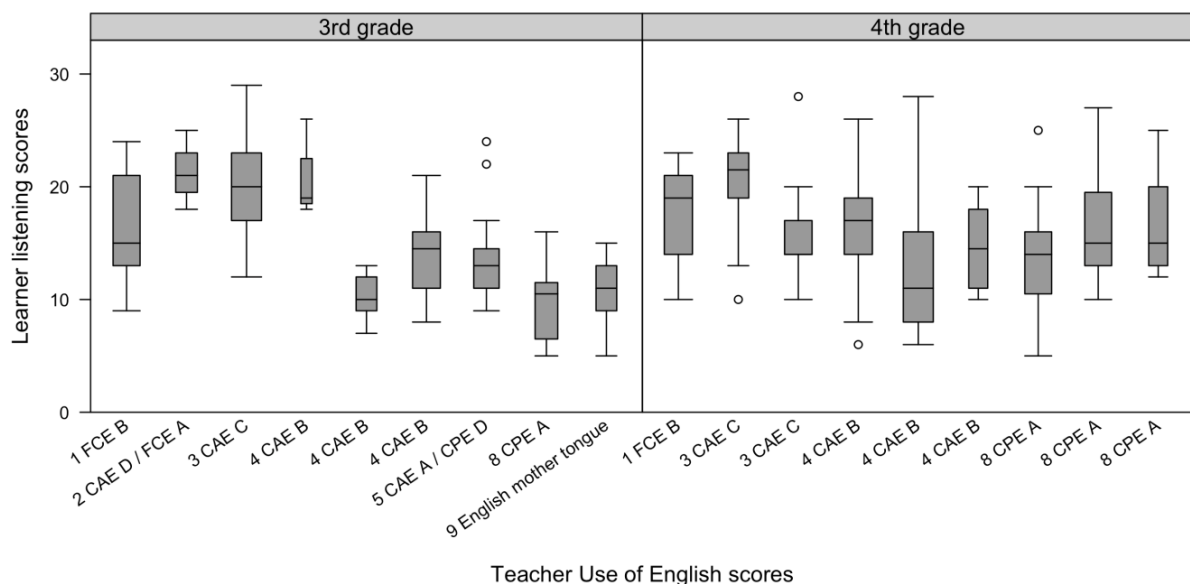
Figure 7: Association between teacher total scores and learner total scores



A negative association between teachers' and learners' total scores has been found, but is not significant. If the teacher's total score increases by one point and the grade level remains the same, the expected mean learner total score decreases by 0.83 points.

**Teacher Use of English scores, learner listening scores.** Teacher scores on the Use of English part of the qualifying exam were associated with learner listening scores as shown in Figure 8. The coefficient for the association was found to be  $\beta_1 = -.77$  ( $p=0.03$ ).

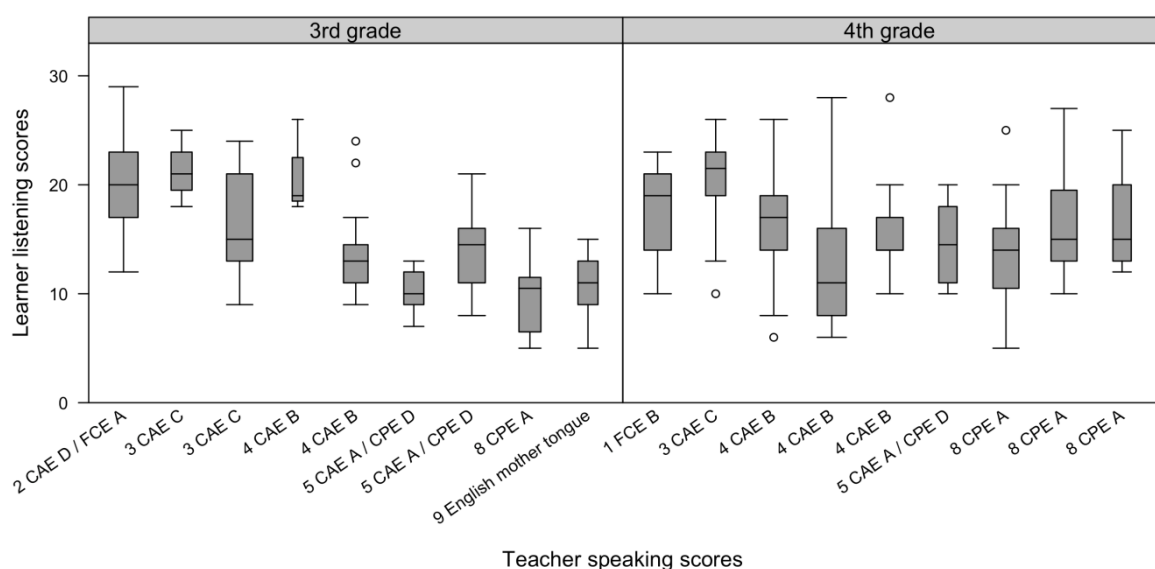
Figure 8: Association between teacher Use of English scores and learner listening scores



In this case, there appears to be a significant negative association between teachers' Use of English scores and learners' listening scores. If the teacher's Use of English score increases by one unit and the learner's grade level remains the same, the expected learner listening score decreases by 0.77 points.

**Teacher speaking scores, learner listening scores.** Again for listening, we can see the association between the teacher speaking scores and the learner listening scores (Figure 9). This coefficient for the association was found to be  $\beta_1 = -0.85$  ( $p=0.01$ ).

Figure 9: Association between teacher speaking scores and learner listening scores



Though not hypothesized to be associated, a significant negative association between teachers' speaking scores and learners' listening scores has been found. If the teacher's speaking score increases by one unit and the grade level remains the same, the expected mean learner listening score decreases by 0.85 points.

As seen in Table 19, the association between learner reading, writing and listening scores to teacher total or listening subtest scores were all negative but non-significant. Learner scores in reading all indicated a general negative association to teacher scores but were not significant. Learner writing scores were positively associated to teacher reading, listening and total scores, but none of these associations were significant.

In summary, to answer this first question, there were no associations found between teachers' scores and learner performance except in the following cases:

- Teacher speaking scores were negatively associated with learner total scores;
- Teacher Use of English scores were negatively associated with learner listening scores;
- Teacher speaking scores were negatively associated with learner listening scores.

Thus it appears that learner listening skills may be more sensitive to teacher skills than other skills measured and that the higher a teacher's speaking and Use of English skills were found to be, the more poorly these beginning classes performed.

### 4.2.2 Teacher learning curve

*Do teachers with the feeling that they are on an upward learning curve have learners who perform better in reading, writing and listening than those learners in classrooms where the teacher feels to be on a downward or stagnated learning curve?*

This research question is again related to the total language competence as well as to the five separate skill subtests of reading, writing, listening, Use of English and speaking of the teachers. For each of these subtest skills, teachers were asked if their respective skills had gotten worse, stayed the same or gotten better, an indication of their learning curve since the date of their official exam. To obtain one variable to reflect a teacher's general feeling of improvement, stagnation or losing of skills, the five questions for the change of teachers' language skills were combined to one score with three categories "1=gotten worse", "2=stayed the same" and "3=gotten better" and this score reflects the median of the five separate variables for each subcategory (worse, same or better) in Table 20. The median shows the general tendency better than the mean and thus there is more variation in the teachers' feelings of change.

*Table 20: Learner total scores grouped by teacher feeling of improvement on all skills*

Grade	Teacher change		Learner total scores				
	Feeling of skills change	N	Mean	Med	Sd	Min	Max
3rd	worse	1	40.44	41.25	9.46	24.00	62.50
	same	8	42.57	43.50	11.22	23.50	68.50
	total	9	42.24	43.00	10.96	23.50	68.50
4th	worse	1	43.45	44.75	6.98	29.00	55.00
	same	6	44.35	43.50	11.58	20.00	71.00
	better	2	50.29	52.50	10.22	25.00	65.00
	total	9	45.49	44.50	11.07	20.00	71.00
All	worse	2	42.02	44.00	8.28	24.00	62.50
	same	14	43.50	43.50	11.42	20.00	71.00
	better	2	50.29	52.50	10.22	25.00	65.00
	Total	18	44.13	44.50	11.12	20.00	71.00

Table 20 shows that no third grade teachers felt to have improved their skills. In looking at the data, two teachers mentioned improving their listening or reading skills

specifically, but then had either stagnated or worsened in the other skills. The majority of the teachers indicated that their skills had not changed since taking their respective qualifying exam. Generally, learner scores here seem to have improved when their teachers mentioned having the feeling of improvement. Learner scores on the listening, reading and writing tests in relation to the teacher's total feeling of improvement are presented in Appendix 8.5. Again, here we see that on all the subtests, the learners who scored the highest were found in classes where teachers indicated having improved the most.

The methods to investigate this research question are almost identical to those from the first question. First, the association between teachers' learning curves (in 3 categories) and learner's language competence were scrutinized through boxplots. Then, for each of the separate language skills, the following linear mixed model was calculated:

$$\text{Learner Score}_{ij} = \beta_0 + \beta_1 \cdot \text{change\_worse}_i + \beta_2 \cdot \text{change\_better}_i + \beta_3 \cdot \text{Grade\_lev\_4}_i + b_i + \varepsilon_{ij}$$

This linear mixed module formula is the same as described in Section 4.1 except that here, instead of teachers' scores as a predictor variable, two dummy variables ( $\text{change\_worse}_i$  and  $\text{change\_better}_i$ ) are used. The coefficient  $\beta_1$  quantifies the difference between the mean learner score in the category "gotten worse" and the reference category "stayed the same". The coefficient  $\beta_2$  quantifies the difference between the mean learner score in the category "gotten better" and the reference category "stayed the same". Significantly positive values for  $\beta_1$  or  $\beta_2$  would confirm the hypothesis to this question. Table 21 contains all possible combinations between different learner scores and teacher change variables.

*Table 21: Associations between teacher feeling of improvement since the official exam and learner scores*

Learner tests	Teacher tests	$\beta_0$	$\beta_1$ <i>worse</i>	$\rho$	$\beta_2$ <i>better</i>	$\rho$	$\beta_3$ 4th	$\rho$
Total	Total	41.92	-1.42	0.76	5.28	0.30	2.87	0.37
Total	Reading	41.35	-2.24	0.74	1.14	0.76	4.35	0.18
Total	Writing	41.15	1.12	0.74	4.87	0.29	2.77	0.42
Total	Listening	42.41	-4.07	0.55	-2.96	0.39	5.11	0.13
Total	UE	42.59	-2.25	0.47	5.70	0.25	3.58	0.23
Total	Speaking	39.25	0.63	0.83	<b>7.86</b>	<b>0.02</b>	3.67	0.15
Listening	Total	15.39	-1.44	0.59	3.48	0.24	-0.06	0.97
Listening	Reading	14.96	-2.41	0.55	0.76	0.69	1.00	0.59
Listening	Writing	14.59	1.17	0.55	3.03	0.26	0.08	0.97
Listening	Listening	15.59	-3.50	0.39	-1.66	0.41	1.45	0.45
Listening	UE	15.17	-0.89	0.59	<b>5.39</b>	<b>0.05</b>	0.56	0.71
Listening	Speaking	13.46	0.37	0.82	<b>5.32</b>	<b>0.00</b>	0.54	0.67
Reading	Total	15.29	0.31	0.88	2.07	0.34	0.89	0.51
Reading	Reading	15.17	-0.33	0.91	0.41	0.77	1.41	0.31
Reading	Writing	15.16	0.32	0.83	1.70	0.38	0.87	0.55
Reading	Listening	15.54	-0.94	0.75	-1.01	0.50	1.65	0.25
Reading	UE	15.52	-0.52	0.72	1.28	0.58	1.24	0.36
Reading	Speaking	14.35	0.82	0.57	2.67	0.07	1.14	0.34
Writing	Total	11.36	-0.32	0.72	-0.13	0.90	<b>1.92</b>	<b>0.01</b>
Writing	Reading	11.34	0.47	0.71	-0.05	0.94	<b>1.84</b>	<b>0.01</b>
Writing	Writing	11.42	-0.20	0.77	0.32	0.72	<b>1.73</b>	<b>0.02</b>
Writing	Listening	11.38	0.36	0.78	-0.28	0.68	<b>1.91</b>	<b>0.01</b>
Writing	UE	11.90	-0.73	0.26	-0.94	0.35	<b>1.72</b>	<b>0.01</b>
Writing	Speaking	11.43	-0.50	0.49	0.06	0.94	<b>1.94</b>	<b>0.00</b>

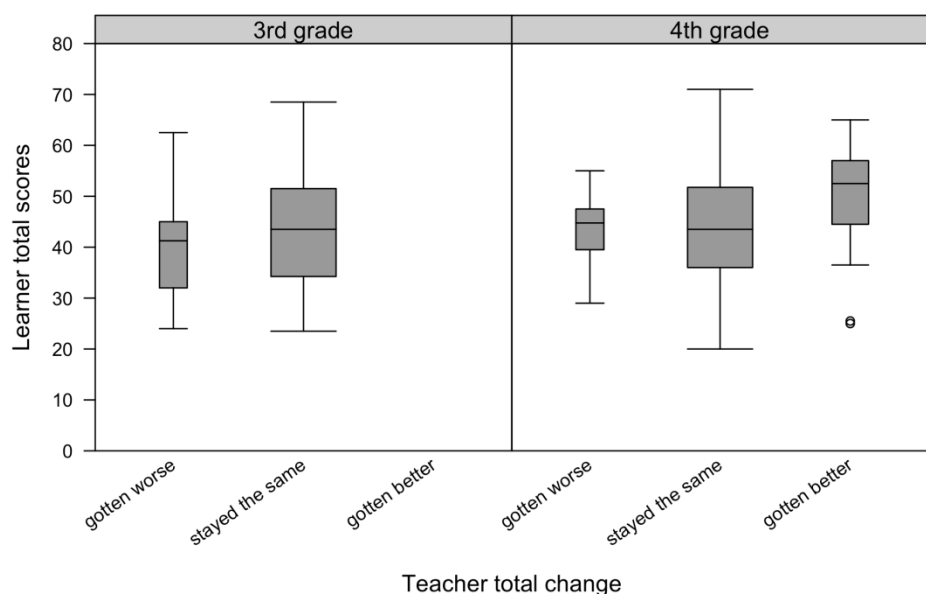
In most cases, there are no significant differences between the learner scores, though the tendencies generally go into the expected directions - a negative sign for the estimated  $\beta_1$  coefficient and a positive sign for the estimated  $\beta_2$  coefficient. This means that when teachers stated that they felt they had lost their English language skills, the learners did not perform as well as when teachers stated that they had improved their own skills. These directions were also expected in looking at the associations between each teacher's subtest score and the same learner subtest score (e.g. teacher reading and learner reading). What came out significant here was

when the teachers felt like they were on a positive learning curve in their speaking skills, then the estimated total learner and listening scores increased significantly as compared to teachers whose speaking level has remained the same since the time of the qualifying exam. This factor (teacher speaking) as well as a teacher's feeling of improving his or her Use of English skills was also positively associated with learner listening skills.

The following boxplots help to visualize the distribution of learner total scores depending on the change variable. The values for all teachers in a specific category of the change variable have been aggregated to one box. The widths of the boxes are scaled relative to the number of learners in the respective category (the narrower the box, the smaller the group size).

**Teacher total change, learner total scores.** Figure 10 clearly shows that in the 4<sup>th</sup> grade, the learners' total scores are higher if the teacher feels that his or her total language skills have improved.

Figure 10: Association between teacher total change and learner total scores



The linear mixed model calculated the corresponding coefficient for “gotten better” to be  $\beta_2 = 5.28$  ( $p=0.30$ ). This association is positive, but not significant. If the teacher total change has “gotten better” and the grade level remains the same, the expected mean learner total score increases by 5.28 points as compared to if the teacher's total skills had “stayed the same”. In the other direction, the coefficient for

“gotten worse” was found to be  $\beta_1 = -1.42$  ( $p=0.76$ ), again not significant but negative as expected.

**Teacher speaking change, learner listening and total scores.** What has been found to be significant is speaking – teachers who felt they have “gotten better” in speaking had learners who performed better generally (on the total scores) and in listening as shown through Figures 11 and 12.

Figure 11: Association between teacher speaking change and learner total scores

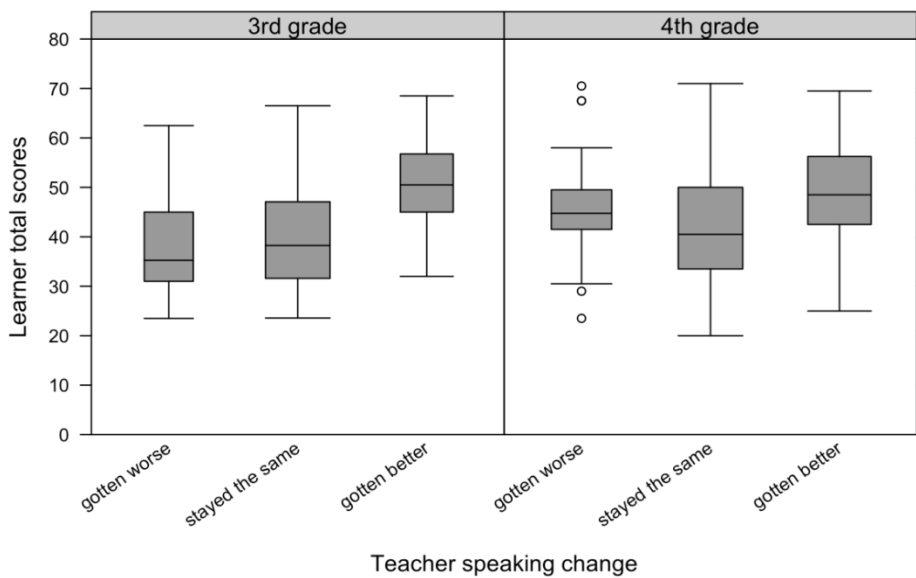
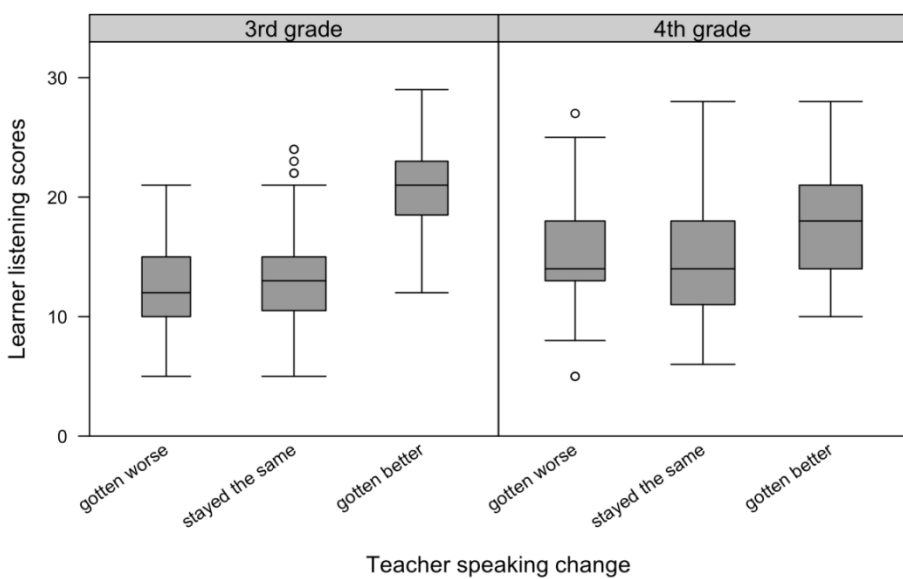


Figure 12: Association between teacher speaking change and learner listening scores



For the learner total scores, the linear mixed model the corresponding coefficient for “gotten better” was calculated to be  $\beta_2 = 7.86$  ( $p=0.02$ ) (Figure 11). For learners’

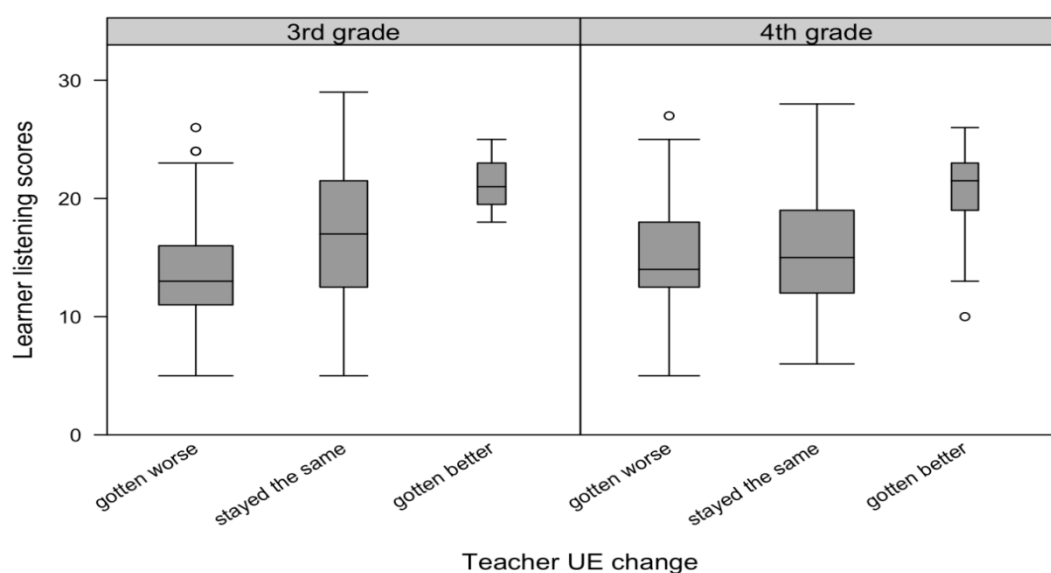


listening scores, the linear mixed model calculated the corresponding coefficient for “gotten better” to be  $\beta_2 = 5.32$  ( $p=0.00$ ) (Figure 12).

**Teacher Use of English change, learner listening scores.** The association between learners’ listening scores and teachers feeling of improvement in Use of English (Figure 13) is clearly positive and significant;  $\beta_2 = 5.39$  ( $p=0.05$ ).

When teachers feel their Use of English skills have improved, learners in this sample tended to perform better on the listening test.

Figure 13: Association between teacher Use of English change and learner listening scores



**Teacher listening change, learner listening scores.** Teacher listening change was found to be negatively associated with learner performance in all subtests, as Table 21 indicates. For third graders, as a teacher’s listening scores were stated to have improved, learner performance in listening worsened, though for fourth graders this was not the case. Similarly, the better a teacher’s listening skills had become, the worse the learners performed in reading and writing and generally on the total score for both third and fourth graders alike.

To answer the question about a teacher’s feeling of improvement, in this sample, a teacher’s feeling of improving his or her speaking skills is positively associated with learner performance in general and that this same positive association can be seen for teachers’ feelings of having improved their speaking and Use of English skills on learner listening performance. This goes in the opposite direction of the findings in

research question 1 where these same associations were negative. Therefore, for this question, the hypothesis is supported, though not confirmed. It was expected that teachers' general feelings of improvement positively correlate with learner skills on all the subtests and this is only partially the case.

### 4.2.3 Teacher exposure to English

*Do learners perform better if their teacher has had an extended stay in an English speaking country or regular contact outside the classroom with the English language?*

This question was divided into several parts due to the nature of the data. The number of years a teacher has lived abroad is measured in years; teacher total regular contact is the number of weeks on vacation per year; and regular contact is ordinal data which indicate having no contact, weekly contact and daily contact and a sum as indicating higher and lower levels of contact in reading, writing, speaking and listening.

As in the previous models, the estimated learner scores have been predicted with the following algorithm.

$$\text{Learner Score}_{ij} = \beta_0 + \beta_1 \cdot \text{years\_abroad}_i + \beta_2 \cdot \text{Grade\_lev\_4}_i + b_i + \varepsilon_{ij}$$

For the second part of this question, this analysis was repeated for the following variables quantifying the regular contact outside the classroom:

- Teacher contact with English in weeks per year
- Teacher regular contact reading (almost never, once or twice a month, daily)
- Teacher regular contact writing (almost never, once or twice a month, daily)
- Teacher regular contact speaking (almost never, once or twice a month, daily)
- Teacher regular contact listening (almost never, once or twice a month, daily)
- Teacher total contact (sum of contact in reading, writing, speaking and listening)

**Years abroad.** Table 22 shows the range of time teachers have spent abroad and the learner total scores grouped accordingly.

*Table 22: Learner total scores grouped by teacher time abroad*

Learner grade	Teacher years abroad	Learner total scores distribution					
		N	Mean	Med	Sd	Min	Max
3rd and 4th	0.08	4	42.05	40.34	11.40	23.50	70.50
	0.50	3	49.25	49.50	10.09	23.50	67.50
	0.60	1	49.64	50.50	9.82	32.00	68.50
	0.92	2	38.21	36.50	9.37	23.58	57.00
	3.00	1	41.27	40.00	10.51	25.00	63.00
	4.00	1	46.32	45.00	10.94	25.50	69.50
	7.00	1	37.50	35.09	9.22	24.80	53.60
	8.00	1	49.17	50.25	10.50	25.00	64.00
	9.00	1	43.45	44.75	6.98	29.00	55.00
	12.00	1	40.44	41.25	9.46	24.00	62.50
	19.00	2	42.01	40.50	12.72	20.00	71.00
	Total	18	44.13	44.50	11.12	20.00	71.00

From looking at this table, no general tendencies can be found, though the classes scoring the highest on the entire test were not those having teachers who lived abroad the longest.

Table 23 shows a summary of all linear mixed models calculated to investigate the question of a teacher's extended stays in an English speaking country with the number of years abroad as covariate. The first four rows correspond to the four models with adjustment for the learners' grade level. The last four rows show the results of the models without including the grade level in the model, a form of sensitivity analysis.

*Table 23: Association between teacher years abroad and learner performance*

Learner subtest	$\beta_0$	$\beta_1$ years abroad	p	$\beta_2$ 4th	p
Total	42.54	-0.28	0.25	5.19	0.10
Listening	15.62	-0.14	0.32	1.31	0.47
Reading	15.72	-0.13	0.19	1.89	0.15
Writing	11.36	-0.01	0.80	<b>1.95</b>	<b>0.00</b>
Total	44.56	-0.14	0.56		
Listening	16.12	-0.11	0.42		
Reading	16.47	-0.08	0.40		
Writing	12.10	0.04	0.46		

No significant effects are found between the number of years a teacher spent abroad in a country where English was the medium of communication and the learners' language skills. In all skills except for learner writing,  $\beta_1$  was also found to be negative. This means that within this sample, the learners' language scores in reading and listening decreased the longer a teacher lived abroad, in the opposite direction as was expected. However, these associations are not significant.

**Teacher regular contact on yearly vacations.** The second analyses investigated the total weekly contact teachers have with English outside the classroom during vacation time and their estimated regular weekly contact. As described in the survey results in Chapter III, teachers spent anywhere from one to five weeks abroad per year in a place where they used their English during their vacation time.

Table 24 shows the results of the model using the predictor variable of teacher holiday contact in weeks a year. The first four rows correspond to the four models with adjustment for the grade level. The last four rows show the results of the models without including the grade level in the model.

*Table 24: Association between teacher holiday contact and learner scores*

Learner subtest	$\beta_0$	$\beta_1$ contact weeks / year	p	$\beta_2$ 4th grade	p
Total	41.64	0.13	0.89	3.95	0.21
Listening	15.48	-0.32	0.58	0.95	0.60
Reading	15.09	0.29	0.47	1.11	0.40
Writing	11.19	0.17	0.36	<b>1.76</b>	<b>0.00</b>
Total	43.34	0.43	0.65		
Listening	15.88	-0.24	0.65		
Reading	15.58	0.38	0.33		
Writing	11.95	0.30	0.17		

Again, no significant effects were found between the number of contact weeks per year on vacation and the learners' language competences.  $\beta_1$  was found to be positive except in the case of listening. This means that the learners' performance on the test was higher with teachers having regular holidays where English is used. Again, none of the associations are significant.

**Teacher regular contact in reading, writing, speaking and listening.** In Table 25, we see the descriptive data for the total scores of the learners when grouped by teacher estimated contact time in reading, writing, speaking and listening outside of the classroom. These data are for the total of all learners, not grouped by grade level.

*Table 25: Learner total scores grouped by teacher frequency of contact in specific language skills*

Teacher contact in...	Teacher contact	Almost never			Monthly			Daily		
	Learner total scores	Mean	Med	SD	Mean	MD	SD	Mean	Med	SD
...Listening		44.85	49.00	11.14	45.74	45.50	10.72	43.51	43.50	11.30
...Reading		44.85	49.00	11.14	45.79	45.50	10.83	43.44	44.00	11.21
...Speaking		45.72	45.60	10.99	46.48	46.00	10.80	41.11	41.00	10.87
...Writing		47.07	47.75	10.56	43.66	44.00	10.65	42.99	42.00	11.58

Table 25 shows that the learners of those teachers with daily contact in listening, reading, speaking and writing in English scored lower than learners in classes where a teacher's daily contact was not as frequent. This does not mean that these teachers had contact in all four skills, but rather that they mentioned using each skills separately on a daily basis (for the sum of their total contact, see Table 27).

Thus Table 26 shows the teacher regular contact in reading, writing, speaking and listening with adjustment for the grade level of the learners. Similarly to the second research question (teacher feeling of improvement), for each variable there are two effects estimated since the regular contact variables each have three categories. Furthermore, there is the additional effect of the learners' grade levels.

*Table 26: Association between teacher regular weekly contact and learner performance*

Learner score	Teacher contact	$\beta_0$	$\beta_1$ once/twice monthly	p	$\beta_2$ every day	p	$\beta_3$ 4th grade	p
Total	Listening	44.85	-1.78	0.78	-6.22	0.36	6.26	0.07
Total	Reading	44.85	-1.68	0.82	-4.03	0.55	4.20	0.20
Total	Speaking	43.33	0.35	0.94	-5.09	0.27	4.60	0.11
Total	Writing	45.27	-4.21	0.27	-7.06	0.09	5.86	0.07
Listening	Listening	16.77	-0.27	0.94	-4.14	0.27	2.49	0.17
Listening	Reading	16.77	-0.07	0.99	-2.24	0.57	0.69	0.71
Listening	Speaking	15.95	0.58	0.82	-3.08	0.25	1.11	0.48
Listening	Writing	17.52	-2.58	0.23	<b>-5.08</b>	<b>0.03</b>	2.10	0.22
Reading	Listening	16.06	-0.54	0.85	-1.31	0.66	1.77	0.25
Reading	Reading	16.06	-0.38	0.90	-0.97	0.74	1.38	0.32
Reading	Speaking	15.97	-0.02	0.99	-1.81	0.37	1.52	0.23
Reading	Writing	16.05	-0.77	0.66	-1.66	0.37	1.81	0.20
Writing	Listening	12.02	-0.85	0.52	-0.75	0.58	<b>1.96</b>	<b>0.01</b>
Writing	Reading	12.02	-0.96	0.49	-0.77	0.55	<b>2.03</b>	<b>0.01</b>
Writing	Speaking	11.44	-0.07	0.94	-0.24	0.80	<b>1.92</b>	<b>0.01</b>
Writing	Writing	11.95	-1.07	0.17	-0.56	0.48	<b>1.92</b>	<b>0.01</b>

There is only one significant association detected: between teachers' regular writing contact and the listening scores of the learners and this is significantly negative. Thus the teachers in this study who have regular contact writing English had learners whose performance in listening is lower than their peers in a classroom with a teacher who has less regular speaking contact time. As with research question one (tested teacher language skills), most associations are negative – the more teachers have contact outside of the classroom with English, the poorer the learners performed in this sample, though again, very little is significant.

Figures 13 through 17 visualize these general trends. Though no amount of contact with English on the level of the teacher is associated with learner performance on the total score, the learner total scores generally decreased with increasing teacher contact.

Figure 14: Association between teacher listening contact and learner total scores

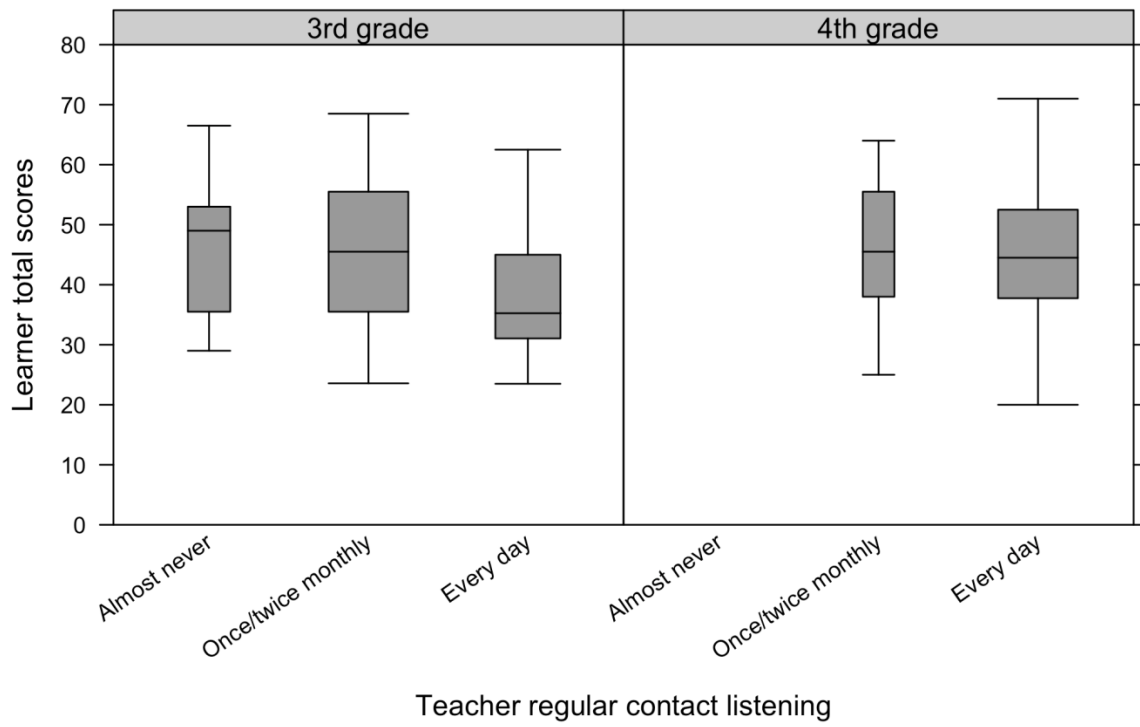


Figure 15: Association between teacher reading contact and learner total scores

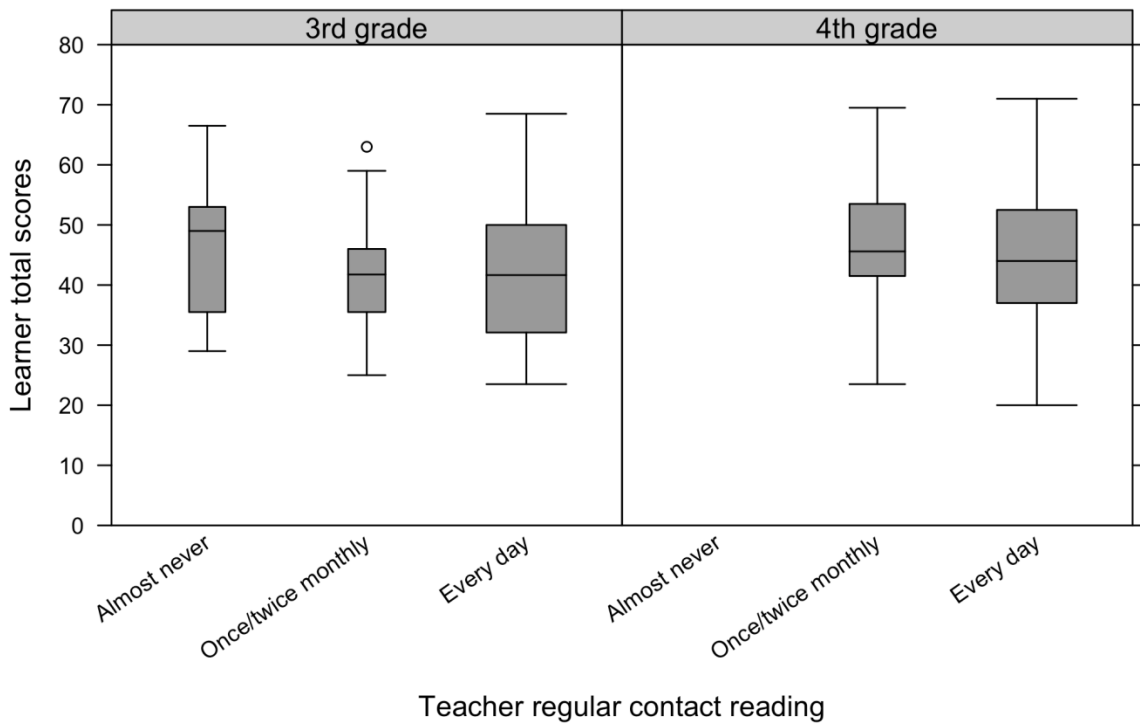




Figure 16: Association between teacher writing contact and learner total scores

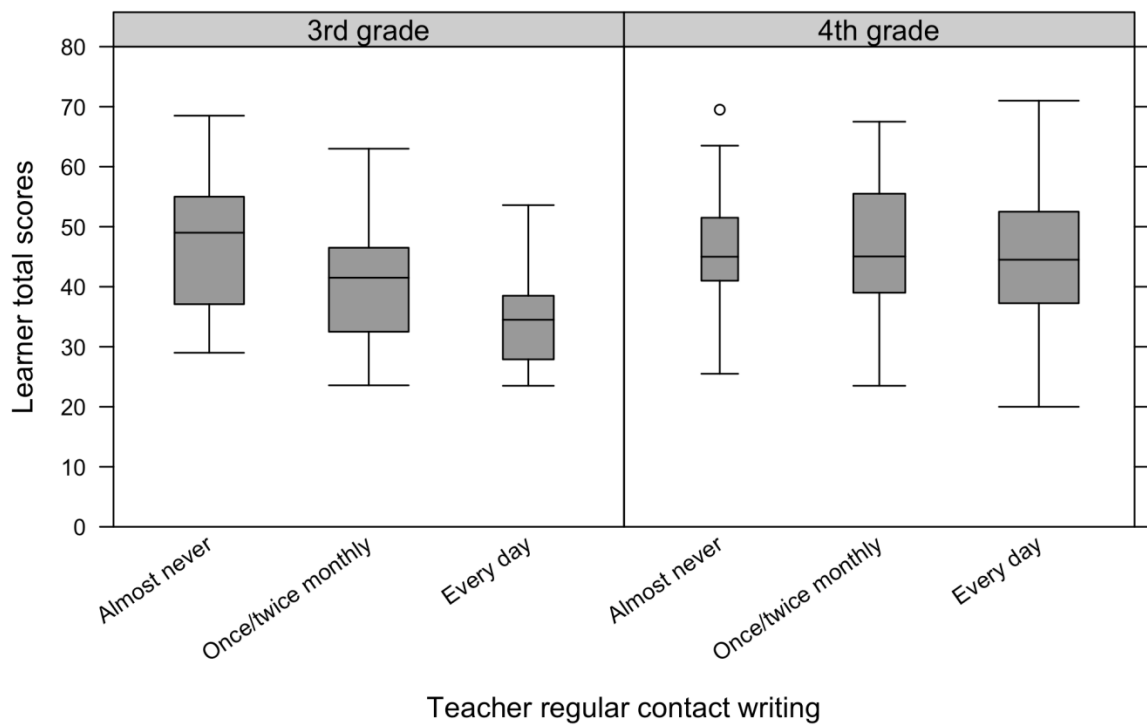
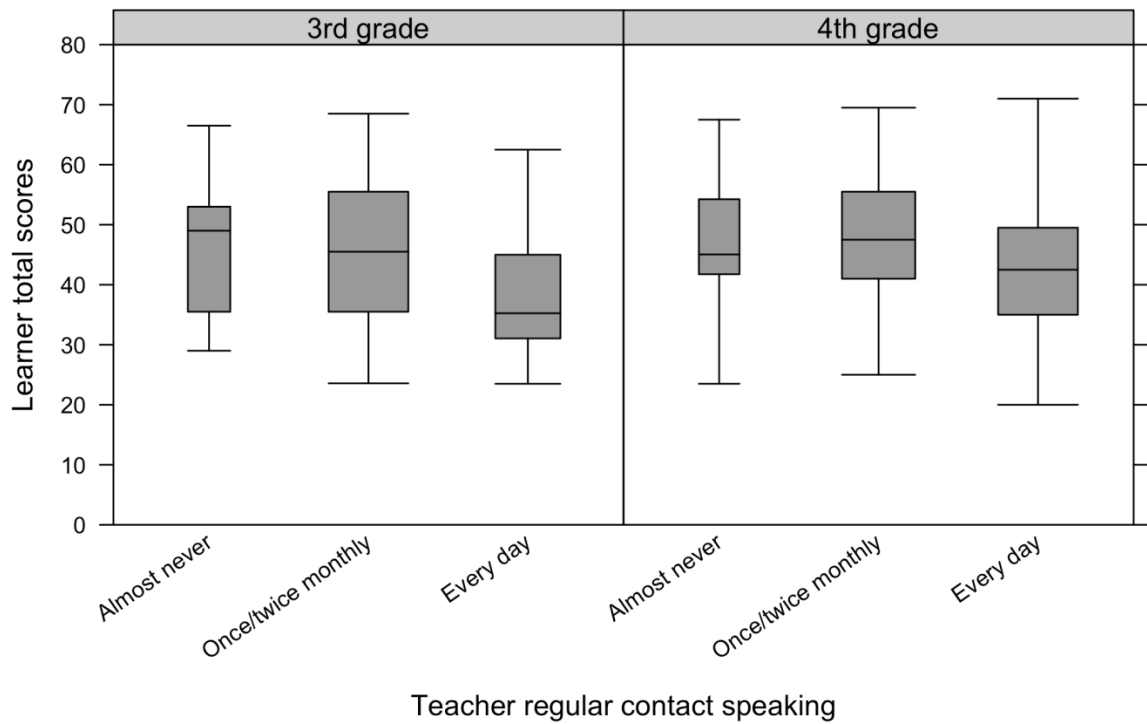


Figure 17: Association between teacher speaking contact and learner total scores



Through the boxplots in Figures 13 through 17, one can clearly see the slightly negative association between teacher contact and learner performance. Though perhaps not as dramatic for fourth grade, this tendency within the sample population here is obvious for all the learner subtests.

Table 27 indicates that the learners of teachers with the least contact did not necessarily have lower scores than those learners of teachers with more contact. This contact number is ordinal and the sum of teacher contact in reading, writing, speaking and

Table 27: Teacher total contact and learner total scores

Teacher total contact	Learner total score		
	Mean	Med	Sd
4	44.85	49.00	11.143
7	46.83	46.50	11.504
8	46.67	45.60	10.722
9	42.76	43.00	10.538
11	46.19	45.59	11.116
12	41.23	41.00	11.151

listening, thus a teacher with a “12” is one who reads, writes, speaks and listens on a daily basis. A teacher with a “4” is one who has slightly more contact in one skill than the others but generally has no regular exposure outside of the classroom.

Table 28 displays the results for the teacher total contact (sum of the reading, writing, speaking and listening contact variables). The first four rows correspond to the four models with adjustment for the grade

Table 28: Association between total teacher regular weekly contact and learner performance

Learner score	$\beta_0$	$\beta_1$ total contact	$p$	$\beta_2$ 4th grade	$p$
Total	51.53	-1.10	0.10	5.49	0.07
Listening	22.22	<b>-0.79</b>	<b>0.04</b>	1.76	0.28
Reading	17.89	-0.28	0.33	1.72	0.20
Writing	11.79	-0.05	0.69	<b>1.97</b>	<b>0.00</b>
Total	50.96	-0.74	0.27		
Listening	22.03	-0.67	0.06		
Reading	17.70	-0.17	0.53		
Writing	11.49	0.09	0.59		

A significant negative association between the teacher total regular contact with English outside of the classroom and the learners' listening scores was found, thus the more frequent overall contact the teacher noted having with English outside the classroom, the lower the learner listening test scores were. This is the only

significant association though the negative coefficients indicate that the more contact the teacher has, the worse the class performance.

It was hypothesized that teacher contact with English outside of the classroom would be positively associated with learner performance and it was not. Indeed, quite the opposite was found. It would appear that the number of years a teacher has lived abroad in an English speaking country and his or her yearly holiday time in an English speaking country are not associated with learner performance. Similarly, regular contact with English outside the classroom in reading, writing, speaking and listening as well as generally is not associated with learner performance except in the case of associations with learners' listening scores: the more writing contact or total contact the teacher has (as defined by being almost daily), the worse the learners performed on the listening test.

#### 4.2.4 Teacher classroom skills profile

*(1) Does the estimated amount of time a teacher speaks German in class detract from learners' performance in reading, writing and listening?*

*(2) Does the estimated amount of time a teacher spends teaching reading, writing, speaking and listening in English interact directly with the learners' performance in each of these same skills?*

**Method:** To investigate this research question, two teacher variables were used: (1) the amount of time a teacher speaks German per forty-five minute lesson and (2) the teacher type which has the following three categories:

- “Traditional”: Teachers who tend to be more traditional – doing reading, writing, and listening activities, but not many speaking activities with the class.
- “Balanced”: Teachers who focus more or less equally on reading, writing, speaking and listening activities in class.
- “No writing”: Teachers who do no writing in class.

Four models for the four different learner scores (reading, writing, listening and total score) were used to see the association of the amount of time a teacher speaks German per lesson:

$$\text{Learner Score}_{ij} = \beta_0 + \beta_1 \cdot \text{DE\_time\_per\_lesson}_i + \beta_2 \cdot \text{Grade\_lev\_4}_i + b_i + \varepsilon_{ij}$$

Two models with the teacher type and reference category ‘Balanced’ (the variables ‘Traditional’ and ‘No\_writing’ are binary dummy variables) were used:

$$\text{Learner Score}_{ij} = \beta_0 + \beta_1 \cdot \text{Traditional}_i + \beta_2 \cdot \text{No\_writing}_i + \beta_3 \cdot \text{Grade\_lev\_4}_i + b_i + \varepsilon_{ij}$$

**Amount of time in German per lesson:** Table 29 provides the breakdown of the various teachers and their time in speaking German as it is associated with learner scores. Both third and fourth grade teachers were grouped together as there does not seem to be an effect of grade level.

Table 29: Learner total scores based on teacher time in German per lesson

Minutes of German	Learner total scores distribution					
	N	Mean	Med	Sd	Min	Max
2	3	42.67	43.00	11.67	23	71
5	2	50.51	51.00	10.00	25	71
6	1	49.17	50.25	10.50	25	64
7	3	46.64	46.25	10.14	24	69
8	1	40.69	39.50	10.60	20	70
10	3	40.83	41.25	8.58	24	57
15	1	37.50	35.09	9.22	25	54
30	1	44.85	49.00	11.14	29	67
Total	18	44.13	44.50	11.12	20	71

Though it would appear in Table 29 that the teachers using the least amount of German do not necessarily have the highest performing learners, when looking at the teachers who use a lot of German, the scores are even lower. With the majority of teachers being somewhere in the middle, we can see that some of the teachers using the least amount of German (5, 6, or 7 minutes) do have the higher performing learners.

Table 30 shows a summary of all linear mixed models that were calculated to investigate the association between the time a teacher spends speaking German per forty-five minute lesson and the learners' scores.

Table 30: Association between teacher time in German and learner performance

Learner score	$\beta_0$	$\beta_1$ Minutes in German	$\rho$	$\beta_2$ 4th grade	$\rho$
Total	41.65	0.01	0.97	4.10	0.21
Listening	15.43	-0.02	0.90	0.64	0.74
Reading	15.53	-0.02	0.86	1.26	0.36
Writing	10.90	0.04	0.42	<b>2.07</b>	<b>0.00</b>
Total	44.72	-0.11	0.67		
Listening	15.89	-0.04	0.79		
Reading	16.50	-0.06	0.59		
Writing	12.45	-0.02	0.76		

There seem to be no significant effects between the amount of time a teacher speaks German and the learners' language competence; indeed  $\beta_1$  is negative for almost all models. This means that the learners' test performance is calculated to decrease with an increasing amount of German used during the lessons (by the teacher), which was to be expected. However, none of the associations are significant.

**Teacher type:** In Table 31, the learner scores categorized by teacher type are presented.

*Table 31: Descriptive statistics for teacher types*

Grade	Teacher type	Learner total scores distribution					
		N	Mean	Med	Sd	Min	Max
3rd	Traditional	3	40.72	38.00	10.83	23.58	61.00
	Balanced	2	42.58	43.00	10.40	24.00	66.50
	No writing	4	42.94	44.00	11.51	23.50	68.50
	Total	9	42.24	43.00	10.96	23.50	68.50
4th	Traditional	2	45.52	44.50	10.24	25.50	69.50
	Balanced	2	49.78	48.75	10.00	25.00	71.00
	No writing	5	43.97	43.50	11.38	20.00	70.50
	Total	9	45.49	44.50	11.07	20.00	71.00
All	Traditional	5	43.12	43.50	10.73	23.58	69.50
	Balanced	4	46.13	46.50	10.76	24.00	71.00
	No writing	9	43.61	43.50	11.40	20.00	70.50
	Total	18	44.13	44.50	11.12	20.00	71.00

It appears that teachers having a “balanced skills” or “traditional” approach to teaching in this study tended to have learners who performed better than teachers having a “no writing” approach to teaching. For fourth graders, the teachers with a balanced skills approach tended to have the highest scoring learners.

Table 32 shows a summary of the linear mixed models that were used to investigate the association between the teacher type and the learners' scores.

Table 32: Association between teacher type and learner performance

Learner score	$\beta_0^*$	$\beta_1^{**}$ traditional	$p$	$\beta_2^{***}$ No writing	$p$	$\beta_3$ 4th grade	$p$
Total	44.11	-3.39	0.43	-2.90	0.45	4.06	0.19
Listening	16.62	-2.31	0.36	-1.45	0.52	0.63	0.72
Reading	15.94	-1.02	0.58	-0.68	0.68	1.33	0.32
Writing	11.68	-0.13	0.87	-0.75	0.29	<b>2.00</b>	<b>0.00</b>
Total	46.12	-3.73	0.40	-2.53	0.52		
Listening	16.93	-2.36	0.34	-1.41	0.52		
Reading	16.60	-1.11	0.55	-0.54	0.74		
Writing	12.64	-0.26	0.80	-0.53	0.57		

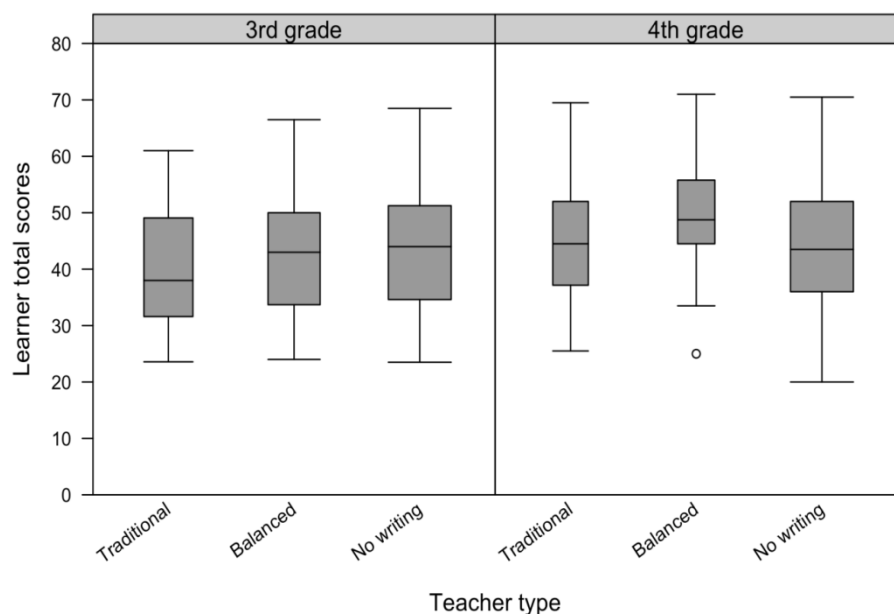
\*mean learner score if the teacher type is “balanced”

\*\* difference between a “traditional” and a “balanced” teacher

\*\*\*the difference between a “no writing” and a “balanced” teacher

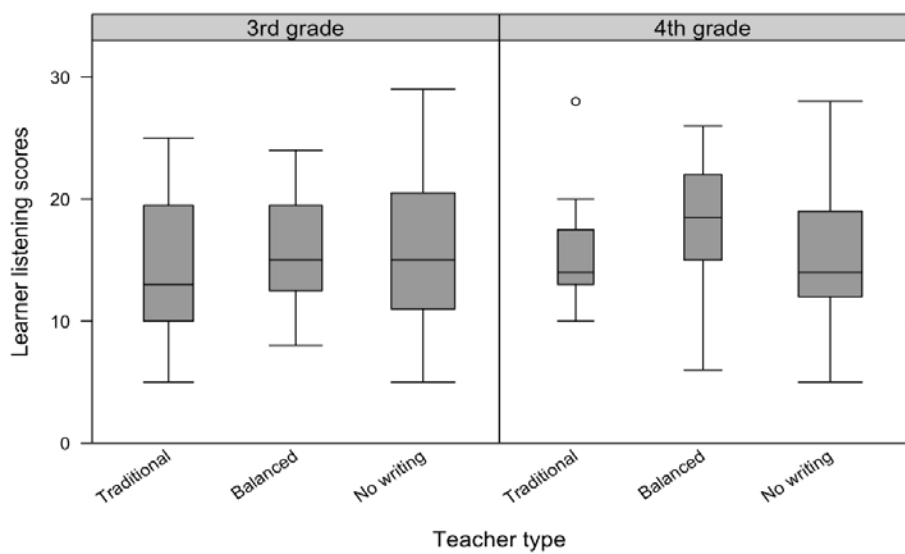
$\beta_1$  and  $\beta_2$  were found to be negative for all the models. This means that within the population of this study, the learners’ language scores are lower when the teacher is not using “balanced” techniques, though the associations are not significant. Figures 18 through 21 visualize the associations between the teacher’s type of teaching and the learners’ performances.

Figure 18: Association between teacher balance of skills and learner total scores



It appears in Figure 18 that being a teacher with a balanced skills approach to teaching English might be the largest predictor of learner performance in the fourth grade. In the third grade, the differences between the approach to teaching and learner performance is not quite as dramatic and neglecting writing does not seem to impact learner performance.

Figure 19: Association between teacher balance of skills and learner listening scores



Again, in Figure 19, we clearly see the positive relation between being a “balanced” teacher and learner performance in the fourth grade. Having a “traditional” or “no writing” approach to teaching English in the third and fourth grades does not seem to be associated with high level results.

Figure 20: Association between teacher balance of skills and learner reading scores

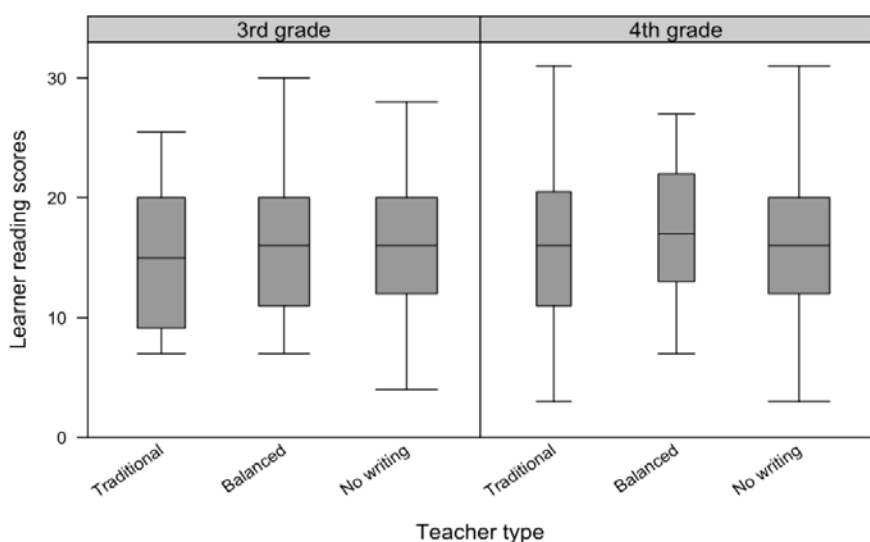
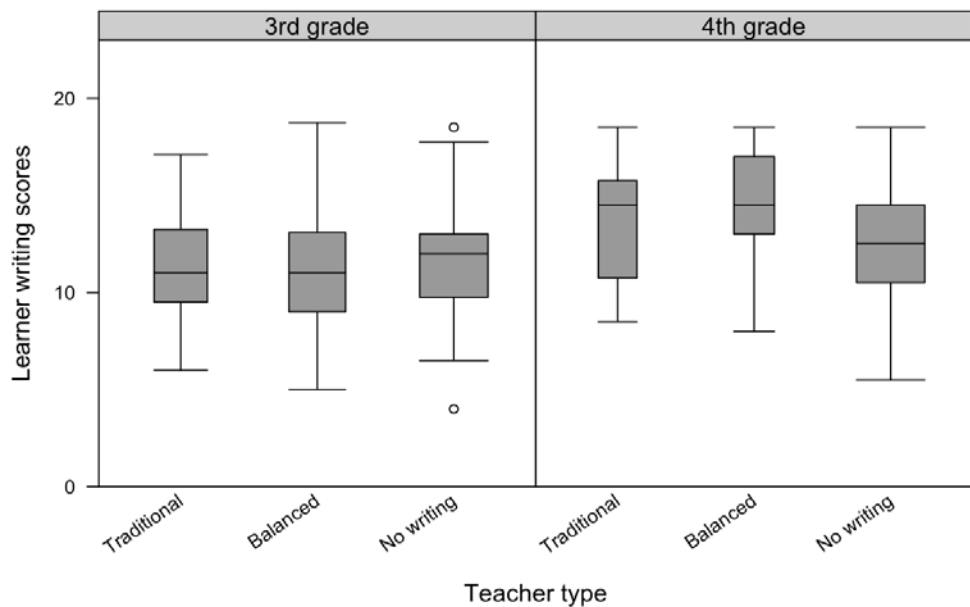




Figure 21: Association between teacher balance of skills and learner writing scores



We see similar results in Figures 20 and 21, though the association to being a balanced skills type of teacher does not seem to have as strong of an association with the reading and writing performance of the learners for learner writing and reading scores as for listening and total scores. For the learner reading and writing scores, the difference between the three different approaches is not as dramatic.

Generally, it would appear that listening tends to be more sensitive to teacher “style” or “type” than the other skills tested. Looking at the descriptive statistics, it would appear that the negative effect of non-balanced instruction is more pronounced for fourth graders. In the fourth grade, being a teacher who spends equal time on each of the four language skills is more often associated with a general higher learner score. Neglecting writing in the fourth grade also seems to be associated with lower learner scores.

The amount of German used in the classroom and the general style of teaching as defined by time on certain language skills do not appear to have any association with learner performance, though being a balanced-skills teacher (equal time teaching reading, writing, speaking and listening) does tend to be associated with learners with better scores, albeit not significantly higher. Thus the original hypothesis here cannot be confirmed and the findings are inconclusive.

### 4.2.5 Summary of individual regressions

Teacher proficiency levels on the Use of English and reading parts of the qualifying exam were expected to be positively associated with learner performance and the results indicate that thus these hypotheses must be rejected. Teacher scores on the Use of English were negatively associated with learner performance on the learner total and listening scores, and teachers speaking scores were negatively associated with learner listening scores. Reading scores, though predicted to be associated with learner performance, were not. Teacher writing and listening scores were not significantly associated with learner performance, as was predicted.

Teachers who felt like they were on an upward learning curve in speaking had learners who performed better on the listening and total score parts of the tests when compared to their peers with teachers who had stagnated or not improved their speaking since the time of the official exam. The hypothesis was that there would be direct correlations amongst teacher and learner skills, which was not the case. Furthermore, feelings of improving speaking skills and Use of English skills were also positively associated with learner's listening scores. All other associations were not significant, thus the hypotheses can neither be accepted nor rejected.

Teacher contact with English outside the classroom also led to very few significant results. It appears that stays abroad had no significant impact on learner performance, as was hypothesized. The only significant association was between teacher contact in writing and teacher total contact and learner performance in listening, and this was shown to be negative which is in the opposite direction as hypothesized. Teacher regular exposure outside of the classroom does not seem to interact with learner performance.

Finally, neither the time on skills profile of the teacher nor the time in German seemed to have any association with learner performance. It seems that being a balanced skills teacher (where reading, writing, speaking and listening are of equal importance), a teacher who neglects writing, or a teacher with a focus on reading, listening and writing skills has no effect on learner performance in the classroom in this study. Though the findings are slightly positive in both instances, no findings here were significant.

The following section, Section 4.3, sets out to see which of the predictor variables looked at individually in Section 4.2 might be important when all the variables are looked at together. Here the results will be analyzed in an attempt to discover which variables on the part of the teacher, controlling for the grade of the learners, have the largest association with learner performance on the reading, writing and listening subtests as well as with the learners' total scores.

### 4.3 Results for a combination of predictors – linear mixed models

#### 4.3.1 Method

The question here was “Which combinations of variables on the side of the teacher, controlling for the years of English the learner has had, contribute the most to learner performance”? The idea here is to see if the variables from the individual models described in section 4.2 are still important when several other variables - both the same as in the individual models and a few additional predictor variables – are introduced and if a profile of teacher characteristics can be made which are conducive to learner performance. For the full models, random intercept linear mixed models were used in order to investigate the associations between the learners’ total and subtest scores and the many teacher factors. The aim is to find out which combinations of teacher factors contribute the most to learner performance and also to gain more insight into the complex interplay of the various factors.

The random forest model was not used as it does not adjust for the repeated measurements (teacher data) in the data structure. Furthermore, as random forests do not estimate coefficients, it would be impossible to see what is significant between the covariate (teacher factors) and the response variable (learner scores). A common linear regression model would have been misleading because in this data set, the observations are not mutually independent – some observations are dependent on one another and a common regression model would not account for this though a random intercept regression model would.

Several variables were included in these analyses that were not included in the individual models in Section 4.2. Firstly, a teacher’s total number of years of teaching was included because in the literature (e.g. Buddin and Zamarro, 2009), it was seen at times as contributing to learner performance. Secondly, variables related to knowledge about language were included here – a teacher’s number of languages spoken is an indication of ability to learn language and the more the languages spoken, perhaps the higher the teacher’s metalinguistic awareness as authors such as Andrews (2001, 2003) imply. In this same sense, having learned English explicitly or implicitly might be a hint at how much a learner knows about English that they explain to learners. This might also be evidence for Andrews’ discussions of the importance of teacher language awareness.

The process of building a model was challenging for several reasons, namely that four models for each of the four learner scores had to be created. Furthermore, there are a large number of teacher variables and although the hypotheses from the literature review was to have helped in indicating certain tendencies, the selection of variables was difficult due to the varying types of data (ordinal as well as binary). Moreover, individual variables from several variable categories are highly correlated with one another (teacher scores, teacher change variables), thus leading to a high risk of multicollinearity – in other words, including the same information twice or several times in the model.

Table 33 gives an overview of which variables from the teacher survey could potentially have been included in the models and which variables were actually selected to be included. The decision to select certain variables and deselect others was complicated. The original idea would have been to see specifically which skills on the side of the teacher interact with the same skills in the learner performance (for example the teachers' feelings of improving their listening skills and the learners' listening performance). However, due to the nature of the data and to simplify the interpretations, it was decided to use the same set of predictor variables for all learner subtests.

Thus, for each learner subtest score and the total score, a full model was created with at least one variable from each of the variable groups. In the case of the variable groups with high correlations (e.g. teacher score on each of the subtests), only one variable per group was included in order to avoid collinearity problems in the regression model.

Table 33: Variable selection for the final mixed model

Variable group & variables	Description / Remarks	Included?
<b>Learner factors</b>		
Grade level (years EN)	Categorical with 2 categories; inclusion of one dummy variable with reference "3rd grade"	Yes
<b>Teacher indicators of (life) experience</b>		
Total years teaching		Yes
Total years English teaching	Not relevant as English is a new subject	No
Number of languages spoken	Categorical with 3 categories (3,4,5), inclusion of two dummy variable with reference "Number languages=3"	Yes
Learning technique	Categorical with 2 categories (explicit or implicit), inclusion of one dummy variable with reference category "implicit"	Yes
Multilingual	Number of languages is more telling than the fact of being monolingual or not.	No
<b>Teachers indicators of teaching characteristics</b>		
Minutes German per lesson		Yes
Teacher type / approach	Categorical with 3 categories, inclusion of two dummies with reference "Approach = balanced"	Yes
Rituals	Not included as only three teachers at the same time grade level (4 <sup>th</sup> ) answered with "rituals=no"	No
<b>Teacher scores on official qualifying exam</b>		
Teacher total scores	Scores are very highly correlated (positively). It does not make sense to include all variables since the same information would thus be included twice. Only teachers' total scores are included since these have the highest correlations with all of the other scores	Yes
Teacher listening scores		No
Teacher reading scores		No
Teacher speaking scores		No
Teacher UE scores		No
Teacher writing scores	variables and are a good indication of the general language level of the teachers.	No
<b>Teacher indicator of level change since exam</b>		
Total change	All variables here had three categories "gotten worse", "stayed the same" and "gotten better". For "total change" the distribution was not well-spread. Thus only speaking change was chosen because this skill is probably the most used in the classroom, had the highest variance amongst the teachers and best represents the general tendency of the individual teacher.	No
Listening change		No
Reading change		No
Speaking change		Yes
UE change		No
Writing change		No
<b>Teacher contact with English outside of school</b>		
Number of years abroad		Yes
Average weekly contact / year		Yes
Teacher contact listening	These variables had three categories "Almost never, "Once/twice monthly" and "every day". The first two categories were aggregated to one resulting in a binary variable "daily contact" vs. "no daily contact" which left little variation. Furthermore, in scrutinizing the data, the teachers who had the highest amount of regular contact were mainly the ones who also had the most weeks per year in an English speaking country and weeks is a better measure than "more or less".	No
Teacher contact reading		No
Teacher contact speaking		No
Teacher contact writing		No
Teacher total contact		No

After determining which variables to include, starting from the full models, the teacher variables and the one learner variable (3<sup>rd</sup> or 4<sup>th</sup> grade) were thrown into the model where a backward selection based on an AIC (Akaike Information Criterion) measure was performed in order to remove variables with large AICs (poor model fit) from the modelling because small values represent better fits. Like this, only those variables which contributed to learners' scores were left in what will be called the reduced model as they make the best predictor variables. This reduced the number of variables to those that contribute the most to the entire model<sup>31</sup>. For this selection, the maximum likelihood (ML) method was used in hopes of finding the best model with the smallest AIC based on ML estimation.

For each of the four different learners' scores (total, listening, reading and writing scores), a random intercept linear mixed model was used as because it was decided to look at the explanatory variables yet at the same time control for teacher variables. The basic equation is as follows and is similar to that described in the introduction to Chapter IV:

$$\text{Learner Score}_{ij} = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + \dots + \beta_k \cdot x_{ik} + b_i + \varepsilon_{ij}$$

To reiterate:

- $i$  and  $j$  stand for the indices for teacher  $i=1, \dots, 18$  and for learner  $j=1, \dots, n_j$
- The intercept  $\beta_0$  can be interpreted as the mean learner score when the all covariates  $x_{i1}, \dots, x_{ik}$  are set to zero.
- The covariates  $x_{i1}, \dots, x_{ik}$  are the teacher-specific characteristics. Since these variables are all constant across a teacher (i.e. each learner of the same teacher who has the same value for  $x_{i1}, \dots, x_{ik}$ ), an additional index ' $j$ ' for the covariate is not necessary.
- The coefficients  $\beta_1, \dots, \beta_k$  are "fixed effects" and quantify the linear effects of teacher factors  $x_{i1}, \dots, x_{ik}$  on learner's language competence. Positive coefficients indicate a positive association between the teacher factor and learner score, negative coefficients a negative association.
- For categorical teacher factors with  $k$  categories, such as teacher type, learning technique or level change, one reference category is fixed and  $k-1$  binary dummy variables for the other  $k-1$  categories are included in the model.

---

<sup>31</sup> See Burham & Anderson (2002) for an in-depth discussion or Noble (2014) for a more concise overview.

- A teacher-specific random intercept  $b_i$  is included. This random intercept is assumed to have a standard Gaussian distribution (mean zero). The association between teachers' and learners' language skills is adjusted for the complex data structure with repeated measurements per teacher.
- $\varepsilon_{ij}$  denotes the error term of the model and is also assumed to have a standard Gaussian distribution (mean zero) and equal variances for each observation (homoscedasticity).

The coefficients  $\beta_1, \dots, \beta_k$  are unknown and are calculated through restricted maximum likelihood estimation (REML). REML is supposed to yield better estimates for the random effects and this method is less biased (more stringent) in that it takes  $df-1$  with small sample sizes. It produces estimations of coefficients and thus  $p$  values can be calculated (see Hox, 2010, p 41).



### 4.3.2 Results

Table 34 shows calculated regression coefficients with the corresponding standard errors in parentheses below for each of the mixed models. Significant coefficients are printed in bold. Both the full and the reduced models are shown, or before and after the AIC backward selection. Both the sign (positive or negative) and the significance can be evaluated and the variables which remain in the model after backward selection can also be scrutinized. Although not always significant, these results provide an idea of which combinations of variables are linked to learner scores.

The AIC totals at the bottom of Table 34 indicate the difference between the full model and the reduced model. What is presented in the reduced model is the combination of predictor variables that contribute the most to the model. Adding further variables would not add anything to the model. Furthermore, the standard deviation of the random intercepts is presented and shows much smaller values in the reduced than in the full models; additional predictor variables increase the standard deviation to quite a degree.

Table 34: Full and reduced multivariable mixed models

	Learners total scores		Learners listening scores		Learners reading scores		Learners writing scores	
	Full	Reduced	Full	Reduced	Full	Reduced	Full	Reduced
<i>Coefficients and standard errors (in parentheses)</i>								
Constant	44.03 (13.50)	45.39 (3.56)	11.64 (5.77)	16.14 (1.68)	18.66 (6.72)	21.25 (3.40)	14.10 (1.61)	12.31 (.57)
<b>Learner factors</b>								
4th grade	5.36 (4.94)	5.10 (2.52)	2.61 (2.12)	1.52 (1.19)	1.17 (2.49)		1.56 (0.66)	<b>2.10**</b> <b>(0.49)</b>
<b>Teacher indicators of (life) experience</b>								
Total years teaching	-0.03 (0.47)		0.12 (0.20)		-0.13 (0.25)	-0.17 (0.08)	-0.05 (0.06)	
Number languages spoken = 4 vs. 3	0.75 (7.48)		-0.68 (3.20)		1.81 (3.76)	2.13 (1.35)	-0.42 (0.96)	
Number languages spoken = 5 vs. 3	1.92 (7.33)		-1.11 (3.14)		2.34 (3.73)	<b>3.40*</b> <b>(1.31)</b>	0.77 (1.02)	
Learning technique = implicit (vs. explicit)	1.52 (7.54)		-0.77 (3.23)		0.29 (3.83)		2.18 (1.03)	<b>1.65*</b> <b>(0.59)</b>
<b>Teachers indicators of teaching characteristics</b>								
Minutes German per lesson	-0.04 (0.49)		0.11 (0.21)		-0.09 (0.25)	-0.16 (0.08)	-0.07 (0.06)	
Approach: no writing (vs. bal. skills)	-6.71 (7.64)	-5.30 (3.11)	-0.68 (3.27)	-1.50 (1.56)	-2.98 (3.83)	<b>-3.85*</b> <b>(1.27)</b>	<b>-3.34*</b> <b>(0.89)</b>	<b>-2.40**</b> <b>(0.69)</b>
Approach: traditional (vs. bal. skills)	-9.07 (8.26)	-7.11 (3.61)	-3.10 (3.53)	-3.42 (1.72)	-3.43 (4.14)	<b>-3.61**</b> <b>(1.22)</b>	<b>-2.85*</b> <b>(1.01)</b>	<b>-2.19*</b> <b>(0.86)</b>
<b>Teacher scores on official qualifying exam</b>								
Teacher total scores	0.51 (1.90)		0.42 (0.81)		0.03 (0.97)		0.06 (0.28)	
<b>Teacher indicator of level change since exam</b>								
Speaking = gotten worse	-1.10 (11.22)	-0.74 (3.15)	-1.17 (4.82)	0.27 (1.52)	0.04 (5.70)		.00 (1.55)	
Speaking = gotten better	7.45 (7.00)	<b>6.73**</b> <b>(2.94)</b>	6.27 (3.00)	<b>5.69**</b> <b>(1.44)</b>	1.52 (3.56)		-.52 (0.97)	
<b>Teacher contact with English outside of school</b>								
Years abroad	-0.40 (0.49)	-0.32 (0.23)	-0.14 (0.21)	-0.14 (0.11)	-0.14 (0.25)	-0.12 (0.07)	-0.13 (0.06)	<b>-0.11*</b> <b>(0.05)</b>
Weekly contact/year	-0.27 (1.52)		-0.90 (0.65)	-0.77 (0.41)	0.28 (0.76)	0.55 (0.25)	0.39 (0.18)	0.31 (0.15)
SD of random intercept	6.81	3.93	2.90	1.85	3.36	2.06	0.53	0.48
AIC (ML-based)	2103.44	2092.79	1646.39	1638.82	1789.40	1782.1	1402.34	1397.28

\* p&lt;.05; \*\*p&lt;.01

**Associations with learner total scores:** In the full model, no significant associations were found to learner total scores. The largest predictor variables here were a teacher's feeling of speaking improvement and the approach to teaching. In the reduced model, there were two positive associations with the total learner scores that remained – the number of years of instruction the learners have had and the teacher's speaking improvement. Though being in the fourth grade is an important contributor to the model, it is not significant. In this population, teachers who feel as if their speaking skills had gotten better since the time of their qualifying exam have learners who performed significantly better than those learners with teachers whose speaking skills have either gotten worse or remained the same. Being a balanced skills teacher can also be assumed to be positive based on the finding here that teachers who are more traditional or who neglect writing are associated with poorer learner results.

Two negative associations with total learners' scores were found in the full model and remained in the reduced models and which were also not significant: teacher type and number of years abroad. Teachers who tend not to include much writing ("no writing") or who tend to neglect speaking ("traditional") have learners with a considerably lower mean than teachers who consider themselves "balanced skills" teachers. Another negative result was the longer the stay abroad, the lower the learners' scores.

Teacher indicators of life experience (languages spoken, years teaching and how English was learned) fall out in the reduced model as do minutes of German and weekly contact per year of English and total scores on the qualifying exam, meaning that they contribute almost nothing to the model.

**Associations with learner listening scores:** Positive associations with the learner listening scores in the full model that were strong enough to stay in the reduced models were again grade level (being in the fourth graders as compared to the third grade) and speaking change. Teachers who had the feeling that their speaking skills had improved had learners who performed significantly better than those learners of teachers whose speaking skills had stagnated or gotten worse.

Negative associations with the total learner scores in both models were again being a teacher who focuses on "no writing" or "traditional" than the reference category of being a balanced skills teacher. Furthermore, teachers who have spent

more time abroad have learners whose scores decrease as do teachers with more weekly contact in English per year – both these variables remained in the model, though were not significant.

When going from the full to the reduced model, neither indicators of teacher life experience nor a teacher's grade on the qualifying exam were strong enough to stay in the reduced model, indicating that they were not as important as the other variables.

**Associations with learner reading scores:** Positive associations with the learner reading scores which remained in the reduced models were the number of languages a teacher speaks – the more languages a teacher speaks, the higher the learner reading scores – being a balanced skills type of teacher and weeks a year abroad in an English-speaking setting. Here, speaking five as compared to three or four languages, and not being a type of teacher who neglects speaking or writing were significant.

Negative associations in both models which decreased learner reading scores were the total number of years a teacher has taught and the amount of German spoken per lesson. Furthermore, as with the other scores, teachers who indicated being “writing focused” or more “traditional” had learners with lower scores. Another negative association was the number of years a teacher has spent abroad – the longer abroad, the lower the learners' reading scores.

What falls out from the full to the reduced model are learners grade level and teacher indicators of speaking level change since the official exam, though both of these were relevant to learner total and listening scores. As in the other exams, teacher total scores on the qualifying exams falls out in the reduced model.

**Associations with learner writing scores:** Positive associations with the learner writing scores in both full and reduced models were grade level – in this test, fourth graders performed significantly better than third graders, unlike in the other subtests and for the total score. Furthermore, this time, teachers who learned English rather implicitly had learners who performed significantly better than those with teachers who mentioned learning English explicitly through language courses. Finally, teacher contact with English in weeks per year was positively associated with learner scores, though not significantly so.

Negative associations with the total learner scores in both full and reduced models were the number of years of teaching, the number of years spent abroad and

the type of teacher. As with the other models, teachers being “no writing” or “traditional” were associated with lower scores and significantly so. Furthermore, teachers who had lived abroad for a longer period of time had learners whose scores were significantly lower than those learners having teachers who had not lived abroad for an extended period of time.

As with the reading test, a teacher’s feeling of change in his or her speaking skills was not associated with learner performance in writing. There is no direct connection between a teacher’s speaking skills and a learner’s writing performance. It was attempted to replace a teacher’s speaking change with a teacher’s writing or reading change or total feeling of change, but these variables also were dropped in both the full and the reduced models, an indication that a teacher’s feeling of change was not relevant to learner performance here. Thus here a teacher’s feeling of change in terms of speaking is presented for the same reasons it was selected for all the models – this variable had the most variation.

### 4.3.3 Summary of findings from the linear mixed models

**Controlling for learner grade:** As to be expected, the learners' grade level was not associated with reading scores in the reduced model, but remained in the models generally for total scores and listening and for writing scores, where it was significant. This was also seen in the individual analyses in Section 4.2.

**No contribution – teacher scores:** In all of the models, a teacher's score on the official exam, though showing positive coefficients in the full models, was never a strong enough predictor variable to stay in the reduced models. For each model run for each of the learner scores, it dropped out of the equation. This was the only variable to interact in such a way.

**Constant contributors – teaching characteristics and teaching contact:** The type of teacher being a balanced skills teacher (as defined by NOT being a “no writing” or “traditional” teacher) stayed in every model for every test and was the variable with the most significance as seen in its constant positive association with all variables and its positive and significant association with the learner reading and writing scores. This means that clearly in this sample, being a teacher who neglects writing or speaking is negatively associated with learner scores.

Furthermore, a constant contributor, though not significant, was the number of years a teacher has lived abroad – lengthier stays were associated with poorer performance on all the tests but especially to learner writing skills. The number of weeks these teachers spend abroad in an English speaking country was both positive and negative, but remained in the models for the individual subtests. It was negatively associated with the listening scores, but positively associated (non-significantly) with learner reading and writing scores – the more the teachers in this study used English on their yearly vacations, the better their class performed on these two tests.

**Strong contributor – speaking change:** A teacher's feeling of having improved his/her speaking skills was strong enough to be positively associated with learners' total scores and was especially important for a learner's listening score, though it dropped out of the models when associated with reading and writing scores.

**Varying contributors:** The number of languages a teacher speaks also had negative and positive associations; speaking five languages was positively and significantly associated with reading scores. The number of years of teaching tended

to drop out of most models except for its association with reading scores, and was mostly negative except in the case of listening scores. The amount of German used in a lesson was similar in that the more the German, the worse the performance except in the case of the association with listening scores. This factor remained in the reduced model only for the reading scores, again insignificant.







4.4 Consolidation of findings..... 154

**Teacher classroom skills profile:** In this study, the estimated amount of time a teacher spends teaching reading, writing, speaking and listening in English interacted directly with the learners' performance in each of these same skills though the time a teacher speaks German in class did not necessarily detract from learners' performance in reading, writing and listening. When looked at independently of other variables, teacher profiles of time on skills show nothing significant, but learners' language scores do decrease when the teacher is not using "balanced" techniques. In the final mixed models, this variable was actually one of the stronger ones. It remained in all four reduced models and was consistently negative for teachers who said they neglected writing or had more traditional approaches to teaching. Thus, the hypothesis can be accepted with some hesitations due to the lack of power of this variable when looked at separately. In terms of the amount of German used in the classroom, the findings are inconclusive, as at times German is negatively associated with learner performance (on the listening and reading tests) but in the mixed models it is never significant and does not stay in all four reduced models.

## V. Discussion

Surprisingly, a teacher's speaking and "Use of English" language skills were found to be negatively associated with learner performance or completely unimportant although they were expected to have no or positive associations (in the case of reading and Use of English). It was also hypothesized that a teacher's feeling of improvement would be important and indeed it was to a certain degree. Other variables, such as regular contact with English outside of the classroom and yearly exposure during holiday time gave mixed messages. Time on combinations of skills in the classroom was expected to be an important predictor variable and it was. Therefore, there were some hypotheses met and some surprises encountered which will be discussed in this chapter.

Part 5.1 provides the results of this study in relation to the larger Freeman, Orzulak and Morissey (2009) emergent model.

Parts 5.2, 5.3 and 5.4 relook at the role of a teacher's language proficiency, methodology and state of learning, as discussed in the literature review in Chapter II in light of these findings.

## 5.1 General discussion

The Freeman, Orzulak and Morissey (2009) emergent model (described in Chapter II, Section 2.3.2) is an assessment model that can well be used to consolidate and discuss the results of the present study. The development of their emergent model is a reflection of studies and various standards set for foreign language teacher education from around the world and thus can be accepted as a reliable reflection of factors important in teaching as defined not necessarily by academics, but by stakeholders in the profession. It is thus understood that factors in such a teacher assessment model are those which are important for teaching and therefore also represent an expected teaching/teacher knowledge base. This knowledge base, as described as to what is seen as necessary for teaching a foreign language, has changed over the years from a static subject-specific language and methodology knowledge system with rigid rules for strict methods and approaches to a more dynamic system of language as mediator between general knowledge about language and methodology:

How teachers engage in the moment of interaction through the medium of language and use of their pedagogical understandings - how they play the language game in class - is connected to three inextricably linked domains we have discussed: knowing about language as content; using the language as *medium* in teaching; and knowing how to teach it, or *methodology*. (Freeman, Orzulak and Morissey, 2009, p. 86)

In combining the Freeman, Orzulak and Morissey (2009) model with the findings from this current study, some ideas for expansion to this model will be addressed with some subtle, yet important differences.

**New addition: Willingness.** The first main contribution from the findings of this research is a general background condition – a teacher’s willingness to learn and positive feelings towards learning to speak English and ability to make improvements visible (to him or herself) regardless of his or her actual level of proficiency. As shown in both the single factor models and the mixed full and reduced models of this dissertation, teachers in this study who stated that their speaking skills had improved have learners who performed better than those learners in classrooms with teachers whose speaking skills had stagnated or worsened. A teacher’s feeling of

improvement in “grammar” (as seen through the Use of English variable) had the same effect on learners’ listening scores; an indication of the idea that knowledge about language (as mentioned in Andrews, 2001, and also Freeman, Orzulak and Morissey, 2009) are conditions for effective teaching, though it is not the knowledge in and of itself per se, but rather the willingness to learn. Of course this leads to the term of the role of “life-long learning” but more specifically, teachers searching for opportunities of feeling like a learner<sup>32</sup> and is mentioned in Borg (2011):

[theories of learning to teach] need to be grounded in the study of the activities of language teaching itself and the social context in which they occur, and that a recognition of the teacher as a learner of language teaching is central to the process. (p. 219)

Imperfect use of the language by the teacher and allowing learners to be aware of this or aware of the learning process of the teacher can provide situations where learners and teachers construct meaning and mutual learning together.

**Medium: Knowing language.** At first glance, the role of “knowing language” in the original emergent model would suggest that the target language of instruction is a medium for which content is negotiated and made comprehensible, and this content is linguistic content – structures, grammar, “correct” models. Were this the definition, then the findings of this present study would counter the idea because here it seems that teacher proficiency in speaking and Use of English actually hinders learner uptake, at least in the beginning. But in looking more closely, the Freeman, Orzulak and Morissey (2009) definition is much more differentiated than it appears in their visualization of the model. They state:

In contrast to conventional grammar-based views, when language is seen as an emerging system, there is no inherent progression or sequential movement toward a target proficiency. Instead, these thinkers argue that as users, learners assemble resources in the moment to act on a particular task and achieve a particular outcome. This view of the unstable and nonstatic nature of language has clear implications for assessment of language competence, and of language as subject matter. (p. 86)

---

<sup>32</sup> This will be discussed in more detail in Section 5.2.

Thus, this “medium” does not have to be a rigid, “perfect” mastery of a linguistic system based on a native-speaker norm.

This supports the idea that language proficiency is not as important as is suggested by Llurda (2006) in what high proficiency actually means for teaching. On the other hand, this definition does not necessarily support the finding that the better a teacher’s language proficiency is, the worse the learners perform unless we look at this from the angle of overtaxing learners.<sup>33</sup> Furthermore, what is also relevant is learning language regardless of the level one starts at, as we see in this study through the association between learner scores and teachers feelings of improving their speaking skills. Thus, I would suggest that in the Freeman, Orzulak and Morissey model, the visual graphic of *Medium: Knowing language* be a bit less present, as it is inseparable from methodology (e.g. how much of the local language to use or not use) and knowing *about* language.

The negative association of teachers’ Use of English scores, speaking scores and amount of regular contact in speaking on learner scores in this study might well indicate that the medium of English – without neglecting the idea that English should actually be used in English lessons – is there for negotiating meaning and input but the teachers levels are not as important in learner performance as, for example, methodology or knowing *about* language. More directly associated to learner performance, both in the individual models and in the mixed model in this current study, were the teachers feeling of having improved their speaking skills (being in an “upward learning curve”, as mentioned above) or teaching style.

**Methodology: Knowing how to teach language.** Secondly, the findings of this study would support methodology being left as it is in the emergent model as a separate, important factor, though to what degree is hard to determine. In its interplay with other factors, as seen in the reduced mixed models in this research, it does seem to interact strongly with learner performance when other factors are involved. In both the Moradkhani, Akbari, Ghafar Samar and Kiany (2013) and Richards (1998) analyses of pedagogical content knowledge, methodology is mentioned in terms of understandings of general theories of teaching, decision-making skills, knowledge of language teaching specific theories and management skills – all of which are elements of a decision to apply a specific methodology.

---

<sup>33</sup> This will be discussed in Section 5.2 of this chapter (Way over their heads).

“Methodology” must be understood as dynamic as well, not static, and as the application of techniques and systems based on the knowledge of the learners and the situations and contexts they are in generally and in their English language use.

In this study, clearly the balanced-skills approach to teaching English to third and fourth graders was shown to be advantageous, and this is only a small indication of methodology – what exactly the settings look like would take further analysis. In scrutinizing the data here, however, one can see that the teachers who had a balanced-skills approach to teaching also had a more controlled approach to teaching. Though each of the four skills was present in each of their lessons, they were rather teacher-centered activities that represented each of the four skills (listening to the teacher as compared to listening to peers, or reading individual words as compared to reading short texts). Thus “time on task” is also supported here – teachers who included writing (balanced-skills teachers) also had learners who performed well in writing.

**Content: Knowing about language.** The role of “knowing about language” from the emergent model cannot be judged in this study though behaviors such as improving in English might point to an increased knowledge about one language. Even though a teacher’s multilingualism does seem to be positively associated in the final mixed models, this is speculative and there are too few data to justify this. One could possibly imagine, though, that a teacher who knows grammar (e.g. high scores on the Use of English part of the CAE) is not the same as a teacher who knows *about* grammar. A teacher who knows English grammar extremely well may not have as positive of an effect in the classroom as the teacher who knows about the functions of grammar (e.g. being more or less polite or direct) in his or her mother tongue and perhaps thus has a better understanding of nuances and precision in language use.

**Motivation.** What is not addressed in the Freeman, Orzulak and Morissey (2009) model but may well be an important aspect of teacher knowledge and learner performance is the aspect of motivation. Motivational strategies used by the teacher or the teacher’s motivation may not only affect learner attitudes towards the foreign language (see namely Dörnyei 1998) but also their performance. This attitude of the teacher can be seen in his or her willingness to continue learning the foreign language him or herself, in the selected methodology (changing based on the needs of the learners) and general knowledge about language in general (being interested

in the curiosities of languages and the underlying uses and structures). It was expected that variables such as teacher contact with English outside of the classroom would have played more of a role in learner performance because this contact can represent a certain curiosity about the language (or at least willingness to use it), a maintenance of level – or motivation. That these factors did not play a significant role leaves room to speculate if regular contact hinders teacher learning because this contact is taken for granted and that a teacher's intrinsic motivation to learn is no longer there; or if motivation should be measured in other ways, for example, if it had been asked if the teachers share experiences of language contact with their class instead of measuring the amount of contact itself.

Similarly, Sugita and Takeuchi (2010) state: "Since the effectiveness of motivational strategies differed according to students' proficiency level, more attention should be paid to the difference in proficiency level when teachers attempt to motivate their students" (p. 31). This indicates that different ages of learners require different combinations of strategies, as we see here in the differences between third and fourth graders' performances. Possibly, the underlying construct that drives teacher teaching and thus learner learning is that of motivation and motivational strategies used and that what are effective motivational behaviors modelled by the teacher for younger learners differs than for those of older learners.

Since pre-service training and further education and teacher evaluations all have elements of language proficiency, qualification, motivational and teaching skills, it was attempted to scrutinize these dynamic interplays in the hopes of refining a general model for the potential re-focusing of important variables of teacher knowledge and necessary qualification. To this end, this study has certainly shed some light on these issues, sometimes suggesting ruling out what is not necessary, sometimes showing what might be important and at times opening other cans of worms that can perhaps be dissected at some point in the future. Showing that certain variables have no relation in this study does also not necessarily mean that they have no relation to learner performance whatsoever, but just were not obviously measured in the study population with the tools used here. In the following sections, the discussion will turn to look at what was significant, as well as not significant in a teacher's language proficiency, methodology and feeling of learning, with each factor being looked at separately, not necessarily in relationship to a larger, more holistic model.



## 5.2 Way over their heads or way below par?

**Negative or no associations.** From the analyses and within this sample, we can see that the general language competence of the teacher has little effect on learner performance, and in fact can have an adverse effect at times. Perhaps Goldhaber and Brewer (2000), Aloe and Becker (2009) and Wayne and Youngs (2003) were perspicacious in their findings indicating that the language skills (in a monolingual setting) of the teacher were negatively or not associated with the performance of the learners in English, although some researchers tended (inconclusively) in the other direction (Andrew, Cobb and Giampietro, 2005). That the Use of English and the speaking sections of the qualifying exam had a significantly negative association on parts of learner performance may be explained. First of all, what is measured on the Use of English exam is the ability to manipulate language and the knowledge of lexical and grammatical items. Teachers who measure high in this are extremely flexible in the language, in a “correct” or “standard English” way. Therefore, in the classroom, if a teacher with a less native standard of English gives a Germanism – “order the strips of paper” as opposed to “put them in order” (so the “put” and “in” may have been tested in the Use of English) – then it is further away from the German miscue, making it easier for the learners to understand.

The speaking test measures fluency, accuracy, discourse management and range – so the same miscue mentioned above would contribute to a lower score in “accuracy”, thus making language more understandable. Furthermore, the higher the fluency and range, the more the teacher is perhaps talking “over the learners’ heads” or producing “incomprehensible input” or “noise”<sup>34</sup>. Scoring high on the speaking test is perhaps an indication of the same construct as the Use of English.

What this possibly means is that the learners may better understand a spoken model which is closer to their own level, which provides them with input in English that is comprehensible, albeit incorrect at times, according to native standards. This understanding might suggest that teachers with higher levels of English perhaps think about simplifying their speech and behaviors to make input more understandable for beginning learners. It could be argued though, that this non-

---

<sup>34</sup> In Krashen’s hallmark work (1985), he describes what optimal input is, and how learners filter out “noise”, or input that is too much for them. He also refers to native speaking teachers who may not have skills in making input comprehensible.

standard input leads to learners understanding non-native teachers, which is not necessarily the goal of English language teaching – learners should understand a range of different types of speakers, not just those with whom they share a mother-tongue.

That teacher's reading scores were considered as a possible predictor variable in this hypothesis makes sense because as of 2015, the Reading and Use of English sections of the Cambridge Advanced Exams will be combined into one part; reading covers some of the same skills which might be considered elements of teacher language awareness. Neither this variable nor listening was significant in the single regression, thus indicating that the receptive skills of the teacher do not play as much of a role as the skills where language must be manipulated and produced.

**Time abroad = high skills?** Unexpectedly, the number of years a teacher lived abroad was consistently negatively associated with learners' scores and significant in the case of writing scores. Higher levels of regular weekly contact were also negatively associated with the learners' listening and scores, though it was hypothesized that this would represent a certain motivation and enjoyment in using the language outside the classroom that perhaps transmits itself inside the classroom. How can this be explained? Perhaps if the teachers had been asked about their cultural competence or pragmatic skills (skills perhaps gained by living abroad and having regular contact) the picture would look different.

So it could be that in this case, the same teachers who lived the longest abroad were also the highest scorers,<sup>35</sup> a similar construct of language proficiency being "over the heads of the learners." Teachers in this study who had lived extensively abroad or who have regular exposure tended to be better in English. Regular contact, especially through extensive time abroad, strengthens language skills, possibly even speaking and Use of English-type skills that were measured via the qualifying exam. Teachers with such extensive exposure may have even picked up oddities and idiosyncrasies of the language which may make them more authentic but less understandable to their class. This may have advantages that metaphoric

---

<sup>35</sup> The total score for those teachers having lived the longest abroad was higher than for those having spent little time abroad, though insignificant ( $r_p = .23, p < .37$ ). The same was found for teacher total weekly contact in English and teacher total grades ( $r_p = .27, p < .001$ ). The number of weeks a teacher goes abroad per year and uses English and her total score was not significant and a very low correlation was found ( $r_p = .04, p < .55$ )

speech is easily acquired (see Bergen 2012, for example), but for this A1 – A2 level of learners, this may be simply too much.

Another insight might be Wang's (2014) findings that teachers who had extensive stays abroad tended to use the local language more often, which was negatively associated with performance in this study, albeit non-significantly.

The short stay teachers' preference for exclusive use of the target language in the classroom probably has to do with a tendency toward rather radical positions by those professionals who feel more exposed to criticism and who therefore experience higher degrees of anxiety. This tendency can also be observed in the phenomenon that NNES teachers are generally stricter than NES teachers in correcting learners' errors and assessing gravity. (p. 76)

Furthermore, none of those teachers having lived abroad for the longest period of time in this study mentioned taking language courses in their free time (others did), and though they were not always highest scorers on the qualifying exams, they were the ones who had not improved since the qualifying exam or had even lost their skills, thus this may be a more important construct than language skills themselves. This suggests that certain teacher experiences influence their behaviors in the classroom but not exactly in ways that are expected – extensive exposure to a target language may trigger teaching decisions that are perhaps not beneficial to all learners.

**Important later?** These findings, of course, do not exclude the idea that the language skills of the teacher play a role at some point in time, especially since it seems that when learners in the fourth grade are accounted for, the skills are still not significant, but, as in the case of writing skills of the learners, may start to be positively associated with learner performance. Similar findings in terms of teacher motivational strategy use were hinted at by Sugita and Takeuchi (2010) where more proficient learners reacted differently to varying teaching strategies – some strategies were over the heads of younger learners. In fact, it could be argued that a teacher's high level skills are a hindrance at first to learner performance, but are either important later in a learner's school career or are important all the time, but only clearly become evident at a later point which would be extremely difficult to measure because a learner's sources of input becomes more complex as the learner gets older and has more exposure outside of the classroom. Maybe the teachers

with higher levels have a negative impact on the learners at the beginning, but with persistence, after a few years, can help to develop learners who perform well after having had a challenging start by having had to work at quite a high level from their earliest instruction. Furthermore, it is unclear as to whether or not there is a threshold – if a teacher’s level is too high in aspects of grammar and in speaking, this could be a bad combination for explaining things in simple terms to beginners and even if this occurs in German, learners may switch off.

**Transfer.** Finally, can these findings be transferred to other subjects or to learning / teaching in general? It could be that in the specific case of English, learners get a lot of exposure outside of the classroom so therefore linguistically “weak” teacher or even peer models inside the classroom are not as important.<sup>36</sup> For learning Chinese in Switzerland, for example, the role of the teacher as a linguistic model may be much more important as learners’ exposure outside the classroom is limited. For French, perhaps this lies somewhere in between – though learners in the eastern part of Switzerland would be able to easily have exposure to French, and because the Swiss German lexicon contains many a French word, it could be the case that the linguistic model inside the classroom is not so important. That being said, creating moments of exposure or motivating learners to find these moments outside the classroom would be part of a teacher’s duty. In the case of very good linguistic models, perhaps more transfer is possible in that anything that is too much above a learner’s head needs breaking down, though not necessarily incorrectly, but into smaller, more digestible steps. And this possibly holds true for languages as well as for math and other subjects.

---

<sup>36</sup> On May 14, 2014, I taught an inservice course to 20 practicing teachers. I overheard the comment “Ja, will ich kai guets Hochtütsch cha, chönnet's Gkind au ned” [“Yeah, because I don’t know standard German well, then neither do the kids”]. So these attitudes about the language proficiency of the teacher persist, even in their native language.

### 5.3. Do what and how matter?

**Balanced skills.** One finding in this study is the role of teacher methodology and the fact that this variable remained in the full model. Clearly being a teacher with a focus on all four skills in teaching had a positive consequence on learner performance on all the three subtests (though not significantly here). Due to the fact that references to reading, writing, speaking and listening strategy-building were almost never checked in the teacher survey, these points were left out of the analyses and this is telling in its own way. This means that only those class activities which are pure time engaged in practicing the skill itself (listening to CDs, reading words, writing short sentences) were included in the grouping of teachers. Though one would expect that teachers teach learning (not necessarily English), the fact that the statements related to the teaching of strategies were not ticked was surprising.

Accordingly, clearly what happens in the classroom contributes to performance. Thus here, it may be inferred that pedagogical content knowledge is what is important in teaching. Though from the data presented in this study it is not clear what forms of skills balance is key (e.g., pair work in listening and speaking or equal treatment of listening and speaking in isolation), this study definitely shows that teachers who do not teach writing are doing a disservice to their learners. Teachers who tend to be more traditional in their teaching, neglecting speaking per se, are also doing a disservice to their learners in this context as seen through the performance test results.

**Writing.** That one year of English language instruction only makes a significant difference in writing skills (and not reading and listening skills where there was no significant difference between learners in the third and fourth grades) can be explained by the fact that there is always a range of abilities in productive skills, even in the mother tongue, and that writing is unfortunately often not focused on at all in the third grades in Switzerland. In fact, it is often completely neglected because there are no curricular standards set for writing in the official curriculum until the end of sixth grade and no report card grades expected for writing until the fourth grade. So although no teacher skill directly interacts with learners' writing skills in this study, the finding that teachers with a balanced skills approach to teaching generally had learners with better performance should encourage teachers to teach writing even in the first year of English. Forthcoming research from the University Heidelberg (see

Rymarczyk) indicates “that an early contact with the written form of English is of advantage for all young learners and that bilingual children show better results than monolingual ones do” (2013).

**Listening.** In most of the research questions in this study, it appears that the strongest associations occurred between teacher factors and learner listening skills. For example, in this population, teacher score on the Use of English and speaking parts of the qualifying exam were negatively associated with learner listening scores. Similarly, when teachers felt like their speaking skills had improved, then so did the listening performance scores of the learners. Thus, developing listening skills and knowing how to teach listening may perhaps be an important key skill that can influence learner performance overall, as Bozorgian (2012) indicates in a study on correlations between listening performance and reading, writing and speaking performance. Another consideration, as Siegel (2014) points out, would be: “Teachers may be less familiar with the intricacies of listening than they are with other language skills, both in terms of what listening entails and how to address it in the classroom” (p. 2) and therefore perhaps the how and what of teaching listening might be an important component in teaching.

**Time in German.** Time in German was not significant in any of the separate regression calculations, nor was it significant in the final mixed models. Although this is often of concern for teachers by the sheer number of times it comes up as a discussion point, the only thing that can be said here is that in the first years of English, it does not hinder performance. A concise overview of the research on the use of the mother tongue in the foreign language classroom can be found in Bruhlmann (2012). Although expecting the use of German to be negatively associated with learner performance is counter to what is written on the subject, this variable did remain in the reduced models as a negative predictor variable for reading and writing performance – the more German used, the worse the learner performance. Perhaps with longer term studies, it could well be that the use of German hinders learners’ attempts of finding comprehension strategies and is thus clearly negative on performance.

Time in German may as well indicate teacher insecurity, or even laziness, which can also negatively impact learning. However, more research into this matter might well show that selected use of the mother tongue is conducive to learning, for example using the mother tongue for contrastive activities – comparing the imperfect

uses in English and German, which ultimately fosters learning. If this is what teachers are doing was not measured. Wang (2014) found that teachers who had stayed abroad for a longer period of time were more willing to use the local language in the classroom but they did so in perhaps meaningful ways:

Teachers with longer stay abroad appear to develop a higher awareness of the current role of English as an International Language and have a deeper appreciation of the enhanced possibilities offered by the use of both L1 and L2 as resources in the L2 classroom. (p. 76)

This study does not support this finding ( $r_s = -.11$ ,  $p < .08$  for the relationship between the number of years a teacher has lived abroad and the amount of German used in the classroom), but the amount of German says nothing about the quality of its use.

**Transfer:** Again, can these findings be transferred to other subjects or to learning and teaching in general? Based on the results of this study, this is a distinct possibility. As Hattie (2003) and others point out, the quality of instruction is based on what is done in the classroom, which settings and atmospheres are created, and how much of an explicit focus is placed on a specific domain, topic, or skill. Hattie indicates that the instructional quality and direct instruction are among the top influences in teaching excellence, and these two variables can well be represented by time on skills and time in English in the English language classroom in Switzerland and poor use of German is time off topic. Therefore, the *what* and the *how* of instructional design do matter – what is being taught should balance reading, writing, speaking and listening.

#### **5.4. Learning along with one's learners**

No matter how high a teacher's language level is measured, s/he still has to have the feeling of learning and perhaps even having something to learn. Though this is perhaps something that research on teacher attitudes and personality can confirm, this point becomes evident in the finding here that teachers who have the feeling that their speaking skills have improved since the qualifying exam have learners who performed significantly better than learners of teachers who have the feeling their speaking skills have deteriorated. In teacher training, especially in compulsory in-service training, some teachers have been away from the field for years to raise families and have not developed their language skills. Others have intentionally avoided English due to the feeling that their skills were not strong enough, so why bother. Contrarily, in pre-service teacher training, some younger students are so convinced of their abilities that they do not see the value in learning with their students. They have worked hard to become experts and are not yet reflective enough to know they need to let their students know that they are learning, too, which Murphey, Asaoka, & Sekiguchi (2004) indicate as having a positive effect in the classroom.

Two of the three teachers in this study with the lowest total scores on their exams mentioned that they are registered in weekly language classes in their free time and had the feeling that their skills had improved. But not all the low-proficiency teachers stated that they had improved since the time of the exam; indeed, some of the teachers with higher levels had also seen improvements. The teachers with the highest scores did not have the highest performing learners, but the picture changed when the same teachers were grouped according to their feelings of improvement (or not). Perhaps a bit of insecurity places teachers in the category of being reflective and thus opens channels to learning and teaching.

Although Darling-Hammond, Holztzman, Gatlin and Heilig (2005) do mention that teachers should be "experts" in their fields – which might be questionable in light of some of the findings in this study – they do emphasize the role of the teacher as a learner. Whether or not the learners know that the teacher is learning with them or from them is not addressed here, but the mere fact of it seems to be enough to be associated with performance. As Huston (2009) points out in many ways, being a non-expert is also not the worst thing in the world and the advantages of non-native



speakers of English hold true here also in the sense of being a “non-expert” or co-learner.

**Transfer.** This point is probably the most transferrable to other subjects when we look at models of professional development and theories of motivation (e.g. Dörnyei 2011) which indicate that teacher factors in motivation are essential to learner success; pedagogical decisions such as providing constructive feedback on how to get there (Black & William, 2009) and the sharing of experiences and creating attitudes of “I’m in this with you” or “I’ve been there before” are motivational teacher strategies conducive to learner learning. This can well be the case for most other subjects and a reflection of the changing of effective teacher models from being omniscient (“sage on the stage”) to facilitators of learning (“guide on the side<sup>37</sup>”).

---

<sup>37</sup> Barry Popik describes the history of these terms on his blog here: [http://www.barrypopik.com/index.php/new\\_york\\_city/entry/sage\\_on\\_the\\_stage\\_or\\_guide\\_on\\_the\\_side\\_teaching\\_terms](http://www.barrypopik.com/index.php/new_york_city/entry/sage_on_the_stage_or_guide_on_the_side_teaching_terms)

## VI. Implications

In this project, very few factors from the side of the teacher were shown to interact with learner performance on a measure of reading, writing and listening performance. At best, a teacher's own language proficiency can be discounted as the main factor in learner performance and this idea should heed way for more scrutiny into teachers' choices of methodology, their knowledge about language(s), their feelings about their own learning and affective factors related to teachers sharing learning experiences and encounters with the target language outside of the classroom. The findings of this study would support teachers focusing on the means of continually improving their language skills, or even learning with their students, as well as their methodology, answering the question, "What is effective instructional design?"

The implications of this research can be applied to three levels. First of all, what might in-service and pre-service training look like based on what was found? Might there be some different models for teacher qualification? What changes might be made? Secondly, what might teacher trainers specifically look at when teaching and observing students and teachers in the field? What topics teacher trainers perhaps place more emphasis on, or assess more critically, in their courses? Finally, in the larger field of teacher selection, are there points that board members might consider in their evaluations and in their hiring of teachers for specific jobs? These points will be discussed in Sections 6.1, 6.2, and 6.3.

### 6.1. In-service and pre-service teacher training

Suggestions from this study for universities of teacher education can provide some ideas on how teachers can be educated, though some goals may be too long-term to really be implemented within a teacher-training program time frame of three to four years. Furthermore, what is relevant to the universities of teacher education Switzerland may not be relevant in other parts of the world due to the variation in populations – Swiss children are fairly exposed to English and other languages in their general environments but Dutch children may be even more exposed and US children less so. Thus what is required from teachers and policies may differ to quite an extent. In this specific Swiss-German situation, on the level of universities of teacher education, implications of the findings from this project might be viewed on the levels of changing language proficiency requirements for graduation, and reviewing in detail the requirements for different cohorts.

**Change the language proficiency requirements.** Lowering the level a teacher needs for certification from a C1 to a B2 or even changing to a more appropriate exam (as in the TKT<sup>38</sup>) would provide less stress on the individual and on the system, and provide more freedom for teaching-relevant issues because, in the author's opinion, much of a student teacher's time is spent working towards language skills mastery instead of mastering language teaching skills. More importantly, taking away or lowering the level of this stipulation may even create a situation where learners perform better because more students and teachers who do not yet meet the current C1 standard may be willing to teach English and be aware of the fact that they are still learning, which is a positive factor in the classroom. Taking away the stress of passing a C1 language exam would provide more funds for more focused instruction about language teaching, classroom language or more methodology in order to enable teachers to teach in a skills-balanced way.

The paradigm of language competence of teachers as based on native speaker norms is not relevant for their beginning learners. Other tests of language competence or adapted usages or interpretations of Cambridge tests (as described by Davies, Hamp-Lyons and Kemp, 2003 or Elder and Davies, 2006) may show that

---

<sup>38</sup> The TKT (Spratt, Pulverness, & Williams, 2011) is the Cambridge Teaching Knowledge Test. The test is written at a B2 level and though it does not test language per se, but rather teaching skills, it would provide a way of "killing two birds with one stone" – testing language and teaching competence at the same time.

language skills are important, but simply not those being tested at the moment. Furthermore, language competence tests based more on intercultural communication skills in English or strategic competence may be more reflective of a teacher's needs but still cover a range of content points and skills for a general public, thus covering the need for teachers to react to various situations and provide spontaneous input on a range of topics.

**Different conditions for different cohorts.** With the introduction of the kindergarten/lower primary cohorts at the various universities of teacher education, it may be possible to say that these future teachers only need a level of B1 to teach English to second graders (where it is taught in this grade) whereas those in the regular primary cohort should have a B2 level as indicated by the data on the varying performance between third and fourth graders in cantons where English starts in the third grade. This sort of decision, however, may cause the same issues that are currently going on in the transitions from primary to secondary (secondary teachers must have the C2 level) where teachers feel they have to “start from the beginning because of the poor models in primary.”<sup>39</sup> As this study indicates, secondary teachers may well do well to look at other variables than teacher proficiency in making this judgment and primary teachers can be reassured that their own language skills are not the key factor in learner performance, as the literature on teacher subject knowledge (as compared to pedagogical knowledge) suggests (see Chapter II, Section 2.3).

**Stay abroad, but why?** Although teachers' long term stays abroad were associated with poor learner performance, potentially due to this representing a “high-level” construct, this does not exclude speaking in favor of short term stays abroad, which are fortunately part of teacher training in Switzerland. A three-week teacher assistantship is a compulsory component for certification and all teachers in the study had the minimum of this. At the PH Zurich, for example, one of the main aims at the moment is listed as being “improving language competence” and it is only this aim that needs to be changed. Perhaps the role of language from a strategic and socio-pragmatic level should be an aim. Teacher assistantships as they are currently required in countries where English is the local language might be

---

<sup>39</sup> Comments in this direction are often overheard in secondary teacher training courses at the PSHS and PHZH.

opened up to countries where the means of communication is in English but it is not the local language, which would be more conducive to traits transferrable back into the Swiss elementary school classroom – learning linguistic and paralinguistic skills in a non-native setting and again, the emphasis might be taken away from the language improvement aspect.

**A language teacher specific profile?** Teacher trainers and university policy makers should keep an eye on the forthcoming „Berufsspezifische Sprachkompetenzprofile für Lehrpersonen für Fremdsprachen“ [Language Knowledge Profile for Foreign Language Teachers] (Egli Cuenat, Klee & Kuster, 2010) which is scheduled to be completed in 2014. Whether the skills necessary to teach, which are to be fairly detailed, are those which truly affect learner performance remains to be seen, but this may be appropriate grounds for foreign language teaching in Europe. Based on these guidelines, already in the Passepartout<sup>40</sup> project, one finds a quite differentiated overview of skills needed to teach French and English but these requirements are for learners starting with French in the third grade and English in the fourth. It remains open if what is needed on the side of the teacher is different for these age groups and these languages.

The catalogue of teaching criteria as described by Egli Cuenat (2010, p.9) provides a thorough overview of skills from the side of the teacher and does allow for the dynamic, situational use of teacher language in the fact that a teacher's level of speaking in spontaneous situations is expected to be B2 but speaking in planned situations should be C1. What is also mentioned is that that proof of proficiency in other skills (reading and writing) should also be more differentiated, not just a general C1 requirement. Whether these are the right levels for the different grades of learners and target languages (the same for French and English in the Egli Cuenat document) requires more research.

In Egli Cuenat (2010), chapter 2 provides a description of *Lernstrategische Kompetenzen und Bewusstheit für Sprachen und Kulturen* [Learning Skills for Language and Cultural Awareness] which goes a bit deeper into the role of the teacher being a good role model for learning; not being a good linguistic model per se, but being willing to admit mistakes and share learning. What is not mentioned in

---

<sup>40</sup> <http://www.passepartout-sprachen.ch/> , see „Konzept Berufsspezifische Sprachkurse C1\* für Lehrpersonen der Primarstufe (Französisch und Englisch)“

this model is the role of knowledge *about* language as could be measured in the mother-tongue, which might be an additional component to be integrated. This model merits development for specific languages (here English) and fine-tuning for specific age groups and can already be used in teacher training for awareness and analyses, if not for setting new standards.

**More discussions about skills' teaching and the use of the local language.**

In pre-service and in-service teacher training there is much more room for analyzing what settings (open or closed), what balance of skills, and how much and when to use the mother tongue to encourage learning are linked to functional (as opposed to notional) language performance of the learners. Although this current study only slightly provokes the discussion about the amount of the local language used by teachers, much more depth can be gone into readings with students about already published works in this regards and analyses of teaching observations as to what may or may not be conducive to learning. Furthermore, really looking at teacher time on what types of language tasks in terms of reading, writing, speaking and listening with their students would enable students to perhaps find more tendencies that encourage learner performance.

## 6.2 Messages to teachers?

There are, as well, implications for students and teachers which can be transmitted during in-service and pre-service courses. The main message to the individual might be:

*Definitely teach reading, writing, speaking and listening right from the start. Definitely keep learning, especially improving your speaking skills but perhaps more generally. However, don't worry about your language level or spending extensive time abroad for the sole sake of learning the target language. Do think about learning more languages, even if not perfectly and also think about the amount of and reasons for the local language (here German) used in class.*

**Keep learning.** From the perspective of teachers and their experiences, it would make sense to provide more opportunities to encourage language learning in general instead of focusing so much on passing the qualifying language competence exams. As is happening in some cantons of Switzerland, a plurilingual approach is encouraged in the elementary school, and a focus on this might be beneficial at the tertiary level. At universities of teacher education these ideas could well be integrated into the courses “Foundations of Language” or “Communication” where students do mostly learn about language and systems, about pragmatic competence, but have very little dealing with contrastive analyses or providing model lessons in exotic languages. Providing insights into systems which students may not have been aware of before can help them define what is hard to teach, what is hard for learners to learn, and gather ideas on learning and teaching strategies.

**Let your learners know you're learning.** How can teacher trainees and experienced teachers be convinced that they might do well to see themselves and even perhaps present themselves in their lessons as co-sharers of learning and not as “experts” but rather as “facilitators” who are also in the process of learning new things? For some teachers this is easy, for others this puts in question their role in the classroom. Simple strategies such as sharing where new language was encountered or sharing funny conversational situations that show communication breakdown strategies could well help model this characteristic. If being on an upward learning curve is especially important for speaking, then the role of in-service training courses not just about methodology but about various aspects of maintaining and become curious about language would be well worth the investment.

So does this mean that teachers with high levels of foreign language ability should simplify or even banalize their English, or purposely attempt to match the level of the learners, as happens in natural communication at times? No, essentially teachers should be allowed to “keep the level they have” but native-speaking or near-native speaking teachers might think to get further training in scaffolding for language learners so to make their language more accessible. And those teachers with low scores on qualifying exams should know that English as a *lingua franca* accepts nonstandard norms and these should be accepted and not “put down” as these skills are not the ones most associated with learner performance and there is quite a range of opinions on what “high” or even “enough” proficiency means.

Therefore, it might be also be wise to suggest that teachers keep taking language courses or courses *in* English not for the aim of improving language skills, but with the aim of learning, keeping on upward learning curves and remembering what learning and studying feel like. No matter how high one’s level gets, it is still good to learn with one’s learners. Taking courses and participating in professional learning communities and keeping up with skills can encourage this.

**Think about the four skills.** It is also important that teachers and student teachers become aware of the various effects on specific skills for general learner performance in reading, writing, speaking and listening and also in the context of differentiating for a specific class’ needs. For example, it appears that writing is a skill that is extremely sensitive to teacher behaviors, especially in terms of how a teacher has learned a language – teachers who learn more implicitly had, in this study, learners who performed better on writing skills. This can indicate that these teachers had a more “holistic” or even “natural” approach to their own learning which impacted the classroom behaviors in teaching writing. It could also mean that they learned spelling a bit more subconsciously and were, for whatever reason related to this, better able to present meaningful writing situations to their learners or simply were not afraid of teaching writing. If children’s listening skills are sensitive to teacher variables as it would appear in this study, then a focus on quality teaching of listening (e.g. by using models presented by Vandergrift 2007) might be worth focusing on in instruction. Time in German essentially takes time away from teaching the other skills if there is no real reason behind the decision to use it, which can have its consequences on learner performance. Finally, there is a definite lack of research on writing in the first years of English language instruction and as writing is not



covered very well in the curriculum or the curricular aims, it is often neglected, thus it merits more development and attention.

**Teacher observations:** Let's turn to how these might look in light of suggestions to be made strictly in teacher training observations. I observed a total of twenty-five teachers in January and February of 2014 and three rather typical, but each different, observations with a short analysis will be presented in Cases A, B, and C. This round of visits took place after the results of the study were obtained but had not yet fully reviewed and evaluated for their implications,<sup>41</sup> thus what might have been said before the findings is noted as well as what, in light of the findings, I might say in similar situations in the future with this new knowledge.

These teachers were either fully qualified teachers who were simply getting certified to teach English or older, part-time students in their final semester of teacher training, but who had been working in schools for a minimum of a year. In these descriptions, the focus is on the teacher though in the actual situation, general feedback included many more specific details on teaching activities geared towards the specific individuals being taught and not just on teacher skills.

---

<sup>41</sup> There are some comments and ideas from the dissertation that flowed into the feedback to the teachers, other thoughts that occurred but were not said, and some ideas that had not changed or I was not ready to change.

**Case A:** Rudolf has an extremely high level of English and he loves the language as he lived abroad for many years and has a lot of regular exposure to English through friends and music. He is highly motivated to keep on learning the language even though he has high level skills (a CAE with an 'A') and could theoretically be content. He loves to sound English and loves being precise. In his fourth grade teaching, he had his class well under control and the students were well behaved, though they are not very active. He has the learners reading and listening, but almost all of their speaking was in German and writing was minimal. He expected a translation of every single word and uses quite a lot of German himself in doing so. In one anecdote, he got extremely excited about the fact that one child had made an association between the word 'climb' and the German word "erklimmen."

**What I said:** I praised him for his skills in English and his love of the language. I was even thinking "wow, you sound like a 'native speaker'" but I bit my tongue. I complimented him on getting the learners excited about their "discovery" of the language and I mentioned that translating words shows knowledge but not usage and suggested that he present more opportunities for communicative settings (in English) in the class. I suggested he watch his "teacher talking time".

**In light of the new findings:** The suggestions provided were perhaps not incorrect, but the emphasis and the message to the teacher need to be re-evaluated. Thus, in this scenario, Rudolf's might be told that he represents a positive teacher learning curve in speaking and that he should continue his own quest in "trying to sound English" as it represents his need to be constantly learning. However, he might be a better example if he let his enthusiasm show through in getting the learners involved in more writing and speaking – controlled or not – so the skills were balanced, and letting the more cultural elements of having lived abroad flow more into the lessons than the pure linguistic joys. Like that, the learners can catch their teacher's enthusiasm in more meaningful settings (getting away from the love of language for the sake of words but towards the love of language for the fun of communicating). In the future, I might suggest that such teachers be encouraged to take up another language for a while, perhaps, to give English a break and to be in the learning situation once again, though not at the cost of discovering English.

**Case B:** Eugene is very poor in English, perhaps a B1 level, never liked the subject in school and struggles for the most basic utterances. However, when teaching his third-grade class, he makes it clear to his students how they can help themselves learn – his lessons are well structured, there are ample dictionaries and the learners are encouraged to try. Although the use of German in his classroom is quite high, German was not used for the same reason other teachers use it – Eugene did not ask for translations, he fostered discussions about how to try when you don't know a word or about the learners' feelings toward learning the language which was initiated by his own willingness to express the fact that he always had a very hard time with the language.

**What I said:** I praised Eugene for his ability to set up meaningful context for language learning because his lesson was an information-gap sort of activity where learners were 'forced' to communicate with one another. I encouraged him to use less German. Although in previous years, I may have suggested that he improve his language skills, this time it was not mentioned. I praised him for his ability to set up a classroom where learning strategies were at the forefront.

**In light of the new findings:** Eugene should be used as an example of good teaching and as a model for other teachers. Although he might be encouraged to continue learning to speak English so he feels and sees some improvement, he should explicitly be told that his "poor" language skills will most likely not hinder learner performance if he keeps on as he is, with incorporating more writing activities. Furthermore, it may not even be worth mentioning the use of German to teachers who appear to be using it meaningfully. In Eugene's case, perhaps also the role of "implicit" versus "explicit" learning should be mentioned. Although this point was only scrutinized in the mixed models, and was only relevant to learner writing, generally it was positively associated with learner performance and could well go hand-in-hand with the feelings of improvement. Should Eugene, instead of taking language courses to improve his skills (and pass the CAE), take courses in English for "fun" (such as adult-education courses on offer in Zurich for/with English speakers like baking in English) then perhaps this would be the access for him that also interacts with his class' performance.

**Case C:** Eliza is a bilingual whose mother and father are English. She grew up in Switzerland and was schooled in German. She is a confident teacher. However, when I visited, she was extremely frustrated that her class did not understand her. In the lesson I saw, the class spent a lot of time doing very simple language exercises such as matching the English word to the German word and a picture. Generally, she checked comprehension by asking for word translations. Furthermore, she explained almost everything in German.

**What I said:** With or without this study, I would not have emphasized her language skills in this case because I knew she had gotten the job because she is a native speaker of English. This teacher was hired in a community where there are many ex-pat families. Therefore, what I mentioned was the fact that there was no need for translations if there were the pictures, that the setting was not overly meaningful and that there should definitely be more of a focus on all four skills, not just reading words and listening to the teacher in English. I reassured her that with time, the class would get used to hearing her speak English.

**In light of the new findings:** In this case, I would clearly emphasize the role of what Eliza is doing in class, as I did, but I would have spent more time working with her to see how the lessons could have been changed. I would have also reviewed the curricular aims as they are not presented as “knowledge of words”. Secondly, I might suggest that she start taking lessons in another language that she does not know or ask her class to teach her (and the rest of the class) a bit of a language that someone in the class knows. This could possibly have the effect that Eliza is more aware of what it means to break language and instructions down into manageable doses quantity-wise, not quality-wise. She might thus become aware that her being an “expert” in the language is not necessarily what is conducive to her learners’ learning.

Native speakers in Switzerland have a privileged position, but may not necessarily be the best for the jobs. In my own observations, the bilingual students at the PH Zurich get opportunities to teach English as they are often contacted for substitute positions in various schools. Some of them are reflective, young people who know they do not know everything about teaching and are studying to learn. Others jump on the opportunity to profit from their linguist prowess and are perhaps not the best teachers.

Thus in teacher training and in lesson observation feedback sessions, teachers should be made aware that high levels do not necessarily bring results. Instead of focusing on what skills teachers seem to be missing, perhaps it is those with high level proficiency, having lived abroad extensively, who need to learn from those who are constantly working at it. Methodology should be emphasized and some slight adaptations to old thinking should be made, without completely revising the curriculum or observed teaching strategies. Furthermore, trainers should not praise people for their language abilities but rather search for ways of finding out how the teachers they are observing feel about where they are going in the language.

### 6.3. Selection

As Andrews (2001) points out in his research, there is also a mismatch in what teachers should know to teach and what boards of education require. Though they should have a high level of language awareness, what they are tested in are their language skills, not knowledge about the language.

**What types of teachers do schools want?** Those in charge of the hiring of teachers need to be aware that teachers selling themselves as good teachers due to their subject-matter knowledge must be taken with a grain of salt. Perhaps board members should be leery of teachers who want to show their high level skills from a subject-knowledge point of view and ask questions to inquire about a teacher's feeling of willingness to learn. They might look at a teacher's stays abroad as an important asset that favors cultural competence but not as a selling point for language teaching. Furthermore, they might do well to ask lower primary teachers about how they envision the role of the four skills in their language classroom and perhaps choose those who would not neglect writing despite many an argument to do so.

An example from Zurich is personnel evaluation (Bildungsdirektion Kanton Zürich) forms to be filled out during teacher observations. Though it would be impossible and overly pedantic to create a form for each specific subject taught, in the training that principals and board members receive, subject-specific differences in effective instructional design should be pointed out. Part B in this form is about instructional practices and states that the teacher: "Hält sich an den Lehrplan und die vorgegebenen Lehrmittel. Ist fachlich auf dem aktuellsten Stand. ... Kann die Lektion anpassen, wenn die Klasse von den eigenen Erwartungen abweicht"<sup>42</sup> (Bildungsdirektion Kanton Zürich. Volksschulamt).

Teachers and observers should know that in the eastern part of Switzerland, the curriculum (Lehrplan) is not clear in terms of the balance of skills to be taught in English. In Schaffhausen, there have been cases of teachers being criticized by board members for having learners write. Furthermore, lay board members evaluating teachers might do well to realize that though there are general elements

---

<sup>42</sup> [*Adheres to the curriculum and the compulsory textbooks. Is up-to-date with the subject-matter. Can adapt lessons when the class performs differently than expected* (Bildungsdirektion Kanton Zürich).]

of good teaching, there are also subject-specific ones. Here, for example they should know that a teacher's subject-matter skills in language may not be as important as those for math. They might do well in realizing that a subject specialist who loves the subject and has lived abroad for many years is not an effective teacher solely because of possessing those two qualities.

**Native speakers or not?** That said, it is possible that in richer areas of Switzerland, the teachers hired may have better language skills (as in examples from Thalwil) than in middle-class Swiss village (for example, Rafz) as the board of education members in charge of hiring are aware of and under pressure from the expectations of the parents. However, these teachers with "higher" level skills than the "country folk" in other areas of Switzerland may have classes of learners who perform fairly well in English, but this performance cannot be attributed to the teacher, but is rather a reflection of their exposure, worldliness, and influence of the parents. Of course maybe it is just the right combination: students with good previous skills and regular exposure to English who have high expectations can perhaps profit from and best deal with higher level teachers. Thus, should native speaking teachers be allowed to teach English? The answer is "yes", but they may well learn from the message of learning with learners and learning other languages as non-native speakers have done and are continually experiencing these processes.

As a closing anecdote, during their final methodology exam to qualify to teach English in Swiss primary schools in May 2014, in-service teachers were asked the question "If you were to apply for a position as an English language teacher (*Fachlehrperson*), what are the qualities that would make you successful?". The first comment from the ten candidates' mouths whose language skills were strong was "I am very good in English – I have lived abroad for a long time." Then came comments about methodology, love of the language, and ability to manage the classroom. For the nine candidates whose language skills were not as strong, they avoided the level of language proficiency and focused on classroom management skills and love of the language. Perhaps nothing is wrong here with these answers, but maybe we can say that what teachers expect those in charge of hiring hope to hear is about their skills in the subject first of all and then their teaching skills and general motivation secondly. I am fairly certain that if a teacher said "My language

skills are poor but I am working on it!" then they'd really have to compensate with some other skills or personality and may not be considered the strongest candidate.



## VII. Conclusions

Freeman, Orzulak and Morissey (2009) put forth the emergent view of language knowledge for/in teaching and also state:

How do we use understanding of language to inform these challenges of documenting and assessing classroom practices across multiple forms of teacher education? Some possible moves in response to this question will include: challenging forms of testing and assessment – both individually or institutionally – that rely on simplistic models of teacher knowledge; developing assessments that truly integrate multiple sources of evidence to gauge teacher preparation and effectiveness; and developing assessments that account for language as both medium and content. These issues, and others like them, will increasingly occupy our thinking as English as a global lingua franca changes our views of what language is and how it works. All of which brings us back to the person of the teacher and how she represents language as content in the act of teaching. It is the challenge of complex assessments to judge the activity of teaching through the person who does it, or in Yeats's words "...to know the dancer from the dance." (p. 87)

Mastering the art of teaching certainly can compensate for any lack of measurable skill, and it is in re-thinking existing views and perhaps even "loosening up" that will allow teachers as artists to refine their art and thus shape their learners. These final words summarize what contributions this work has made to the field of language teaching with a healthy dose of skepticism about interpreting the results.

## 7.1 Looking back

First and foremost, the interpretations from this research should by no means be taken as causal. The methods used here, regression models, help support variable reduction in order to better tweak the combination of variables that might be used for prediction. Having a hint at which variables on the side of the teacher do NOT vary with learner performance does not mean that the teacher variables which DO vary with learner performance are the reasons for learner performance; these measures may be predictive for whatever construct they represent and the interpretation of the construct represented by the measure is a reflection of the subjective eyes of the researcher. There is a lot of projection of what unobservable variable is “hidden” behind an observed one. One variable may be attributable to something else – perhaps a teacher’s score on the exam is related to general exposure and living abroad but it could also be linked to ability generally, simple laziness or poor test taking skills. Furthermore, the whole field of motivation could well have provided constructs relevant to this study. Thus, the regression models used here are for model reduction and for the predictive strength of a measure and it would be too pompous to suggest anything other than associations and co-variations between teacher attributes and learner performance which may be explained by myriad constructs, though indeed the researcher is always searching for causes.

Looking back on the level of the test for learners, it would have been important to have had information on the individual learner’s language aptitude. Knowing the learners’ grades in German and English or having other information such as from an aptitude test could help to control for a child’s ability. The main research questions could have been answered similarly by having learners’ grades in English to accompany a teacher survey, but grade information was not very differentiated as at the time the Swiss cantons approached this very differently (in some cantons teachers had to report reading, listening and speaking skills, in other only listening and speaking and yet in others simply one grade). Furthermore, as English is a relatively new subject, there may have been quite a variation in grading systems as there are no standardized measures and teachers are new at collecting data on learner English.

Furthermore, knowing more about the learners’ exposure to English may have really put it all into perspective despite teacher attributes, but this is difficult to

measure and native speakers were not included in this study. Because this study looks at performance as measured through a test, it must be said that it is a considerable feat to measure knowledge, given the dynamic nature of it and the multiple angles by which it can be defined. We can only take a snapshot of certain elements of a person's "knowledge" at any one given moment of time. Of course there are things that may correlate highly (teacher perceived subject knowledge) but there are other things which interact individually (so-called hygiene factors according to Brown, 2004).

On the level of the teachers and the teacher survey, using teacher knowledge about language (teacher language awareness) or language teaching knowledge may have been a better measure than language skills or estimations of time teaching specific skills. It would be interesting to redo this study by asking the teachers to take some form of metalinguistic awareness standardized measure or a standardized measure of teaching knowledge like the Cambridge Teaching Knowledge Test – Knowledge About Language. Similarly, a more differentiated profile of certain subskills (decoding skills) would have provided more information about the teachers' skills (e.g. phonological awareness in reading), but this information is not available. If this information had been available by having the item scores of teachers' qualifying exams, it may have been insightful to correlate it to the learner results as this information on the learners is available.

In asking teachers about their instructional practices, a survey of the frequency of certain behaviors is simpler to use and can collect data efficiently. However, teacher observations, or some control of the survey items filled out and a classroom visit, might have been a better measure but are much more cumbersome and time-consuming.

Moreover, there are certainly not enough data to generalize these results. There were approximately thirty participants per explanatory variable, clearly a small sample size. These results are also perhaps not generalizable outside central Europe and to other languages because Switzerland is not a low income area of the world, teacher training is different here, children have a lot of exposure to English and the teacher is, in many cases, no longer the primary source of input.

Methodologically, there may have been better ways of finding out which teacher variables were the most highly associated with learner performance and the issues of the multicollinearity may well have been resolved through residualizing. Though it

would have been a problem if the scores of the teacher's subtests were not collinear because if a teacher scores "very high" in reading and "very low" in listening, it would leave open the question of whether the tests were measuring the same "level" or not which is not what Cambridge would want.

Selecting the predictor variables for the mixed model was difficult as well and although it was assumed that selecting a teacher's feeling of improving his or her speaking skills from this data set would be a representation of general feelings of improvement; this may not have been the case. It may have been better to select predictor variables specifically for each learner subtest instead of selecting the same ones for all models. Furthermore,  $p < .05$  was chosen as the level of significance. Setting  $p < .1$  would have been less strict but provided some other results as many of the associations would have been significant had this level been chosen.

Finally, this dissertation was essentially driven by questioning a political decision in terms of its relevance and effectiveness on teaching quality as measured by learner performance in English. English for the sake of learning English is only one official aim of teaching English to young learners – teaching them to learn and increasing metalinguistic awareness are also aims of English language teaching which are not measured here. Although the drive was a political one and the study itself was rather explorative, models could be fit to various points and some indications of more reasoned, as opposed to intuitive decision-making have been found. Perhaps it is a chance for politics and academia to meet.

## 7.2 Looking forward

There is a definite lack of research in the field of teacher effectiveness in foreign language education, especially with beginning level learners. Though there are suggestions for measuring effective teaching as suggested by models of teacher knowledge (Richards, 1998 or Moradkhani, Akbari, Ghafar Samar and Kiany, 2013), standardized tests such as the Educational Testing Service's Praxis II World Language Pedagogy and the Cambridge TKT, and country-specific models which are a mix of political agendas and academic research, there is too much criticism around any of these measures and too little research to better define them. Thus it is difficult to find one model which projects combinations of elements of teaching styles and teacher attributes onto learner performance.

This research may contribute to refining the existing conditions in Switzerland. The one key finding, that the language skills of the teacher are not the most important part of effective teaching and could even be detrimental at times, is essential and shows that at least at the beginning of language instruction with younger learners, this should not be a deciding factor for certification and hiring. This piece of work could well contribute some insights on teacher knowledge to profiles being developed for future language teachers. Teacher trainers and those in charge of hiring might as well be able to reexamine their foci.

The test developed during the piloting phase of this study could also be used by teachers and teacher trainers as an example of what children can realistically do after a year or two of English. Teachers can also use it as a model for creating their own such tests because in the classroom there is a lot of controversy around testing young learners, with many teachers not yet comfortable with a functional, "can-do" approach to testing (many stick to translation tests or simple cloze activities).

This research is also a stepping stone for many other questions that could be followed up upon, mainly:

- Is there a threshold in learner age where the negative associations between learner performance and teacher speaking and Use of English scores actually become positive?
- What elements of teacher contact with the target language outside of the classroom can be found that actually positively influence what happens inside of the classroom?

- What is the interplay between a teacher's skills and abilities and his or her choices of teaching methods?
- Can DaZ (Deutsch als Zweitsprache) teachers teach modern foreign language teachers about what is important in teaching? What qualities might non-native German speaking DaZ teachers bring to the classroom that native speaking DaZ teachers do not?
- What is the effect of different uses of the mother tongue in the classroom on learner performance in the target language?
- What other teacher variables related to a teacher's language biography might be associated with learner performance?
- What specific differences between groups of teachers with very low level foreign language skills and those with very high level foreign language skills are associated with class performance?
- Are there ways for teachers to avoid the negative impact of their high-level language skills on student performance or even make it a positive impact?

It has been all too easy up to now to attribute quality English language teaching to teachers' language proficiency. In the future, stakeholders in the English language education of young learners need to change their focus away from subject-matter knowledge onto pedagogical knowledge. In this specific case, we need to focus less on the "what" of the amount of teacher language knowledge and rather more on the "how" of teaching and learning the language. If we want pedagogical knowledge to come to the forefront, then more rigorous qualifications related to teaching practices (as opposed to subject-matter knowledge) and much more research are necessary to clarify the dynamic interplay between teacher pedagogical knowledge and motivation to learn and its impact on learner performance.

## Appendices

### 8.1 Teacher survey and survey summary data

#### Umfrage – Lehrperson: Allgemeine Informationen

Ihre Initialen \_\_\_\_\_ Kanton: \_\_\_\_\_ Klasse: \_\_\_\_\_ Email : \_\_\_\_\_

Seit wie vielen Jahren unterrichten Sie? \_\_\_\_\_ Seit wie vielen Jahren geben Sie Englischunterricht (alle Ebenen)? \_\_\_\_\_

Mit welchen Sprachen sind Sie aufgewachsen? \_\_\_\_\_

In welchen Sprachen haben Sie jetzt Kenntnisse? \_\_\_\_\_

Würden Sie sich als zwei- oder mehrsprachig bezeichnen? **Ja/Nein**. Falls ja, in welchen Sprachen? \_\_\_\_\_

Freuen Sie sich meistens darauf, Englisch zu unterrichten? **Ja/Nein**

#### Teil 1: Ihre Fertigkeiten

1. Spezifische Sprachfertigkeiten: Diese Fragen beziehen sich auf die aktuellsten standardisierten Englisch-Sprachprüfungen, die Sie abgelegt haben. In den meisten Fällen ist dies ein Cambridge Advanced Exam. Falls Sie kein Cambridge Exam gemacht haben, fahren Sie mit Nummer 2 weiter. Sie können gerne einen Kopie Ihrer Resultate beilegen.

Geben Sie Ihre zuletzt abgelegte Cambridge-Prüfung an: **FCE / CAE / CPE** : In welchem Jahr? \_\_\_\_ Was war Ihre „score“: \_\_\_\_ Was war Ihre „grade“: \_\_\_\_

Bitte kreuzen Sie die Fertigkeiten genauso an, wie sie auf Ihrem CAE-Profil stehen.

	Reading	Writing	Use of English	Listening	Speaking
Exceptional					
Good					
Borderline					
Weak					

Wie würden Sie dieselben Fertigkeiten heute beurteilen?

	Reading	Writing	Use of English	Listening	Speaking
Exceptional					
Good					
Borderline					
Weak					

#### 2. Falls Ihre zuletzt abgelegte Prüfung keine Cambridge-Prüfung war, geben Sie bitte die folgenden Informationen (möglichst detailliert) an:

Bezeichnung der Prüfung: \_\_\_\_\_ Jahr: \_\_\_\_\_ Gesamtbewertung: \_\_\_\_\_

Einzelbewertungen: Reading: \_\_\_\_\_ Writing: \_\_\_\_\_ Speaking: \_\_\_\_\_ Listening: \_\_\_\_\_ Grammar: \_\_\_\_\_

Würden Sie heute die getesteten Fertigkeiten höher, gleich oder tiefer bewerten?

Reading: ..... **höher / gleich / tiefer**

Writing: **höher / gleich / tiefer**

Speaking: ..... **höher / gleich / tiefer**

Listening: **höher / gleich / tiefer**

Grammar: ..... **höher / gleich / tiefer**



**Teil 2: Kontakt mit Englisch ausserhalb der Schule****1. Ausserhalb der Schweiz**

A. Nennen Sie die Dauer Ihres längsten Auslandsaufenthaltes, während dem Sie Englisch gesprochen haben. \_\_\_\_\_

Wenn war dieser Aufenthalt (Jahr)? ..... \_\_\_\_\_

Wo waren Sie während dieses Aufenthalts? ..... \_\_\_\_\_

Welche Sprachen haben Sie dort verwendet (in der Reihenfolge ihrer Benutzungshäufigkeit, häufigste zuerst)? \_\_\_\_\_

B. Verbringen Sie Ihre Ferien regelmässig in einem Land oder in Ländern, wo Sie Englisch verwenden?

Ja / Nein

Wenn ja, in welchem/welchen? \_\_\_\_\_

Wie oft tun Sie das (z.B. 3 Wochen pro Jahr)? \_\_\_\_\_

**2. In der Schweiz**

Wie oft lesen/schreiben/sprechen/hören Sie Englisch ausserhalb Ihres Englisch-Schulunterrichts? Kreuzen Sie bitte die zutreffenden Felder an.

	selten (punktuell, weniger als 1 Mal pro Monat)	ca. 1-2 Mal pro Monat	jeden Tag
Lesen			
Schreiben			
Sprechen			
Hören			

Bitte beschreiben Sie Ihre aktuellen regelmässigen Kontakte mit der Sprache Englisch (Beispiele: Ihr Ehepartner ist englischsprachig, ihre Nachbarn sind es, Sie lesen regelmässig englische Literatur, nehmen Englischunterricht).

---



---



---

Wie würden Sie sagen, dass Sie Ihre Englischkenntnisse eher erworben haben:

A) Sprachkurse und Selbststudium (eher explizit)

B) Kontakt mit Englischsprechenden und Zeit im Ausland (eher implizit)?

**Teil 3: Methoden:** Diese Fragen beziehen sich auf die Zeit gegen Ende des Schuljahres, nicht auf den Jahresanfang.

1. Welches Lehrmittel verwenden Sie? \_\_\_\_\_  
Befolgen Sie die Anweisungen im Kommentar für Lehrpersonen ziemlich wörtlich? **Ja/Nein**  
Ergänzen Sie das Lehrmittel häufig mit eigenem Material aus anderen Quellen? **Ja/Nein**
2. Würden Sie sagen, dass Ihr Unterricht immer ähnlich aufgebaut ist (z.B. mit Ritualen, etwa mit einem Lied oder Spiel am Anfang)? **Ja/Nein**
3. Auf Ihre zwei oder drei Wochenstunden Englisch verteilt, schätzen Sie, wie viel Zeit Ihre Kinder mit den folgenden Tätigkeiten auf Englisch verbringen.

An explicit focus is set on getting the children to...	Ordnen Sie die Tätigkeiten in absteigender Reihenfolge gemäss der Zeit, die sie in Ihrem Unterricht beansprucht. Falls mehrere Tätigkeiten in etwa gleich viel Zeit beanspruchen, können Sie mehrmals die gleiche Ziffer setzen.	Kreuzen Sie dieses Feld an, wenn die Tätigkeit in jeder Stunde vorkommt.
...speak / sing in chorus		
...speak / listen to English in pairs/groups.		
...speak in pairs / to you individually.		
...develop speaking strategies (practice in front of a mirror, etc...).		
...present to the whole class.		
...listen to you.		
...listen to a CD or other media.		
...develop listening strategies.		
...read words.		
...read short texts / stories.		
...develop reading strategies.		
...write words.		
...write short texts (e.g. sentence completion).		
...develop writing strategies.		

4. Schätzen Sie, während wie viel Zeit Sie selber im Englischunterricht Deutsch sprechen (Minuten pro Lektion). \_\_\_\_\_  
Falls Sie im Englischunterricht Deutsch verwenden, befolgen Sie dabei bestimmte Grundsätze (kreuzen Sie die zutreffenden Grundsätze an)?

- \_\_\_\_\_ Ich spreche Englisch, wiederhole aber (für die Gesamtklasse) auf Deutsch, wenn Kinder es nicht verstehen.
- \_\_\_\_\_ Ich spreche einzelnen Kindern gegenüber Deutsch, wenn Sie etwas auf Englisch nicht verstehen, jedoch nicht der Gesamtklasse gegenüber.
- \_\_\_\_\_ Ich gebe die meisten Anweisungen auf Deutsch.
- \_\_\_\_\_ Ich gebe nur komplexere Anweisungen auf Deutsch.
- \_\_\_\_\_ Ich gebe Erklärungen zur Grammatik nur auf Deutsch.
- \_\_\_\_\_ andere: \_\_\_\_\_

**Thank you for your time! Don't hesitate to contact me if you have any questions!**

## Survey results Summary

	<i>Range</i>	<i>Mean</i>	<i>SD</i>
<b>General teacher predictors</b>			
Total number of years teaching	1 - 28	16.27	7.56
Number of language spoken	3-5	3.95	0.86
<b>Survey Section 1: Teacher levels as measured on standardized exam (low [1]-high [9])</b>			
Level on FCE / CAE / CPE (total score)	1-9	4.38	2.09
Level on reading test	1-9	4.34	2.28
Level on writing test	1-9	4.29	2.38
Level on Use of English test	1-9	4.61	2.28
Level on listening test	1-9	4.40	2.11
Level on speaking test	1-9	4.67	2.30
<b>Teacher estimated levels in 2011/2012 as compared to time of test (1 [worsened]; 2 [stagnated]; 3 [improved])</b>			
Reading level change since exam	1-3	2.26	0.58
Writing level change since exam	1-3	1.88	0.71
Use of English level change since exam	1-3	1.60	0.64
Listening level change since exam	1-3	2.20	0.55
Speaking level change since exam	1-3	2.04	0.78
<b>Survey Section 2 : Teacher contact with English outside of the classroom</b>			
Years abroad in a country where EN was language of communication	1-19	5.9	6.90
Weeks per year spent on holiday in a place where EN used	0-5	1.3	1.69
<b>Frequency of contact (1 (never), 2 (a few times a month); 3 (daily))</b>			
Reading frequency in EN outside of school time	1-3	2.62	0.60
Writing frequency in EN outside of school time	1-3	2.21	0.78
Speaking frequency in EN outside of school time	1-3	2.30	0.69
Listening frequency in EN outside of school time	1-3	2.56	0.61
Total contact	1-12	9.71	2.27
<b>Survey Section 3: Teacher characteristics in the classroom</b>			
Are rituals important – yes (1) or no (1)?	0 or 1	.8	
Balanced skills teacher - yes (1) or no (1)?	0 or 1	.25	
Traditional teacher - yes (1) or no (1)?	0 or 1	.22	
No writing teacher - yes (1) or no (1)?	0 or 1	.53	
Minutes German per lesson using German	2-30	8.24	6.25

## 8.2 Learner test

(color and originally formatted version found on <http://www.edacross.org/index.php/dissertation>)

### Kinder - Einführungsblatt

Was sind deine Initialen? \_\_\_\_\_

In welcher Klasse bist du? \_\_\_\_\_

Bist du ein Junge oder ein Mädchen? \_\_\_\_\_

Welche Sprachen sprichst du? \_\_\_\_\_

Hörst du Englisch gerne? \_\_\_\_\_

Liest du gerne auf Englisch? \_\_\_\_\_

Kreuze an, ob du die folgenden Dinge auf Englisch tun kannst:

	Ja	Nein
Wenn langsam und deutlich gesprochen wird, kann ich kurze Gespräche über Themen verstehen, die mir gut bekannt sind (z.B. Schule, Familie, Freizeit, Häuser, Autos).		
Ich kann es verstehen, wenn jemand langsam und mit einfachen Worten von sich, seinen Kolleginnen und seiner Familie erzählt.		
Ich kann einfacheren, gut aufgebauten und mit Bildern oder Folien illustrierten Vorträgen, z.B. in der Klasse, über vertraute Themen folgen; es sollte aber langsam und deutlich gesprochen werden.		
Ich kann in kurzen Artikeln in Zeitschriften für Kinder und Jugendliche die wichtigsten Punkte verstehen, wenn ich mich mit dem Thema einigermaßen gut auskenne.		
Ich kann kurze Texte in einfachen Bildergeschichten grob verstehen, wenn ich schon vieles aus den Bildern erraten kann.		
Ich kann in einfachen Porträts von Menschen (z.B. eines «Stars») viele Informationen verstehen.		
Ich kann mich selbst, meine Familie und meine Hobbys in einem Brief kurz vorstellen.		

**Hören 1****Deine Initialen:** \_\_\_\_\_**Vorbereitung :**

Lies zuerst dieses Blatt durch. Dafür hast du 2 Minuten Zeit.

**Aufgabe**

Hör dir das Gespräch zwischen Zoe und James an. Sie sprechen über Wintersport. Beantworte dazu die Fragen. Kreuze jeweils die beste Antwort an.

→Hinweis:

- Du hörst das Gespräch einmal.
- Während des Hörens gibt es einmal eine kurze Pause, damit du die Fragen 3-5 noch einmal lesen kannst.

1. Wann hat Zoe mit dem Snowboarden begonnen?

- vor einem Jahr
- im Januar
- dieses Jahr

2. Zoe hat einen Unfall gehabt. Was war die Folge?

- Bruch eines Fussknochens
- Bruch eines Handgelenks
- Bruch des Ellbogens

Pause

3. Findet Zoe das Snowboarden schwierig?

- Nein, aber am Anfang war das Gleichgewicht ein Problem.
- Nein, aber ein gutes Board ist unbedingt nötig.
- Nur das Drehen findet sie schwierig.

4. Wie gefällt James das Skifahren?

- Er ist begeistert davon.
- Er findet es noch sehr schwierig.
- Er hat noch etwas Angst davor.

5. Fährt James auch Snowboard?

- Noch nicht, aber er fährt seit langem Ski.
- Ja, aber er sucht einen neuen Lehrer.
- Nein, aber vielleicht später einmal.

## Hören 2 Deine Initialen: \_\_\_\_\_

### Vorbereitung :

Lies zuerst dieses Blatt durch. Dafür hast du 1½ Minuten Zeit.

### Aufgabe

Hör dir an, was Alison über eine spannende Woche in der Schule erzählt.

- Du findest im Folgenden eine Zusammenfassung von dem, was du hörst. Fülle die Lücken, so dass die Geschichte richtig zusammengefasst ist, auf Englisch oder Deutsch. Du hörst das Gespräch einmal.

**Teil 1** Alison geht's \_\_\_\_\_ heute.

Die Woche war über \_\_\_\_\_.

Alison hat \_\_\_\_\_ wie "doomah" gelernt, und sie hat diese auf Zettel geschrieben und in ihre \_\_\_\_\_ gesteckt.



**Teil 2** Alison beschreibt hier eine \_\_\_\_\_ über

\_\_\_\_\_.

Am Schluss sind alle \_\_\_\_\_.



**Teil 3** Alison beschreibt hier \_\_\_\_\_.

Sie hat diese Dinge aus \_\_\_\_\_ gemacht.



**Teil 4** Alison spricht hier über \_\_\_\_\_.

**Teil 5** Beschreibe eine Sache, welche die Kinder gegessen haben:

\_\_\_\_\_

Alles in allem, wie fand Alison die Woche? \_\_\_\_\_.



**Hören 3****Deine Initialen:** \_\_\_\_\_**Vorbereitung :**

Lies zuerst dieses Blatt genau durch. Dafür hast du 1½ Minuten Zeit.

**Aufgabe**

Carl kommt aus Irland und ist in der Schweiz zu Besuch. Er beschreibt einige seiner Bekannten.

Hör dir Carls Beschreibung an und überprüfe, wie gut die Notizen in der Tabelle sind.

Sind sie richtig, nur zum Teil richtig oder falsch? Kreuze das Zutreffende an.

→ Hinweise

- Achtung: Es macht nichts, wenn in den Notizen nicht alles steht, was gesagt wird.
- Zwischen den einzelnen Beschreibungen gibt es immer eine kurze Pause zum Ankreuzen und zum Weiterlesen.
- Du hörst die Beschreibungen einmal.

Name	Notizen zum Überprüfen	richtig	Vielleicht/ zum Teil richtig	falsch
<b>Bsp. John</b>	<b>klein – klug</b>		<b>x</b>	
<b>Yvonne</b>	blond – ziemlich mager			
<b>Fred</b>	kurze Haare – etwas dick			
<b>Sylvia</b>	rote Haare – blaue Augen			
<b>Emmanuel</b>	klein – treibt viel Sport			
<b>David</b>	sehr gross – lernt nicht viel			
<b>Allison</b>	ziemlich gross – witzig			

## Hören 4

Deine Initialen: \_\_\_\_\_

**Vorbereitung :**

Lies zuerst dieses Blatt genau durch. Studiere und vergleiche die Bilder. Dafür hast du 1½ Minuten Zeit.

**Aufgabe**

In dieser Aufgabe hörst du 9 Tonaufnahmen, die erste ist für dich schon gelöst. Schreib die Nummer jeder Tonaufnahme ins Kästchen links vom Bild, das in der Aufnahme beschrieben wird. Du hörst jede Aufnahme einmal.





**Lesen 1, Seite 1**      **Deine Initialen:** \_\_\_\_\_

Lies die folgende Geschichte. Beantworte die Fragen auf der nächsten Seite. Du hast 10 Minuten Zeit.

**PAT'S ST. PATRICK'S DAY SURPRISE**

Pat can't find anything green to wear for St. Patrick's Day. He's afraid children at school will be mean to him and he's right.

When he gets to school, the children pinch and chase him. Pat runs around and looks for something green to wear. He thinks he finds a scarf. He puts it around his neck. The children don't pinch him. They run away and scream. Pat doesn't know why.

Then, a teacher tells Pat there is a snake around his neck! Pat stands very still while the teacher makes a phone call. A man comes from the zoo. He takes the poisonous snake off Pat's neck. He tells Pat he is very lucky the snake did not bite him! This was the most surprising St. Patrick's Day Pat ever had in his eight years!

**Lesen 1, Seite 2****Deine Initialen:** \_\_\_\_\_

**Aufgabe 1:** Du findest im Folgenden eine Zusammenfassung von dem, was du gelesen hast. Fülle die Lücken (auf Englisch oder Deutsch), so dass die Geschichte richtig zusammengefasst ist.

Am St. Patrick's Day muss du \_\_\_\_\_ tragen. Pat geht \_\_\_\_\_ und sucht dort etwas. Pat dachte, \_\_\_\_\_ gefunden zu haben, aber tatsächlich war es \_\_\_\_\_! Die Kinder haben \_\_\_\_\_ Pat! Am Schluss der Geschichte ruft der Lehrer \_\_\_\_\_ an.

**Aufgabe 2:** Beantworte die folgenden Fragen kurz auf Deutsch:

Wie fühlt sich Pat am Schluss des Tages? \_\_\_\_\_

Wie alt ist Pat? \_\_\_\_\_

Wie oft ist St. Patrick's Day? \_\_\_\_\_

**Aufgabe 3:** Lies die folgenden Aussagen. Kreuze an, ob sie richtig oder falsch sind. Wenn du im Text keine Antwort darauf findest, kannst du „weiss nicht“ ankreuzen.

	Ja	Nein
<del>Die Bilder oben auf der Seite sind gemäss der Geschichte richtig geordnet.</del>		
Wer am St. Patrick's Day nichts Grünes an hat, wird nach Hause geschickt.		
Pat ist ein Junge.		

**Lesen 2, Seite 1**

**Deine Initialen:** \_\_\_\_\_

Lies die Ausschnitte von verschiedenen amerikanischen Zeitungen für Kinder. Löse Aufgabe 1 und 2. Du hast dafür 10 Minuten Zeit.

**Aufgabe 1:** Zeichne eine Linie von jedem Textausschnitt zum Titel, der dazu gehört. Nicht alle Titel können verwendet werden.

**A:** Did you ever dream of building a roller coaster? Cartoon buddies Phineas and Ferb turn that dream into reality! The characters take on one crazy project after another. They become rock stars. They build a time machine. Cartoonist pals Dan Povenmire and Jeff "Swampy" Marsh created *Phineas and Ferb*.

**B:** Are you the sort of student who can't get enough of school? Move to Chile! Here are the countries where kids ages 9 to 11 spend the most time in school each year.

- (1) Chile 1,094 hours per year
- (2) Italy 1,023 hours per year
- (3) The Netherlands 1,000 hours per year...

**C:** Cleopatra ruled the Egyptians over 2,000 years ago. Much is known about the powerful, beautiful queen. But a mystery remains: Where was she buried? Last week it was announced that they may have found her tomb. The team found coins, a mask and the head of a statue. "If this tomb is found, it will be one of the most important discoveries of the century," ...

**D:** Workers in the Gulf of Mexico are racing to clean up a big oily mess. The Deepwater Horizons oil rig exploded off the coast of Louisiana. An oil rig is a platform where oil drilling occurs. The blast is causing huge amounts of oil to leak into the sea. The waters are home to dolphins, sea turtles and more....

**E:** The 2010 Winter Olympic Games came to a happy end yesterday for Canada and for the USA. The Canadian men's ice hockey team beat the Americans in a 3-2 victory. When Sidney Crosby hit the puck through the legs of American Ryan Miller to score the winning goal, fans in their red-and-white jerseys—the national colors of Canada—jumped from their seats and yelled, "Go, Canada!" ...

**A Big Oil Spill**

**Machines buried in tombs**

**The Mystery of the Winter Games**

**Studying Hard?**

**Goodbye, Olympics!**

**School's out on the rig!**

**Cartoon Friends**

**Buried Beauty?**

**Lesen 2, Seite 2****Deine Initialen:** \_\_\_\_\_

**Aufgabe 2:** Beantworte die folgenden Fragen, indem du jeweils den Buchstaben des richtigen Textausschnitts hinschreibst. Es kann sein, dass mehr als eine Antwort korrekt ist.

**Beispiel:** Welcher Ausschnitt ist über Kanada? E

1. Welcher Ausschnitt ist über eine Entdeckung? \_\_\_\_\_
2. Welcher Ausschnitt ist über eine Katastrophe? \_\_\_\_\_
3. Welcher Ausschnitt ist über eine Sportart? \_\_\_\_\_
4. Welcher Ausschnitt ist über ein Volk, das vor langer Zeit lebte? \_\_\_\_\_
5. Welcher Ausschnitt ist über etwas, das in der Zukunft geschehen könnte? \_\_\_\_\_
6. Welcher Ausschnitt vergleicht Länder? \_\_\_\_\_
7. Welcher Ausschnitt erwähnt eine Freundschaft? \_\_\_\_\_
8. Welcher Ausschnitt beschreibt technische Konstruktionen? \_\_\_\_\_
9. Welcher Ausschnitt handelt von Fernsehen? \_\_\_\_\_
- ~~10. Welcher Ausschnitt stammt aus einem Buch über Fakten und Rekorde? \_\_\_\_\_~~

**Lesen 3****Deine Initialen:** \_\_\_\_\_

Lies die folgende Geschichte. Beantworte die Fragen darunter. Du hast 5 Minuten Zeit.



*Hi! I'm Michael Stevens! I live in Colorado, in the United States, but I'm actually from China. I've lived here for five years. I'm ten. I have one older sister and a younger brother who drive me crazy.*

*In my spare time, I do lots of things, but my favourite things to do are listening to music or watching movies – anime to be exact! I'm really good on the computer and at drawing and I like to write stories, but those take too much time with school and play!*

**Aufgabe 1:** Beantworte die folgenden Fragen kurz auf Englisch oder Deutsch:

1. What is the boy's first name? \_\_\_\_\_

2. How old is the boy? \_\_\_\_\_

3. Where is the boy originally from? \_\_\_\_\_

4. How many sisters and brothers does the boy have? \_\_\_\_\_

5. What is one of the boy's hobbies? \_\_\_\_\_

6. Was möchtest deiner Meinung nach der Junge einmal werden, wenn er gross ist?

\_\_\_\_\_

**Schreibaufgabe****Initialen:** \_\_\_\_\_**Anweisungen**

Stell dir vor, dass du einen neuen Brieffreund in Amerika hast. Du willst ihm jetzt in einem ersten Brief etwas über dich und dein Leben erzählen. Was würdest du schreiben?

Schreibe alles, was dir in 15 Minuten oder weniger einfällt. Schreib deinen Brief bitte ohne Hilfe - dein Englisch ist sicher gut genug dafür!

Wir haben den Brief schon mal mit "Dear Mark," angefangen. Du darfst aber den Brief anders anfangen.

Dear Mark,

### 8.3 Instructions to teachers

*This information was found online at: <http://edacross.org/study> but was removed in 2013.*

#### **Doctoral study: Association between learners' listening, reading and writing performance and teacher proficiency and experience with English**

Download : **Full abstract**

Below you will find information about the tests. The study requires the following parts:

**Teachers** : Survey (4 pages, ca. 20 minutes)

#### **Children**

Listening test – (5 pages, ca. 25 minutes)

Reading test – (6 pages, ca. 30 minutes)

Writing sample – (1 letter that can be done in class or for homework – 15 minutes)

[DOWNLOAD TESTS AND SURVEY - ONE DOCUMENT](#) (PDF, 16 A4 pages, 1 MB)

[DOWNLOAD RECORDINGS FOR THE LISTENING TEST](#) (MP3, 40 MB)

In the parts below, you have

- a description of the test difficulty level
- descriptions and more precise instructions on how to carry out each of the parts of the test.

#### **TEACHERS: Survey**

- 4 pages, 20 minutes. You will need a copy of your CAE (or other test) results or know your scores.
- You will need an estimation of how much teaching time per week you spend on reading, writing, listening and speaking. [DOWNLOAD SURVEY](#) (PDF)

#### **CHILDREN: Tests**

- Please do the tests with the children during the last four weeks of school (mid-June-mid July).
- If a child is not there for one test or part, do not worry about it.
- The answers are not included. If you would like them, I would be happy to send them to you when the school year is over. If you would like to use the test again with the same class at the beginning of the fourth grade, please feel free. This might be a good activity to see how they start the new school year and see what they remember.
- It is better not to do all the tests on one day – you can do the various tasks over the course of a week or two and take breaks between parts.
- Please keep to the time constraints!
- Some tests are difficult. Please tell the children that they shouldn't get frustrated! If everybody gets 100%, I won't learn anything! That said, for many questions, more than one answer can be correct.

**Listening test**

The children should:

- Fill in the “Einführungsblatt“ [Introductory sheet]
- Turn to the first page. Start playing the track immediately.

***THE TIME TO READ THE INSTRUCTIONS AND THROUGH THE TASK IS INCLUDED ON THE RECORDING.***

- Please do not play any recordings twice though you are welcome to use them again later, after you have sent the test to me.
- Cloze answers (Lückentexte) can be done in German OR English, it doesn't matter.

**Task descriptions**

	Time	Source	Points	Level
Einführungsblatt	5 Minutes	Laura		
Hören 1	4 Minutes	Lingualevel	5	A2.1
Hören 2	6 Minutes	Laura + daughter	12	A1.2 – A2.1
Hören 3	4 Minutes	Lingualevel	6	A1.2
Hören 4	5 Minutes	Laura	8	A1.2

[DOWNLOAD LISTENING TESTS](#) (PDF, <1MB, 5 pages - print color)

[DOWNLOAD RECORDINGS](#) (MP3, 40 MB, ca. 10 seconds between recordings).

**Reading test**

- This test is slightly longer than the listening tests, so make sure the children get a break between the tests. The first two are more difficult than the third one.
- Tests 1 and 2 are two pages each.
- Cloze answers (Lückentexte) can be done in German OR English, it doesn't matter.

**Task descriptions**

	Time	Source	Points	Level
Einführungsblatt	5 Minutes	Laura		
Lesen 1	10 Minutes	ABCteach	13	A1.2-A2.1
Lesen 2	10 Minutes	Time Magazine for Kids	15	A1.2
Lesen 3	5 Minutes	Laura	6	A1.2

[DOWNLOAD READING TESTS](#) (PDF, 6 pages)

**Writing test**

This test can be printed out, copied, and given as homework or you can have the learners do this in class.

[DOWNLOAD WRITING PROMPT](#) (PDF, 1 page)

[DOWNLOAD LETTER TO THE PARENTS](#) (PDF, 1 page)



## 8.4 Sample R Codes

Below are samples that are representative of A) the separate tests run for each question and B) the linear mixed model for a combination of variables.

*A) Separate analyses of variables – Example from Research Question 1: Do teachers with higher levels of English have learners whose competence in reading, writing and listening are also higher than those learners in classrooms with teachers having lower levels of language competence?*

```
mod_rq1_grade <- function(tvar="TGrade",
                        lvar="total_scores_learners"){
  myform3 <- as.formula(paste(lvar, "~", tvar, "+ factor(No_years_EN)"))

  mod <- lme(myform3,
            random=~1|Teacher_id, data=mydata)

  return(mod)
}
```

*B) Multilevel model from Research Question 5: Which combinations of variables on the side of the teacher, controlling for the years of English the learner has had, contribute the most to learner performance?(Same analyses run for reading, writing and listening scores, variables replaced)*

```
fullform <- as.formula(total_scores_learners ~
  T_measured_UE +
  t_sp_change +
  T_longest_stay +
  #contact_tot +
  contact_weeks +
  DE_Zeit_min +
  teacher_type +
  T_teach +
  T_lang_total +
  learning +
  No_years_EN)
# ML estimation for the AIC
mod <- lme(fullform, random=~1|Teacher_id, data=mydata, method="ML")
AIC(mod)
# REML estimation for the output
mod <- lme(fullform, random=~1|Teacher_id, data=mydata, method="REML")
summary(mod)
zeroform <- as.formula(total_scores_learners ~ 1)

mod <- lme(fullform, random=~1|Teacher_id,
          data=mydata, method="ML")

# Stepwise backward selection based on the AIC
modstep <- stepAIC(mod, scope = zeroform,
                  direction="backward")
summary(modstep)
AIC(modstep)

# Model with REML after backward selection (Variables changed based on stepwise findings)
form_final <- as.formula(total_scores_learners ~ t_sp_change +
  T_longest_stay + teacher_type + No_years_EN)
mod <- lme(form_final, random=~1|Teacher_id,
          data=mydata, method="REML")
summary(mod)
```

**8.5 Associations – Teacher feeling of change (total) and learner subtest scores***Learner listening scores grouped by teacher feeling of improvement on all skills*

	Teacher change	Mean	Med	Sd	Min	Max
3rd graders	gotten worse	14.28	14.50	3.83	8.00	21.00
	stayed the same	15.55	14.00	5.72	5.00	29.00
	Total	15.35	14.00	5.48	5.00	29.00
4th graders	gotten worse	13.55	14.00	4.77	5.00	25.00
	stayed the same	15.21	14.00	4.82	6.00	28.00
	gotten better	18.41	19.00	4.55	10.00	26.00
	Total	15.68	15.00	4.97	5.00	28.00
3rd and 4th graders	gotten worse	13.89	14.00	4.31	5.00	25.00
	stayed the same	15.37	14.00	5.26	5.00	29.00
	gotten better	18.41	19.00	4.55	10.00	26.00
	Total	15.54	15.00	5.18	5.00	29.00

*Learner reading scores grouped by teacher feeling of improvement on all skills*

	Teacher change	Mean	Med	Sd	Min	Max
3rd graders	gotten worse	15.83	16.50	6.56	7.00	30.00
	stayed the same	15.47	16.00	5.72	4.00	28.00
	Total	15.53	16.00	5.83	4.00	30.00
4th graders	gotten worse	16.25	17.00	3.85	6.00	22.00
	stayed the same	15.99	16.00	6.31	3.00	31.00
	gotten better	18.65	20.00	6.01	5.00	28.00
	Total	16.58	16.00	6.07	3.00	31.00
3rd and 4th graders	gotten worse	16.05	17.00	5.24	6.00	30.00
	stayed the same	15.74	16.00	6.03	3.00	31.00
	gotten better	18.65	20.00	6.01	5.00	28.00
	Total	16.14	16.00	5.98	3.00	31.00

*Learner writing scores grouped by teacher feeling of improvement on all skills*

	Teacher change	Mean	Med	Sd	Min	Max
3rd graders	gotten worse	10.33	10.75	3.03	5.00	17.00
	stayed the same	11.56	11.50	3.05	4.00	18.74
	Total	11.37	11.50	3.06	4.00	18.74
4th graders	gotten worse	13.65	14.00	1.79	10.00	16.50
	stayed the same	13.15	13.46	3.20	5.50	18.50
	gotten better	13.23	13.00	2.81	7.19	17.50
	Total	13.23	13.46	2.97	5.50	18.50
3rd and 4th graders	gotten worse	12.08	12.50	2.94	5.00	17.00
	stayed the same	12.39	12.00	3.22	4.00	18.74
	gotten better	13.23	13.00	2.81	7.19	17.50
	Total	12.45	12.50	3.14	4.00	18.74

## References

- Achermann, B., & Sprague, K. (2006). *Explorers*. Zürich: Lehrmittelverlag des Kantons Zürich.
- Aloe, A. M., & Becker, B.J. (2009). Teacher verbal ability and school outcomes: Where is the evidence? *Educational Researcher*, 38(8), 612–624. doi: [10.3102/0013189X09353939](https://doi.org/10.3102/0013189X09353939)
- American Council on the Teaching of Foreign Languages. (2010). *Position statement on the use of the target language in the classroom*. Retrieved from <http://www.actfl.org/i4a/pages/index.cfm?pageid=4368#targetlang>
- American Council on the Teaching of Foreign Languages. (2012). *Proficiency guidelines*. Retrieved from [http://actfl.org/files/public/ACTFLProficiencyGuidelines2012\\_FINAL.pdf](http://actfl.org/files/public/ACTFLProficiencyGuidelines2012_FINAL.pdf)
- Andrew, M.D., Cobb, C.D., & Giampietro, P.J. (2005). Verbal ability and teacher effectiveness. *Journal of Teacher Education*, 56(4), 343–354. doi:[10.1177/0022487105279928](https://doi.org/10.1177/0022487105279928)
- Andrews, S. (2001). The language awareness of the L2 teacher: Its impact upon pedagogical practice. *Language Awareness*, 10(2-3), 75–90. doi:[10.1080/09658410108667027](https://doi.org/10.1080/09658410108667027)
- Andrews, S. (2003). Teacher language awareness and the professional knowledge base of the L2 teacher. *Language Awareness*, 12(2), 81–95. doi:[10.1080/09658410308667068](https://doi.org/10.1080/09658410308667068)
- Andrews, S. (2006). The evolution of teachers' language awareness. *Language Awareness*, 15(1), 1-19. doi:[10.1080/09658410608668846](https://doi.org/10.1080/09658410608668846)
- Anton, M., & DiCamilla, F. (1998). Socio-cognitive functions of L1 collaborative interaction in the L2 classroom. *Canadian Modern Language Review*, 54 (3), 314 - 353.
- Arnet-Clark, I., Lanz, C., Koller-Grimm, U., Hartmann, F., Sigel, B., Schüpfer, I., & Stampfi-Vienny, C. (2005). *Young World 1*. Klett Verlag. Switzerland.
- Arva, V., & Medgyes, P. (2000). Native and non-native teachers in the classroom. *System*, 28, 355-372.
- Bachman, L. (2004). *Statistical Analyses for Language Assessment*. Cambridge: Cambridge University Press.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59, 389-407. doi:[10.1177/0022487108324554](https://doi.org/10.1177/0022487108324554)
- Barton, A., Bragg, J., & Serratrice, L. (2009). 'Discovering language' in primary school: an evaluation of a language awareness programme. *Language Learning Journal*, 37(2). doi:[10.1080/09571730902928029](https://doi.org/10.1080/09571730902928029)
- Bell, T. (2005). Behaviours and attitudes of effective foreign language teachers: Results of a questionnaire study. 38(2), 259-270. doi:[10.1111/j.1944-9720.2005.tb02490.x](https://doi.org/10.1111/j.1944-9720.2005.tb02490.x)
- Bergen, B. (2012). *Louder than words*. Basic Books. Perseus. New York.
- Berry, R. (1990). The role of language improvement in in-service teacher training programmes: Killing two birds with one stone. *System*, 18(1), 97-105.
- Berthele, R. (2009). Überlegungen zur quasi totalen aber vollkommen normalen Nutzlosigkeit sprachwissenschaftlicher Forschung für die Unterrichtspraxis. *Zeitschrift für Literaturwissenschaft und Linguistik*, (39, 153), 10-25.

- Bildungsdirektion Kanton Zürich. Volksschulamt. Mitarbeiterbeurteilung. [Internal publication]  
Retrieved from <http://vsa.zh.ch>
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. doi:[10.1007/s11092-008-9068-5](https://doi.org/10.1007/s11092-008-9068-5)
- Blomeke, S., Suhl, U., & Kaiser, G. (2011). Teacher education effectiveness: quality and equity of future primary teachers' mathematics and mathematics pedagogical content Knowledge. *Journal of Teacher Education*, 62(2), 154–171. doi:[10.1177/0022487110386798](https://doi.org/10.1177/0022487110386798)
- Boonen, T., Van Damme, J., & Onghena, P. (2013). Teacher effects on student achievement in first grade: Which aspects matter most? *School Effectiveness and School Improvement*, 25(1), 126–152. doi:[10.1080/09243453.2013.778297](https://doi.org/10.1080/09243453.2013.778297)
- Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research*, 10(1), 3–31. doi:[10.1191/1362168806lr182oa](https://doi.org/10.1191/1362168806lr182oa)
- Borg, S. (2011). Language teacher education. In J. Simpson (Ed.), *The Routledge handbook of applied linguistics* (pp. 215-228). London: Routledge
- Bozorgian, H. (2012). The relationship between listening and other language skills in international English Language Testing System. *Theory and Practice in Language Studies*, 2(4), 657-663. doi:[10.4304/tpis.2.4.657-663](https://doi.org/10.4304/tpis.2.4.657-663)
- Brady, B. (2007). [Review of the book: *Non-native Language Teachers: Perceptions, Challenges and Contributions to the Profession* by Lurda, E. ed.]. *International Journal of Multilingualism* 4:3, 236-238, doi:[10.2167/ijm014b.0](https://doi.org/10.2167/ijm014b.0)
- Braine, G. (1999). *Non-native educators in English language teaching*. Mahwah, NJ: L. Erlbaum Associates.
- Breiman L. (2001) Statistical modeling: the two cultures. *Statistical Science* 2001;16:199 – 231.
- Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed.; pp. 328–375). New York: Macmillan.
- Brosh, H. (1996). Perceived characteristics of the effective language teacher. *Foreign Language Annals*, 29(2), 125–136. doi:[10.1111/j.1944-9720.1996.tb02322.x](https://doi.org/10.1111/j.1944-9720.1996.tb02322.x)
- Brown, A. V. (2009). Students' and teachers' perceptions of effective foreign language teaching: A comparison of ideals. *The Modern Language Journal*, 93(1), 46–60. doi:[10.1111/j.1540-4781.2009.00827.x](https://doi.org/10.1111/j.1540-4781.2009.00827.x)
- Brown, N. (2004). What makes a good educator? The relevance of meta programmes. *Assessment & Evaluation in Higher Education*, 29(5), 515–533. doi:[10.1080/0260293042000197618](https://doi.org/10.1080/0260293042000197618)
- Bruce, C. D., Esmonde, I., Ross, J., Dookie, L., & Beatty, R. (2010). The effects of sustained classroom-embedded teacher professional learning on teacher efficacy and related student achievement. *Teaching and Teacher Education*, 26(8), 1598–1608. doi:[10.1016/j.tate.2010.06.011](https://doi.org/10.1016/j.tate.2010.06.011)
- Bruhlmann, A. (2012). Does the L1 have a role in the foreign language classroom? A review of the literature. *Working papers in TESOL and Applied Linguistics*. 12(2), 55-80. Retrieved from [www.tc.columbia.edu/tesolwebjournal](http://www.tc.columbia.edu/tesolwebjournal)
- Bruner, J. (1996). *The culture of education*. Cambridge: MA. Harvard University Press.

- Bruton, A. (2010). Are the differences between CLIL and Non-CLIL groups in Andalusia due to CLIL? A reply to Lorenzo, Casal and Moore (2010). *Applied Linguistics*, 32(2), 236–41. doi:[10.1093/applin/amr007](https://doi.org/10.1093/applin/amr007)
- Buddin, R., & Zamarro, G. (2009). Teacher qualifications and student achievement in urban elementary schools. *Journal of Urban Economics*, 66, 103–115. doi:[10.1016/j.jue.2009.05.001](https://doi.org/10.1016/j.jue.2009.05.001)
- Buechel, L. & Sprague, K. (2013). *Teaching English final exam grading rubric at the PHZH*. [Internal document]. Zurich University of Teacher Education.
- Buechel, L. (2012). Non-Native English speaker accents in Swiss elementary Schools: A summary of pre-service teacher research. In Y. Bayyurt & S: Akcan (Eds.). *Proceedings of the Fifth International Conference of English as a Lingua Franca*.
- Büeler, X., Stebler, R., Stockli, G. & Stotz, D. (2001). *Schulprojekt 21. Lernen für das 21. Jahrhundert? Wissenschaftliche Evaluation*. Schlussbericht zuhanden der Bildungsdirektion des Kantons Zurich. Zurich/Winterthur: Universität Zurich und Zürcher Hochschule Winterthur.
- Burham, K.P. & Anderson, D.R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*, 2nd Ed. Springer, New York.
- Butler, Y. G. (2004). What level of English proficiency do elementary school teachers need to attain to teach EFL? Case studies from Korea, Taiwan, and Japan. *Tesol Quarterly*, 38(2), 245-278. doi:[10.2307/3588380](https://doi.org/10.2307/3588380)
- Butzkamm, W. & Caldwell, J. (2009). *The Bilingual Reform: A paradigm shift in foreign language teaching*. Tübingen: Gunter Narr Verlag.
- Butzkamm, W. (2003). We only learn language once. The role of the mother tongues in FL classrooms: death of a dogma. *Language Learning Journal*, 28(1): 29–39. doi:[10.1080/09571730385200181](https://doi.org/10.1080/09571730385200181)
- Cambridge English Language Assessment. (2014). *Exam results converter*. Retrieved from: <http://www.cambridgeenglish.org/exams-and-qualifications/advanced/results/exam-results-converter/>
- Cambridge English Language Assessment. (2014). What's in Cambridge English: Advanced? Retrieved from: <http://www.cambridgeenglish.org/exams-and-qualifications/advanced/whats-in-the-exam/>
- Cambridge ESOL (2013). Principles of good practice: Quality management and validation in language assessment. Retrieved from <http://www.cambridgeenglish.org/images/22695-principles-of-good-practice.pdf>
- Cambridge Young Learners English tests. Handbook for teachers. (2007). <http://www.cambridgeenglish.org/images/84670-vocabulary-list.pdf>
- Carnoy, M., & Arends, F. (2012). Explaining mathematics achievement gains in Botswana and South Africa. *PROSPECTS*, 42(4), 453–468. doi:[10.1007/s11125-012-9246-6](https://doi.org/10.1007/s11125-012-9246-6)
- Carroll, J. B. (1975). *The teaching of French as a foreign language in eight countries*. NY: John Wiley and Sons.
- Celik, S. (2006). A concise examination of the artificial battle between native and non-native speaker teachers of English in Turkey. *Kostamonu Education Journal*, 14 (2), 371–376.
- Cetin, Y., & Flamand, L. (2013). Posters, self-directed learning, and L2 vocabulary acquisition. *ELT Journal*, 67(1)52–61. doi:[10.1093/elt/ccs053](https://doi.org/10.1093/elt/ccs053)

- Chambless, K. (2012). Teachers' oral Proficiency in the target language: Research on its role in language teaching and learning. *Foreign Language Annals*, (45, 1), 141–162. doi:[10.1111/j.1944-9720.2012.01183.x](https://doi.org/10.1111/j.1944-9720.2012.01183.x)
- Clotfelter, C., Ladd, H., & Vigdor, J. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, Elsevier, 26(6), 673-682. doi:10.1016/j.econedurev.2007.10.002
- Coniam, D., & Falvey, P. (1999b). The English language benchmarking initiative: A validation study of the Classroom Language Assessment component. *Asia Pacific Journal of Language in Education*, 2(2), 1–35.
- Coniam, D., & Falvey, P. (2013). Ten years on: The Hong Kong Language Proficiency Assessment for Teachers of English (LPATE). *Language Testing*, 30(1), 147–155. doi:[10.1177/0265532212459485](https://doi.org/10.1177/0265532212459485).
- Council of Europe Council for Cultural Co-operation. (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Darling-Hammond, L., & John Bransford, J. (2007). *Preparing Teachers for a Changing World: What Teachers Should Learn and Be Able to Do*. San Francisco, CA: Jossey-Bass
- Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Heilig, J. V. (2005). Does teacher certification matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. Chapel Hill, NC: The Southeast Center for Teaching Quality. Retrieved from <http://epaa.asu.edu/ojs/article/view/147/273>
- Darling-Hammond, L., Berry, B., & Thoreson, A. (2001). Does teacher certification matter? Evaluating the evidence. *Educational Evaluation and Policy Analysis*, 23:(1), 57–77. doi: [10.3102/01623737023001057](https://doi.org/10.3102/01623737023001057)
- Darling-Hammond, L. (1999). Teacher quality and student achievement: A review of state policy evidence. Seattle: Center for the Study of Teaching Policy.
- Davies, A. (2003). *The native speaker: Myth and reality*. London: Cromwell Press Ltd.
- Davies, A., Hamp-Lyons, L. & Kemp, C. (2003). Whose norms? International proficiency tests in English. *World Englishes*, 22(4), 571-584. doi:[10.1111/j.1467-971X.2003.00324.x](https://doi.org/10.1111/j.1467-971X.2003.00324.x)
- de Lima, D. C. (2001). English-as-a-foreign-language teacher-training programs: An overview. *Linguagem & Ensino*, 4(2), 143-153.
- Diehr, B., & Rymarczyk, J. (2012). Current developments in teaching English as a foreign language in German primary schools. *Anglistik*, 23(1), 13-23.
- Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, 31(3), 117-135. doi : <http://dx.doi.org/10.1017/S026144480001315X>
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd edition). Harlow: Longman.
- Edelenbos, P., Johnstone, R., & Kubanek, A. (2006). *The main pedagogical principles underlying the teaching of languages to very young learners*. Final report of the EAC 89/04, Lot 1 study European Commission Education and Culture, Retrieved from [http://ec.europa.eu/languages/policy/language-policy/documents/young\\_en.pdf](http://ec.europa.eu/languages/policy/language-policy/documents/young_en.pdf)

- Education, Audiovisual and Culture Executive Agency. (2008, 2012). Key data on teaching languages at school in Europe. Retrieved from [http://eacea.ec.europa.eu/education/eurydice/key\\_data\\_en.php](http://eacea.ec.europa.eu/education/eurydice/key_data_en.php). doi:10.2797/77414
- Educational Testing Service (ETS). (2014). World languages pedagogy. Retrieved from <https://www.ets.org/praxis/prepare/materials/584>
- Egli Cuenat, M. (2010). Konzept berufsspezifische Sprachkurse C1 für Lehrpersonen der Primarstufe (Französisch und Englisch). Retrieved from <http://www.passepartout-sprachen.ch>
- Egli Cuenat, M., Klee, P., & Kuster, W. (2010). Berufsspezifische fremdsprachenkompetenzen in der lehrpersonenbildung. *Babylonia*, 1, 41-48.
- Ehrenberg, R. G., & Brewer, D. J. (1995). Did teachers' race and verbal ability matter in the 1960's? Coleman revisited. *Economics of Education Review*, 14(1), 1-21.
- Ehrenberg, R. G., & Brewer, D. J. (1994). Do school and teacher characteristics matter? Evidence from high school and beyond. *Economics of Education Review*, 13(1), 1-17.
- Elder, C., & Davies, A. (2006). Assessing English as a Lingua Franca. *Annual Review of Applied Linguistics*, 26, 282-301.
- Elder, C. (2001). Assessing the language proficiency of teachers: are there any border controls? *Language Testing*. 18(2): 149-170. doi:[10.1177/026553220101800203](https://doi.org/10.1177/026553220101800203)
- Enever, J. (2014). Primary English teacher education in Europe. *ELT Journal* (68,3) 231-42. doi:[10.1093/elt/cct079](https://doi.org/10.1093/elt/cct079).
- Erziehungsdirektoren-Konferenz Ostschweiz. (2010). *Lehrplan Englisch: Primarstufe und sekundarstufe 1*. Available from: <http://www.vsa.zh.ch>
- European Council. (2002, March). Barcelona presidency conclusions. Nr: 100/1/02
- Evaluation & Accreditation of Quality in Language Services. (2008). EAQUALS bank of descriptors. *EAQUALS/ALTE Portfolio Descriptor Revision Project*. Retrieved from [http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Key\\_reference/EAQUALSBank\\_levels\\_EN.pdf](http://www.coe.int/t/dg4/education/elp/elp-reg/Source/Key_reference/EAQUALSBank_levels_EN.pdf)
- FAQ Zu Passepartout. (2014 Febuary). Publication of the Erziehungsdirektion des Kantons Bern: Amt für Kindergarten, Volksschule und Beratung. Retrieved from <http://www.erz.be.ch>
- Ferguson, R., & Ladd, H. F. (1996). How and why money matters: An analysis of Alabama schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 265-298). Washington, DC: Brookings.
- Freeman, D., Orzulak, M.M., & Morissey, G. (2009). Assessment in second language teacher education. In A. Burns & J.C. Richards (eds.), *The Cambridge guide to second language teacher education*. (pp. 77-90). Cambridge: Cambridge University Press.
- Gelman, A. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge ; New York: Cambridge University Press.
- Goldhaber, D., & Brewer, D.J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis*, 22, 129-145. doi:[10.3102/01623737022002129](https://doi.org/10.3102/01623737022002129)
- Goldhaber, D. & Brewer D. (1996). Evaluating the effect of teacher degree level on educational performance. In Fowler, W. (ed). *Developments in School Finance*. (pp. 197-210). Washington, D.C.: National Center for Education Statistics.

- Goldhaber, D., & Brewer, D. (1997). Why don't schools and teachers seem to matter? Assessing the impact of unobservables on educational productivity. *Journal of Human Resources*, 32(3), 505-523. doi:10.2307/146181
- Grant, L. (1997). Testing the language proficiency of bilingual teachers: Arizona's Spanish proficiency test. *Language Testing*, 14(1): 23-46.
- Grossman, P.L. & Shulman, L.S. (1994). Knowing, believing, and the teaching of English. In Shanahan, T. (ed.), *Teachers thinking, teachers knowing: reflections as literacy and language education*. Council of Teachers of English, 3–22.
- Hammadou-Sullivan, J.A. (2004). Identifying the best foreign language teachers: Teacher standards and professional portfolios. *Modern Language Journal*, 88(3), 390-402. doi: [10.1111/j.0026-7902.2004.00236.x](https://doi.org/10.1111/j.0026-7902.2004.00236.x)
- Hanushek, E.A. (1986) Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature*, 24 (3) 1141-1177.
- Hanushek, E.A. (1997) Assessing the Effects of School Resources on Student Performance: An update. *Educational Evaluation and Policy Analysis* 19 (2) 141-164. doi: [01623737019002141](https://doi.org/10.1623737019002141)
- Harris, J. & Ó Duibhir, P. (2011). *Effective language teaching: A synthesis of the research*. Research Report No. 13. Trinity College Dublin, Ireland.
- Hasler, D. (2006). *Fruhenglisch in Lichenstein*. [Diploma assignment at the Schaffhausen University of Teacher Education].
- Hattie, J. (2003). Teachers make a difference: What is the research evidence? Paper presented at the Australian Council for Educational Research Annual Conference on Building Teacher Quality, Melbourne, Australia.
- Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.
- He, E. (2012). Systematic use of mother tongue as learning/teaching resources in target language instruction. *Multilingual Education*, 2(1). doi:[10.1186/2191-5059-2-1](https://doi.org/10.1186/2191-5059-2-1)
- Hicks, D., & Littlejohn, A. (2003). *Primary Colours*. Cambridge University Press.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371–406. doi:[10.3102/00028312042002371](https://doi.org/10.3102/00028312042002371)
- Hlas, A., & Hlas, C. A review of high-leverage teaching practices: Making connections between mathematics and foreign languages. *Foreign Language Annals*, (45, 1), 76–97. doi:[10.1111/j.1944-9720.2012.01180.x](https://doi.org/10.1111/j.1944-9720.2012.01180.x)
- Hox, J. (2010). *Multilevel analysis: Techniques and applications* (2<sup>nd</sup> ed.). New York, USA: Routledge.
- Huang, F. L., & Moon, T. R. (2009). Is experience the best teacher? A multilevel analysis of teacher characteristics and student achievement in low performing schools. *Educational Assessment, Evaluation and Accountability*, 21(3), 209-234. doi:[10.1007/s11092-009-9074-2](https://doi.org/10.1007/s11092-009-9074-2)
- Hulstijn, J. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91 (4), 663–67. doi:[10.1111/j.1540-4781.2007.00627\\_5.x](https://doi.org/10.1111/j.1540-4781.2007.00627_5.x)



- Huston, T. (2009). *Teaching what you don't know*. Cambridge, USA: Harvard University Press.
- In'nami, Y., & Koizumi, R. (2009). A meta-analysis of test format effects on reading and listening test performance: Focus on multiple-choice and open-ended formats. *Language Testing*, 26(2), 219–244. doi:[10.1177/0265532208101006](https://doi.org/10.1177/0265532208101006)
- Jacob, B. A., & Lefgren, L. (2004b). The impact of teacher training on student achievement: quasi-experimental evidence from school reform efforts in Chicago. *Journal of Human Resources* 39(1): 50–79. doi: [10.3386/w8916](https://doi.org/10.3386/w8916)
- Jarvis, S. (2013). How effective is it to teach a foreign language in the foundation stage through songs and rhymes? *Education 3-13: International Journal of Primary, Elementary and Early Years Education*. 41(1), 47–54. doi:[10.1080/03004279.2012.710099](https://doi.org/10.1080/03004279.2012.710099)
- Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an International Language. *Applied Linguistics* 23, 83–103. doi: [10.1093/applin/23.1.83](https://doi.org/10.1093/applin/23.1.83)
- Kase, K., & Jensen, JC. (2013). Can a non-native speaker train students in English phoneme acquisition? A classroom study. *Kinki university center for liberal arts and foreign language education journal. Foreign language edition*, 3(1), 239-251. Retrieved from <http://kurepo.clib.kindai.ac.jp>
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2012). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, 64(1), 90-106. doi:[10.1177/0022487112460398](https://doi.org/10.1177/0022487112460398)
- Kourieos, S. (2014). The knowledge base of primary EFL teachers – Pre-service and in-service teachers' perceptions. *Journal of Language Teaching and Research*, 5(2). doi:[10.4304/jltr.5.2.291-300](https://doi.org/10.4304/jltr.5.2.291-300)
- Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.
- Kreis, A., Williner, M., & Maeder, C. (2014). Englischunterricht in der Primarschule des Kantons Thurgau. Schlussbericht der Evaluation. *Forschungsbericht Nr. 12*. Pädagogische Hochschule Thurgau, Switzerland.
- Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Educational Evaluation and Policy Analysis*, 25(3), 287–298. doi:[10.3102/01623737025003287](https://doi.org/10.3102/01623737025003287)
- Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. New York, NY: Routledge.
- Lazaraton, A., & Ishihara, N. (2005). Understanding second language teacher practice using microanalysis and self-reflection: A collaborative case study. *Modern Language Journal*, 89(4), 529-542. doi:[10.1111/j.1540-4781.2005.00328.x](https://doi.org/10.1111/j.1540-4781.2005.00328.x)
- Lenz, P. (2007). *Lingualevel: Instrumente zur Evaluation von Fremdsprachenkompetenzen : 5.-9. Schuljahr*. Bern: Schulverlag. Available at: <http://lingualevel.ch>
- Lipovsky, C., & Mahboob, A. (2007). Examining attitudes towards NESTs and NNESTs: A comparison of a thematic vs. an appraisal analysis. In Christina Gitsaki (Ed.) *Language and Languages: Global and Local Tensions*. Newcastle upon Tyne: Cambridge Scholars Press.

- Llurda, E., & Huguet, A. (2003). Self-awareness in NNS EFL primary and secondary school teachers. *Language Awareness*, 12(3): 220-233. doi:[10.1080/09658410308667078](https://doi.org/10.1080/09658410308667078)
- Llurda, E. (2004). Non-native teachers and English as an international language. *International Journal of Applied Linguistics*, 14(3), 314-323. doi:[10.1111/j.1473-4192.2004.00068.x](https://doi.org/10.1111/j.1473-4192.2004.00068.x)
- Llurda, E. (2005). Non-native TESOL students as seen by practicum supervisors. In Llurda, E. (Ed.), *Non-native language teachers: Perceptions, challenges, and contributions to the profession* (pp. 131-154). Boston, MA: Springer.
- Llurda, E. (ed). (2006). *Non-Native language teachers: Perceptions, challenges, and contributions to the profession*. New York: Springer.
- Llurda, E. (2008). The effects of studying abroad on self-perceptions of Catalan EFL teachers. In Dogancay-Aktuna, S. & J. Hardman (eds.), *Global English Teaching and Teacher Education: Praxis and Possibility*, (pp. 99-111). Alexandria: Teachers of English to Speakers of Other Languages, Inc.
- Llurda, E. (2009). Attitudes towards English as an international language: The pervasiveness of native models among L2 users and teachers. In Sharifian, F. (ed.). *English as an International Language. Perspectives and Pedagogical Issues*. Bristol: Multilingual Matters.
- Macaro, E. (2001). Analysing student teachers? Codeswitching in foreign language classrooms: Theories and decision making. *The Modern Language Journal*, 85, (4), 531–48. doi:[10.1111/0026-7902.00124](https://doi.org/10.1111/0026-7902.00124)
- Mark. (2006, March 29). Re: Teaching English qualification [Forum comment]. Retrieved from <http://www.englishforum.ch/employment/388-teaching-english-qualification.html>
- Matsuda, A., & Matsuda, P. K. (2001). Autonomy and collaboration in NES/NNES teacher education. *The CATESOL Journal*, 13(1), 109-121.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, (1, 1), 30-46 (Correction, Vol. 1, No. 4, 390).
- Medgyes, P. (1999). *The non-native teacher*. Max Hueber Verlag, Germany.
- Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.
- Metzler, J., & Woessmann, L. (2010). The Impact of teacher subject knowledge on student achievement: evidence from within-teacher within-student variation. *IZA Discussion Papers 4999*, Institute for the Study of Labor (IZA). doi:[10.1016/j.jdeveco.2012.06.002](https://doi.org/10.1016/j.jdeveco.2012.06.002)
- Moradkhani, S., Akbari, R., Ghafar Samar, R., & Kiany, G. R. (2013). English language teacher educators' pedagogical knowledge base: The macro and micro categories. *Australian Journal of Teacher Education*, 38(10). doi:[10.14221/ajte.2013v38n10.7](https://doi.org/10.14221/ajte.2013v38n10.7)
- Muñoz, M. A., Prather, J. R., & Stronge, J. H. (2011). Exploring teacher effectiveness using hierarchical linear models: Student- and classroom-level predictors and cross-year stability in elementary school reading. *Planning and Changing*, 42(3/4), 241-273.
- Murphey, T., Asaoka, C. & Sekiguchi, M. (2004). Primary teachers co-learning English with their students. *The Language Teacher* 28(2) 15-18.
- Murphey, T. (2003). NNS primary school teachers learning English with their students. *TESOL Matters*, 13(4), 1-6.

- Murray, H. (2002) Developing language awareness and error detection. What can we expect of novice trainees? In H. Trappes-Lomax & G. Ferguson (eds) (2002). *Language in language teacher education* (pp. 187-198). Amsterdam: John Benjamins. doi: 10.1075/llt.4.13mur
- Nation, P. (2003). The role of the first language in foreign language learning. *Asian EFL Journal*, 5(2), 1-8.
- Newby, D., Allan, R., Fenner, A.-B., Jones, B., Komorowska, H. & Soghikyan, K. (2007). European portfolio for student teachers of languages. Graz: European Centre for Modern Languages of the Council of Europe. Retrieved from: <http://epostl2.ecml.at/>
- Noble, J. (2014). *Akaike's information criterion; model reduction*. [Lecture during course "Research design and multivariate data".] Retrieved from: <http://users.ecs.soton.ac.uk/jn2/statistics.php>
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237–257. doi:[10.3102/01623737026003237](https://doi.org/10.3102/01623737026003237)
- Oketch, M., Mutisya, M., Ngware, M., & Sagwe, J. (2010). Teacher math test scores, classroom teaching practices and student achievement in Kenya. African Population and Health Research Center (APHRC). Working Paper No. 45, Nairobi, Kenya, pp. 1-22.
- Organisation for Economic Co-operation and Development (OECD). (2009). Creating effective teaching and learning environments : first results from TALIS. [internal publication]: <http://www.oecd.org/about/publishing/corrigenda.htm>
- Palardy, G. J. ,& Rumberger, R. W. (2008). Teacher effectiveness in first grade: The importance of background qualifications, attitudes, and instructional practices for student learning. *Educational Evaluation and Policy Analysis*, 30(2), 111-140. doi:[10.3102/0162373708317680](https://doi.org/10.3102/0162373708317680)
- Park, S. & Oliver, S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38, 261-284. doi:10.1007/s11165-007-9049-6
- Peltzer-Karpf, A., & Zangl, R. (1998). *Die Dynamik des frühen Fremdspracherwerbs*. Tübingen, Germany: Gunter Narr Verlag.
- Phelps, G. (2009). Just knowing how to read isn't enough! Assessing knowledge for teaching reading. *Educational Assessment, Evaluation and Accountability*, 21(2), 137-154. doi:[10.1007/s11092-009-9070-6](https://doi.org/10.1007/s11092-009-9070-6)
- Phelps, G., & Schilling, S. (2004). Developing measures of content knowledge for teaching reading. *The Elementary School Journal*, 105(1), 31–48. doi:[10.1086/428764](https://doi.org/10.1086/428764)
- Poulson, L., Avramidis, E., Fox, R., Medwell, & Wray, D. (2002). The theoretical beliefs of effective teachers of literacy in primary schools: An exploratory study of orientations to reading and writing. *Research Papers in Education*. 16(3). 271–292. doi:[10.1080/02671520110058697](https://doi.org/10.1080/02671520110058697)
- Reber, A.S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.
- Richards, J. C. (1998). *Beyond training*. Cambridge, UK: Cambridge University Press.
- Rymarczyk, J. (April, 2013). *Early literacy in the foreign language classroom*. Paper presented at CLIL – explore the difference conference at the Kirchliche Pädagogische Hochschule Graz, Austria.

- Rymarczyk, J. (forthcoming). *Lese- und Schreibleistungen von Grundschulern bei simultaner und zeitversetzter Alphabetisierung Deutsch und Englisch*.
- Sadtono, E. (1995). The standardization of teacher trainees in EFL countries. *NETWORK*, 2(2), 17–24.
- Santibañez, L. (2006). Why we should care if teachers get As: Teacher test scores and student achievement in Mexico. *Economics of Education Review*, 25(5), 510–520. doi:[10.1016/j.econedurev.2005.08.001](https://doi.org/10.1016/j.econedurev.2005.08.001)
- Schweizerische Konferenz der Rektorinnen und Rektoren der Pädagogischen Hochschulen. (2013). Fremdsprachenunterricht in der obligatorischen Schule. Retrieved from: [http://www.edudoc.ch/static/web/arbeiten/sprach\\_unterr/ftktbl\\_sprachen\\_d.pdf](http://www.edudoc.ch/static/web/arbeiten/sprach_unterr/ftktbl_sprachen_d.pdf)
- Schweizerische Konferenz der Rektorinnen und Rektoren der Pädagogischen Hochschulen (SKPH). (2004). Empfehlungen der SKPH zum Koordinationsbedarf im Bereich der Fremdsprachkompetenzen zu Handen D-EDK. Bern. Retrieved from <http://www.cohep.ch/de/publikationen/empfehlungen/>
- Schwerdt, G., & Wuppermann, A.C. (2011). Is traditional teaching really all that bad? A within-student between-subject approach. *Economics of Education Review*, (30), 365–379. doi:[10.1016/j.econedurev.2010.11.005](https://doi.org/10.1016/j.econedurev.2010.11.005)
- Seidel, T. & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research*, 77, 454–499. doi:[10.3102/0034654307310317](https://doi.org/10.3102/0034654307310317)
- Senn-Carroll, M. (2006): *First Choice: One world, many people*. Lehrmittelverlag des Kantons Zürich.
- Shin, S. J. (2008). Preparing non-native English-speaking ESL teachers. *Teacher Development*, 12(1), 57–65. doi:[10.1080/13664530701827749](https://doi.org/10.1080/13664530701827749)
- Shono, S. (2004). *Good ESL teachers: From the perspectives of teachers and adult learners*. (Doctoral dissertation, The University of Texas at Austin). Retrieved from <https://www.lib.utexas.edu/etd/d/2004/shonos042/shonos042.pdf>
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–22.
- Siegel, J. (2014). Exploring L2 listening instruction: Examinations of practice. *ELT Journal*, 68(1), 22–30. doi:[10.1093/elt/cct058](https://doi.org/10.1093/elt/cct058)
- Song, M.-Y. (2008). Do divisible subskills exist in second language (L2) comprehension? A structural equation modeling approach. *Language Testing*, 25 (4), 435–464. doi:[10.1177/0265532208094272](https://doi.org/10.1177/0265532208094272)
- Spratt, M., Pulverness, A., & Williams, M. (2011). The TKT course: Modules 1, 2 and 3. Cambridge University Press. doi: [10.1017/CBO9781139062398](https://doi.org/10.1017/CBO9781139062398)
- Stronge, J. H., Ward, T. J., & Grant, L. W. (2011). What makes good teachers good? A cross-case analysis of the connection between teacher effectiveness and student achievement. *Journal of Teacher Education*, 62(4), 339–355. doi:[10.1177/0022487111404241](https://doi.org/10.1177/0022487111404241)
- Sugita, M. & Takeuchi, O. (2010): What can teachers do to motivate their students? A classroom research on motivational strategy use in the Japanese EFL context. *Innovation in Language Learning and Teaching*, 4(1), 21–35. doi:<http://dx.doi.org/10.1080/17501220802450470>
- Thornbury, S. (1999). *How to Teach Grammar*. Harlow: Longman.

- University of Cambridge Local Examinations Syndicate (UCLES). (2007). Cambridge young learners English tests handbook: Starters, movers and flyers. Cambridge: UCLES. Retrieved from <http://www.cambridgeenglish.org/images/84670-vocabulary-list.pdf>.
- Vandergrift, L. & Goh, C. (2011). *Teaching and learning second language listening: Metacognition in Action*. Routledge.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40, pp 191-210. doi:[10.1017/S0261444807004338](https://doi.org/10.1017/S0261444807004338)
- Viakinnou-Brinson, L., Herron, C., Cole, S., & Haight, C. (2012). The effect of target language and code-switching on the grammatical performance and perceptions of elementary-level college French students. *Foreign Language Annals*. doi:[10.1111/j.1944-9720.2012.01170.x](https://doi.org/10.1111/j.1944-9720.2012.01170.x)
- Von Ow, A., Husfeldt, V., & Bader-Lehmann, U. (2012). Einflussfaktoren für den Lernerfolg von Englisch an der Primarschule. Eine untersuchung in fünf Schweizer Kantonen und dem Fürstentum Liechtenstein. *Babylonia* 1(12), 52-57.
- Walker, R. (2006, June). Native or non-native? Or the best of both worlds? [Forum post]. Retrieved from <http://tesl-ej.org/ej37/f1.html>
- Wang, D. (2014). Effects of study abroad on teachers' self-perceptions: A study of Chinese EFL teachers. *Journal of Language Teaching and Research*, 5(1), 70-79. doi:[10.4304/jltr.5.1](https://doi.org/10.4304/jltr.5.1)
- Wayne, A., & Youngs, P. (2003). Teacher characteristics and student achievement gains: a review. *Review of Educational Research*, 73(1), 89–122. Retrieved from <http://www.jstor.org/stable/3516044>
- Wendelspiess, M. (2010, April 6). *Verwendung der obligatorischen Lehrmittel*. [Letter to regional (Zürich) board of education members]. Bildungsdirektion Kanton Zurich, Switzerland.
- Wright, S., Horn., & Sanders, W. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education*, 11, 57-67. doi:[10.1023/A:1007999204543](https://doi.org/10.1023/A:1007999204543)
- Zemp, B., & Strittmatter, A. (2007 November 12). Leitlinien Sprachkompetenzniveaus in der Grundausbildung der Lehrkräfte der Obligatorischen Schule. Stellungnahme des LCH zur EDK-Vorlage. Retrieved from <http://www.lch.ch>