University of Reading

Assessment for Learning in teaching English to Young Learners: teachers' understanding, classroom practice and impact on interactions

Agnieszka Maria Britton

Thesis submitted for the degree of Doctor of Philosophy

in Applied Linguistics

May 2015

Declaration

I confirm that this is my own work and the use of all materials from other sources has been properly and fully acknowledged.

A M Britton

1st May 2015

Acknowledgements

My thanks go firstly to my supervisors, Prof Paravneh Tavakoli and Prof Janet Enever (University of Umea, Sweden) for their continuous support, guidance and encouragement. Our relationship goes back to 2005 when I began studying for my Masters Degree at the London Metropolitan University. Janet and Parvaneh, you have been inspiring and supportive during that time, the years between MA and beginning my doctoral research and especially since I began to grapple with the intricacies of conducting postgraduate study in the complex and largely uncharted area of assessment in Teaching English to Young Learners. Your support and advice made it possible to stay on course.

My thanks go to Prof Suzanne Graham and Prof Clare Furneaux for the constructive critique and encouragement which I received from them at the confirmation stage at the University of Reading. I would also like to acknowledge the support in creating conditions for the development of postgraduate researchers within the Department of English Language and Applied Linguistics by Prof Jacqueline Laws.

I would also like to express my gratitude to the teachers and administrators at the study school. Thanks go especially to the teachers who were generous with their time and insightful comments and who opened their classrooms for me and made this study possible. Thank you.

Finally, my thanks go to my husband, Neil Britton, who, most importantly, survived this adventure and relentlessly supported me every step of the way.

Abstract

It has been suggested that Assessment for Learning (AfL) plays a significant role in enhancing teaching and learning in mainstream educational contexts. However, little empirical evidence can support these claims. As AfL has been shown to be enacted predominantly through interactions in primary classes, there is a need to understand if it is appropriate, whether it can be efficiently used in teaching English to Young Learners (TEYL) and how it can facilitate learning in such a context. This emerging research focus gains currency especially in the light of SLA research, which suggests the important role of interactions in foreign language learning.

This mixed-method, descriptive and exploratory study aims to investigate how teachers of learners aged 7-11 understand AfL; how they implement it; and the impact that such implementation could have on interactions which occur during lessons. The data were collected through lesson observations, scrutiny of school documents, semi-structured interviews and a focus group interview with teachers.

The findings indicate that fitness for purpose guides the implementation of AfL in TEYL classrooms. Significantly, the study has revealed differences in the implementation of AfL between classes of 7-9 and 10-11 year olds within each of the three purposes (setting objectives and expectations; monitoring performance; and checking achievement) identified through the data. Another important finding of this study is the empirical evidence suggesting that the use of AfL could facilitate creating conditions conducive to learning in TEYL classes during collaborative and expert/novice interactions. The findings suggest that teachers' understanding of AfL is largely aligned with the theoretical frameworks (Black & Wiliam, 2009; Swaffield, 2011) already available. However, they also demonstrate that there are TEYL specific characteristics. This research has important pedagogical implications and indicates a number of areas for further research.

Table of Contents

List of Appendices

List of Tables

List of Figures

List of Abbreviations

AfL	Assessment for Learning
AMTB	Attitudes/Motivation Test Battery
AO	Age of onset
AoL	Assessment of Learning
BAF	Barcelona Age Factor Project
BERA	British Educational Research Association
BSM	Bilingual Syntax Measure
CA	Conversation analysis
CBA	Classroom-based assessment
CBOS	Centrum Badania Opinii Spolecznej (English: Public Opinion Research Centre)
CC	Colour Coding
CEFR	Common European Framework of Reference
CP	Critical Period
CPD	Continuous Professional Development
CPH	Critical Period Hypothesis
EAL	English as an additional language
EAP	English for Academic Purposes
EFL	English as a foreign language
ELLiE	Early Language Learning in Europe Project
ESL	English as a second language
EVA	EVAluation of English in Schools Project
FA	Formative Assessment
FL	Foreign language
FLL	Foreign Language Learning
GUS	Glowny Urzad Statystyczny (English: Central Statistical Office)
ICS	'I Can' Statements
IDs	Individual Differences
IDZ	Intermental Development Zone
IFA	Informal formative assessment
IMWE	Indicating Mistakes Without Explanation (an AfL technique)
INT	Interview

IRF	Initiation – Response – Feedback/Follow up
ISCED	International Standard Classification of Education
ITT	Increased Thinking Time (an AfL technique)
L	Learner
L1	First language
L2	Second language
L3	Third language
LOR	Length of residence
LP	Learning partner (an AfL technique)
LRE	Language-related episode
MAN	Managerial mode
MAT	Materials mode
MM	Mind Maps (an AfL technique)
NNS	Non-native speaker
NS	Native speaker
NST	Next Steps (an AfL technique)
Q	Question
PPRR	Perfect Purple Red to Remember (an AfL technique)
ROWD	Record of Work Done
RQ	Research question
SA	Summative Assessment
SC	Success Criteria (an AfL technique)
SCH	Star Charts (an AfL technique)
SF	Smiley Faces (an AfL technique)
SGBM	Sharing Good and Bad Model (an AfL technique)
SLA	Second language acquisition
SP	Sensitive Period
SS	Skills and systems mode
SST	Sherif's Star (an AfL technique)
T	Teacher
TEFL	Teaching English as a Foreign Language
TEYL	Teaching English to Young Learners
THUD	Thumbs Up or Down (an AfL technique)

TL	Traffic Lights (an AfL technique)
TSAW	Two Stars and a Wish (an AfL technique)
YL	Young learner
YLE	Young Learners Exams
YLL	Young language learner
VYL	Very young learner
WALT	What Are We Learning Today? - type questions (an AfL technique)
WM	Working memory
WTC	Willingness to Communicate
ZPD	Zone of Proximal Development

Chapter One: Introduction

1.1 Background

In September 2008, I moved to Poland and started teaching English as a foreign language to young learners, having worked as a primary class teacher in London for a number of years. The language school which I worked for organised the courses by grouping learners based on their age into 5-6 year olds, 7-9, 10-11, 12-14, 15-17, 18 and older. I was surprised to find that many of my colleagues preferred not to teach children aged 11 and younger. The anecdotal evidence which I began to gather through staffroom discussions indicated that there was a rather specific reason why teachers preferred working with adolescents and adults. What my colleagues seemed to be saying was that children did not do, what I believe they described as, 'real' or 'serious' learning. Furthermore, they seemed to be of the opinion that it was difficult for teachers to demonstrate to parents at the end of a term what learners achieved because children's work remained at a basic level for a long time. A number of issues were certainly at play there. These included at least the following: accountability to parents, teachers' beliefs about what constitutes learning and achievement in a foreign language context in general and by young children in particular.

At that time, I also became aware that when my colleagues discussed assessment, they tended to refer predominantly to summative assessment, mostly pen and pencil tests, based on units of work sometimes as long as a whole term (4 months). In the autumn term of 2008, nobody in the staffroom or during in-service training sessions was talking about Assessment for Learning (AfL), which I had experience of using in the primary classrooms in England. I became interested in this area and soon discovered that very little was known about AfL in TEYL contexts.

Two years later, the school invited an TEYL researcher from a local university to deliver a whole day, in-service training to all teachers at the school. And so, the school embarked on its AfL journey and I found myself in a unique context for conducting research on AfL. In October 2011, I started collecting data for the current study.

The current chapter provides the background for the study reported in this thesis. It introduces the key terms (1.2), outlines the current study's context (1.3), the aims (1.4),

the research questions (1.5) and the methodology (1.6). The final section (1.7) provides an overview of the structure of the thesis.

1.2 Key terms

A number of key terms are used throughout the thesis. A brief definition of each of them is presented here. However, it should be noted that more detailed discussion of terminology is provided in the relevant sections of Chapter 2.

Assessment for Learning

The theoretical framework adopted in the current study was discussed by Black and Wiliam (2009). A useful definition, consistent with that framework, was proposed by the Third International Conference on Assessment for Learning in Dunedin, New Zealand in March 2009 as a 'part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance on-going learning' (Klenowski, 2009, p. 2). This approach includes five aspects: 1) clarifying and understanding the learning intentions and the criteria for success; 2) engineering effective classroom discussions and tasks to elicit evidence of student understanding; 3) providing feedback that moves learning forward; 4) activating students as learning resources for one another; 5) activating learners as the owners of their own learning (Black & Wiliam, 2009). These strategies can be pragmatically implemented in the classrooms by deploying a range of AfL techniques. The overarching purpose of using AfL is to facilitate on-going learning as opposed to only making judgments about what has already been learnt. The latter is the purpose of Assessment of Learning (AoL), a term which AfL is often contrasted with. AoL is normally conducted at the end of a unit of work to gain insights into students' attainment, which may later be used for reporting purposes (Wiliam, 2001).

Teaching English to Young Learners

There seems to be a number of terms used to describe contexts in which children learn English as a foreign language (Ellis, 2014). For the sake of consistency and simplicity the term *Teaching English to Young Learners* (TEYL) is used in this thesis. Johnstone (2009) identifies four models of implementation of teaching foreign languages to young learners. These include 1) coursebook based programmes; 2) more flexible programmes

where aspects of content from other curriculum areas are also taught through a FL; 3) an awareness raising model which usually does not aim to develop language proficiency but instead 'seeks to sensitize children to languages in general, with particular attention to the variety of languages that are actually used in the local community' (p. 35) and, finally, 4) bilingual and immersion programmes. The term TEYL is used in this thesis with reference to the first two models, where the FL is English. In these contexts, English is usually a separate subject and the amount of time in classroom is limited. A discussion of pedagogical considerations of TEYL is presented in Chapter 2.

Young Language Learners

In the context of this study, the term y*oung language learners* (YLLs) refers to children aged 7-11. This is consistent with the terminology used in the European context, where primary school children are usually referred to as y*oung learners* (YLs) and pre-school children as v*ery young learners* (VYLs) (Nikolov & Mihaljević Djigunović, 2011). It is useful to note that primary school starting ages differ in various educational systems. In Europe, children start primary education as early as the age of four (Northern Ireland) or as late as seven (e.g. Finland or Sweden) with the majority of educational systems requiring children to start primary school at the age of six (EURIDICE, 2014). Hence, any reference to primary age children is relative to the context in which it is being used. In the year when the data were collected (2011/12), and in the country where this study was based, children started primary education at the age of seven. It was lowered to six in 2014/15.

Teaching English as a Foreign Language

The term *teaching English as a foreign language* (TEFL) is also used in the current study. It refers to contexts where English is a curriculum subject, i.e. according to Models 1 and 2 in Johnstone (2009). It should be noted that the term TEFL does not refer to any specific age group. Hence, it includes primary, secondary and tertiary education as well as private sector language schools. This use is consistent with the widely adopted terminology in the field of English language teaching (Ellis, 2014). The language which is taught in such a context is referred to as a *foreign language* (FL).

English as a Second Language

The term *English as a second language (*ESL) is used when discussing contexts where English is the means of instruction through which learners are taught the curriculum. This includes contexts where either all or almost all the teaching is delivered through English. This definition corresponds with immersion and bilingual teaching as discussed by Murphy (2014, Chapter 6) and Johnstone's (2009) Model 4. The language taught in such contexts is referred to as a *second language* (L2).

English as an Additional Language

The term *English as an additional language* (EAL) is used in the educational systems of the British Isles to refer to the language needs of children who are educated in mainstream, English medium schools but who do not speak English as their L1. In EAL contexts, children are expected to learn the curriculum content alongside acquiring L2. Cummins (1986) argues that it can take a child as long as 5-7 years to acquire L2 to a level which enables learners to engage with the academic language of the school.

First language

The term *first language* (L1) is used with reference to the language(s) which children learn from birth, usually in family homes and which they acquire, at least in oral form, before they enter the educational system. In some contexts, L1 is also known as the *home language* or *mother tongue*. This may include one (monolingualism), two (simultaneous bilingualism) or more (multilingualism) languages (Murphy, 2014).

1.3 Research context

This study was conducted in a private language school in Poland which belongs to a chain of well-established schools of English as a foreign language with over seventy years of experience worldwide and branches in major Polish cities. The model of teaching followed the first model described by Johnstone (2009), namely general topics were used to teach English, the curriculum relied on a coursebook and the amount of time spent in class was limited to 120 minutes per week.

In state education in Poland one foreign language is taught from Year One of primary school (learners aged 6-7) and a second foreign language is introduced on a compulsory

basis in Year Four (10-11 year olds). One of the two foreign languages must be English. It is also popular for parents to enrol children on language courses in private language schools. Section 3.1 provides detailed discussion of the context.

1.4 Research motivation and aims

Despite the claims that AfL can play an important role in raising achievement (Black & Wiliam 1998; Wiliam, 2011), its implementation in TEFL in general, and in TEYL classrooms in particular, has rarely been researched. AfL continues to receive a lot of researchers' and policy makers' attention globally (Bennett, 2011; Klenowski, 2009). Simultaneously, language teaching at primary school level seems to be widespread worldwide (Pinter, 2011). For instance, in Europe the average number of foreign languages learned per pupil at International Standard Classification of Education (ISCED) Level 1, i.e. in primary education, increased to 0.8 in 2011 from 0.5 in 1998 (Eurostat, 2013). Yet, little is known about assessment which could capture foreign language learning (FLL) in childhood (McKay, 2006), especially the type of assessment which has a formative function, i.e. aims to move learning forward.

The current study aims to address that gap in research and to contribute new insights to the field of TEYL by investigating AfL in a TEYL context. The originality of this study lies in the fact that it explores a largely under-researched area and makes connections between two areas of research: foreign language learning, on the one hand, and research into AfL conducted in different educational settings, on the other.

1.5 Research questions

An overarching aim of the present study is to report how teachers understand and implement AfL in a TEYL setting and what impact of AfL on interactions can be observed in the classrooms of 7-11 year olds. Elaborating further on that aim, the study seeks to:

- Report how teachers understand AfL in a TEYL context and whether any TEYL-specific features can be identified according to their understanding.
- Identify classroom embedded AfL techniques that are used in TEYL classrooms.
- Identify if the type, frequency and purpose of using AfL techniques vary across a) age groups; b) teachers; c) different language skills and d) over time.

- Develop insights into the impact that AfL may have on learning through impacting on interactions in TEYL lessons.

The above aims were used to develop the research questions (RQs) that inform the current study:

- **RQ1:** How do teachers understand AfL after receiving a limited amount of training and being encouraged to use AfL techniques for at least one academic year when teaching English to young learners aged 7-11?
- **RQ2:**
  - **2.1:** How do teachers translate their understanding of AfL into classroom practice in a TEYL context with students aged 7-11 in a private language school in Poland?
  - **2.2:** Do teachers report any changes in their practice of using AfL over time?
- **RQ3:** What is the observable impact of AfL on classroom interactions in a TEYL context?

These research questions are discussed further in Section 2.4.

1.6 Research methodology

As indicated so far and discussed in Chapter 2, there is little research available about AfL in TEYL contexts. Hence, the study reported here was of an exploratory and descriptive nature. This required careful consideration of methodological issues in order to design a model which would offer valid insights into this largely under-researched area.

A mixed methods approach has been adopted in the present study, incorporating quantitative and qualitative approaches. It is believed that by aiming for a good fit between the research questions and the research methods, the design of the current study ensures internal validity. To ensure fitness for purpose, the adopted interpretive framework permitted gaining insights into relationships between the practical implementation of AfL and interactions in TEYL classrooms, adopting quantitative methods. Furthermore, it offered an opportunity for in-depth analysis and thus enabled a deeper, interpretive understanding of the researched issues through adopting qualitative methods. A detailed discussion of research methodology is reported in Chapter 3.

1.7 Organisation of the thesis

This thesis is organised in six chapters. The current chapter provides an introduction to the study. It is followed by a review of the relevant literature in Chapter 2. Subsequently, Chapter 3 discusses methodological considerations, including the study design, the pilot study and methods and procedures used for data collection and analysis. Subsequently, Chapter 4 reports the findings of the present study. Discussion of the findings is provided in Chapter 5. Finally, the implications and limitations of the present study are discussed in Chapter 6.

Throughout the thesis, where appropriate, the discussion may refer the reader to another section within this thesis. This is done by providing the number of the relevant section in brackets, e.g. (5.2) refers the reader to Section 5.2 in Chapter 5.

Chapter Two: Literature review

2.1 Introduction and structure of the chapter

Teachers of children studying English as a foreign language face the challenge of assessing progress in an effective and appropriate manner according to age. The results of their assessments provide information about what the learners have learnt and inform decisions to be made and steps to be taken in order to move the learning forward. When assessment serves the purpose of facilitating learning, this is where AfL practices occur. The aim of this chapter is to review what is already known about the intersection of learning and assessment.

To discuss what is already known about assessment that promotes learning, it is first important to clarify what is understood by learning and, specifically for the purpose of this thesis, the concept of FLL in childhood. However, any discussion of learning processes in childhood would not be complete without first considering the cognitive development of young learners. Hence, this chapter begins by exploring the theories of cognitive development and of foreign language learning in childhood that are relevant to understanding the conditions for effective language learning in classes of 7-11 year olds. The review in Section 2.2 demonstrates that there is a need for, and value in, gaining a deep understanding of the processes involved in FLL in childhood. Such insights can inform assessment practices that are appropriate to TEYL contexts and that not only provide summative information about the progress that learners make but, perhaps more importantly, could contribute to facilitating FLL.

Section 2.3 begins by discussing the theoretical framework of AfL adopted in the present study. The review indicates that a well-established theoretical framework and consistency in the use of terminology is lacking in the AfL literature. To address this gap, the framework including selected terminology is discussed in detail. Then, the attention shifts to reviewing what is already known about assessment in TEYL contexts. The review indicates that research from TEYL contexts provides insights predominantly into summative assessment. Only very few studies seem to have explored its formative function. The final part of Section 2.3 shifts to reviewing what is known about AfL in similar educational contexts. These include TEFL with older learners and EAL in primary schools. The aim is to identify issues concerned with the implementation and impact of AfL in similar educational contexts.

The chapter concludes with Section 2.4 where the research questions driving the study are identified.

2.2 Young Language Learners

As McKay (2006) accurately observed:

> (i)t is axiomatic that the way that children learn best be reflected in the way that they are assessed, and the knowledge of how young learners learn language is therefore fundamental for those involved in the language assessment of young learners (p. 47).

Hence, it is important for the present study to review what is known about how children learn languages before exploring assessment in a TEYL context. This section begins by reviewing influential theories of child cognitive development and research into the processes of FLL in childhood (2.2.1). The attention then shifts to considering the socio-cultural theory of learning and the impact of interactions on FLL in childhood (2.2.2). Finally, the issue of affect and its relationship to learning is reviewed in Section 2.2.3.

2.2.1 Cognitive development in childhood and FLL

Children think, function and learn differently from adults and they develop their cognition as they mature. The most influential theories of child cognitive development have attempted to capture the nature and timing of those changes. This section considers age as a factor in cognitive development (2.2.1.1), the starting age of instruction and its relationship to FLL (2.2.1.2), and language processing in childhood (2.2.1.3). It is believed that, by reviewing these areas, it will be possible to identify important considerations for implementing assessment in TEYL contexts.

2.2.1.1 Stages of cognitive development

One of the most influential, though not unchallenged, theories of learning was developed by Jean Piaget (1896-1980): a Swiss child psychologist. Basing the theory on observations of his own children, Piaget aimed to identify aspects of intelligence development that could be generalised to other children. Piaget's four stages of development (Table 2.1) are demarcated with significant changes in thought processing, summarised in Column 3 of Table 2.1.

Table 2.1: Stages of development according to Piaget; adapted from Pinter, 2011

1. Stage name	2. Age	3. Characteristics of the stage
Sensori-motor	Birth - 2	Goal oriented behaviour Imitation Repetitive motor habits Curiosity
Pre-operational	2 – 7	Animism, egocentrism, centration (see discussion below for definitions)
Concrete operational	7 – 11	Development of logical thinking, hierarchical classification Understanding of causality, reversibility of processes Development of symbolic thought and analogy
Formal operational	11 – 12 and older	Ability to complete formal operations without relying on concrete objects Ability to deduce and hypothesise Development of abstract thinking and systematic investigation

The concrete operational and formal operational stages are relevant to the cohort of children in the current study.

According to Piaget, the concrete operational stage is demarcated by a so-called intellectual revolution, which occurs at around the age of seven (Wood, 1998, p. 23). At this stage, children become aware that one task may be approached or seen from more than one perspective and are able to categorise objects according to certain criteria, recognising that the same object may belong to more than one category. Furthermore, children begin to display less egocentrism, centration and animism. *Egocentrism* is the inability to distinguish between the subjective and the objective (Berk, 2000). *Centration* refers a tendency to focus on one aspect of a task and ignore the others (ibid.). Finally, *animism* entails assigning animate qualities to inanimate objects (ibid.). As they mature, children start to develop the ability to think in abstract terms, deduce and hypothesise. This marks the final developmental stage in Piagetian theory, which is referred to as the formal operations stage and happens around the age of eleven or twelve.

Piaget's work has attracted a considerable amount of research and discussion. The main critique centres on the methodological issues. Most significantly, the design of tasks used by Piaget was evaluated by other researchers as inaccessible and not ecologically

authentic, especially for younger children (Donaldson, 1978). Ecologically authentic tasks engage children in activities which they routinely perform in real life such as playing games, drawing and talking about age-relevant topics (Turek, 2013). Some researchers have stressed that even adults might have performed badly on those of Piaget's tasks that were designed to test formal operational thinking in the cases where the instructions were misleadingly expressed (Winer, Craig & Weinbaum, 1992). Further critique is connected with the ways in which the experiments were conducted: that they were ambiguous or potentially misleading in the connections between the actions of the interviewer and the questions; and that the possible impact of contextual factors might have not been accounted for. For example, schooling can facilitate the development of concrete and formal operational thinking and, together with other life experiences, often embedded in culture and unique to each child, can have an impact on how operational thinking develops (Ceci & Roazzi, 1994). Wood (1998) argued convincingly that there seem to be further changes in thinking past the age of 12, especially during puberty, and that some of those may result from development in literacy, since reading and writing can aid such changes. Furthermore, Piaget's experiments (Piaget, 1969) were conducted without taking social interaction into account, which excluded any effect that interaction could have on children's performance. Overall, it is currently believed (e.g. Donaldson, 1978) that Piaget underestimated the abilities of children.

This section has indicated that learners aged 7-11, the age of the participants in the current study, might require support through the use of concrete props and visual aids before they reach the formal operational stage. To aid this, it may be important for teaching and assessment to be located in a context that is meaningful for children, through the provision of such activities as storytelling or songs. Additionally, it may be appropriate to use teaching as a context for assessment, as opposed to employing delayed assessment procedures that lack a learning context. In other words, it may be useful to integrate assessment procedures into classroom work.

However, the stages of development may not be as easily catalogued as Piaget (1969) interpreted them. Additionally, they may differ depending on the individual characteristics of children. Hence, it seems worth considering more closely what research has shown about age as a factor in FLL and the relationship between age and individual differences (IDs). Consequently, cognitive IDs – cognition, memory and metacognition

– are discussed in Section 2.2.1.3 while affective IDs – anxiety, motivation, attitude and self-concept – are discussed separately in Section 2.2.3.

Piaget's theory and the outcomes of research based on it suggest that explicit teaching and testing of language forms may not be appropriate in a TEYL classroom if this requires learners being able to process abstract rules, i.e. before they reach the formal operational stage. Furthermore, the complexity of language used in assessment tasks seems to be an important consideration where the language of instructions could serve to inhibit learners' performance. All these considerations suggest that the structure of tasks can have an effect on young learners' performance, which has direct implications for assessment. Hence, in Section 2.2.2, the review will include studies that have investigated task characteristics and their impact on performance.

2.2.1.2 The age factor in L2 learning

While there seems to be an agreement that age is an important factor in language learning, 'scholars have not been able to establish the exact pattern or nature of age-related change, let alone identify the specific causes and mediators of the process' (Dörnyei, 2009, p. 233). This section discusses studies that provide insights into how different ages of commencing instruction affect FLL in childhood, especially at ages 7-11. Overall, the research findings suggest that older beginners have an initial advantage in L2 learning, while younger learners tend to outperform the older ones in the long term. Importantly, however, the picture that emerges from the research suggests that factors other than age alone may play an important role in determining success in language learning.

2.2.1.2.1 Critical Period Hypothesis

One of the widely researched theories of the relationship between age and language learning is the Critical Period Hypothesis (CPH). The theory was developed in late 1950s and 1960s and was based on the argument that the brain loses its plasticity in childhood and that 'children are better second language learners than adults because their brains are specially organized to learn language, whereas those of adults are not' (Bialystok & Hakuta, 1999, p.176). Critical periods (CPs) are understood to have three-stage windows of time of high sensitivity to a certain type of learning. The three stages are onset (the beginning of sensitivity), peak (when sensitivity is high) and offset (a decline in sensitivity). This understanding of CPs is problematic in second of foreign language learning because the offset stage of sensitivity to language learning does not seem to be

constant and with maturation a more gradual decline towards a levelling out can be observed. For example, Long (2013) argues that from early adolescent years a slow and less noticeable decline continues until death and depends 'only partly on age, and more on other factors, such as amount of exposure, usually operationalised as length of residence (LOR), and the proportions of L1 to L2 use' (p. 4). Hence, more accurate seems to be Long's (ibid.) suggestion that the nature of development and changes in the ability to learn a language is better captured by the notion of sensitivity rather than critical periods. Bialystok and Hakuta (1999) refer to sensitive periods (SPs) as a weaker version of CPs. The notion of SPs suggests that for those who start after a certain age, learning a foreign or second language becomes significantly less successful and that achieving a native-like proficiency may not be possible (Granena & Long, 2013b).

There seem to be three issues of importance here. First, it seems useful to investigate whether any research provides evidence that there exist sensitive periods which end with childhood and after which foreign language learning is less successful. The second issue is related to factors that may impact on the level of success in FLL at different ages. These could include neurological maturation of the brain; factors connected to individual learner differences (attitudes, motivation, anxiety, self-concept, aptitude, memory); or those social in nature connected to the amount and/or type of exposure to L2. The third issue is defining how success is measured. The majority of research in this area has adopted a 'native-like' mastery of English as the measure of success. This is also referred to as *ultimate attainment*. It is hoped that by considering these three areas, the present review will identify age-related issues which may be of importance to implementing AfL in TEYL classrooms.

With regards to the notion of ultimate attainment, it is important to note that, as Nikolov and Mihaljević Djigunović (2011) notice, in TEYL programmes 'YLs are not expected to achieve native levels of proficiency' (p. 97). Hence, the notion of ultimate attainment seems problematic in such contexts. Presumably, young children's L2 should be compared to the language of native speakers (NSs) of a similar age. This would entail accepting non-grammatical forms uttered by a learner as satisfactory in cases when NSs of a similar age tend to make similar mistakes. In order to establish that, language corpora for different age (groups) would be needed. However, research exploring ultimate attainment did not consider children's use of L2 and did not compare it to the language used by child NSs. Instead, the majority of research into ultimate attainment has been

conducted with adults whose language acquisition history and language proficiency were analysed. Not surprisingly, research focusing on ultimate attainment has been carried out predominantly in immersion and immigrant settings, which means that the findings are not easily applicable to TEYL classrooms and as such are only briefly summarised in Section 2.2.1.2.2.

Following from Krashen, Long and Scarcella (1979), many studies have investigated sensitive periods by attempting to measure success in FLL as *the rate of progress*, and not just the ultimate attainment, and by comparing the results of learners who started at different ages. The studies exploring the impact of the starting age on progression in FL indicate that older learners achieve more at the initial stages (e.g. Muňoz, 2006) than younger beginners do, whereas younger beginners may achieve native-like proficiency in the long term, especially in pronunciation, accent (e.g. Flege, Munro & McKay, 1995) and grammar (e.g. Birdsong & Molis, 2001). However, research also suggests that factors other than age, namely formal training, personal motivation and access to authentic input, may account for exceptional success in foreign or second language learning in developing native-like pronunciation for instance, even if it begins after puberty (Bongaerts, 1999). These studies indicate that age is an important factor in FLL but other factors may also play a significant role.

The following two subsections review studies that offer insights into rate of progress and ultimate attainment. The aim is to consider the three issues related to the notion of sensitive periods that have been identified above as relevant to the current study. It is also important to note that the research into age as a factor in language learning provides insights into how young children learn foreign and second languages. Hence, it seems useful to explore this area to inform assessment practices in TEYL contexts.

2.2.1.2.2 Ultimate Attainment – younger learners' advantage

A number of studies have looked at the issue of ultimate attainment. Much research has adopted the age of onset (AO) as the variable that can predict ultimate attainment. Granena and Long (2013a) report that the typical value attributed to AO in predicting variance in ultimate attainment is around 30% (p. ix). Overall, the findings of AO research are rather complex but do indicate some areas in which an earlier start seems to yield higher results in the mastery of pronunciation and accent (Fledge et al. 1995; Flege, Yeni-Komshian & Liu, 1999; Long, 2005;) and of grammar (DeKeyser, 2000) in the long

term. However, some researchers claim that contextual factors are a stronger predictor than age in the acquisition of grammar (Jia & Fuse, 2007) or pronunciation and accent (Moyer, 2004).

Regarding accent and pronunciation, Long (1990, 2005) suggests that the peak period for acquiring native-like pronunciation falls between the ages of 0-6, with offset at 6-12 and becoming significantly more difficult after the age of 12. Flege et al. (1995) concluded that an age of arrival of Italian speakers between 3.1 and 11.6 years in Canada allowed them to develop native-like accents. Later, Flege et al. (1999) obtained similar results with Korean immigrants in the US.

Research has also pointed to a number of periods in a child's development that are sensitive to developing the specific language areas of phonology, morphosyntax, lexis and collocations. Regarding lexis and collocations, evidence for SPs has been provided by a number of studies (Munnich & Landau, 2010; Spadaro, 2013) showing the following age-related performance: peak 0-6, offset 6-12 and significantly less effective past 12. Regarding morphosyntactic development, Long (1990) quotes the following: peak 0-6, offset 6 to mid-teens and significantly difficult past the age of 16/17. These findings corroborate with those reported by DeKeyser (2000) who, in a study of grammatical accuracy judgements by 57 adult Hungarian speakers, found that, up to the age of 17, ultimate attainment strongly correlated with the age of beginning to acquire L2 but not starting at the age of 17 or later. However, Jia and Fuse's (2007) study, which accounted for contextual factors in investigating the development of grammar in ten immigrant children, concludes that, after 5 years of immersion, the language environment predicted success in acquiring L2 better than the starting age, which points to the importance of contextual factors at school and at home.

A recent study by Granena and Long (2013b) considered both the potential amount of L2 exposure in terms of length of residence (LOR) as a contextual factor, and language aptitude in investigating the scope and timing of maturational constrains in acquiring Spanish by 65 Chinese speakers. The study included a control group of 12 native speakers. The participants were divided into three groups based on AO (3-6, 7-15, 16-29 year olds). The findings confirmed that sensitive periods end first for phonology, followed by lexis and collocations and later for morphology and syntax. Language aptitude was measured using the computer based Swansea Language Aptitude Test

(Meara, 2005). Significant correlations were found between *language aptitude* and the acquisition of phonology and, in the oldest group, between aptitude and lexis and collocations. The only significant correlation between LOR and ultimate attainment was found in the domain of lexis and collocations and only in the groups with AO 3-6 and 7-15. This study provides evidence that AO is an important predictor of ultimate attainment and that other factors, such as length of exposure to L2 and language aptitude, can also play an important role but their impact may differ according to language domain and age.

Although research in this area has a long tradition, with the exception of the Granena and Long's (2013b) study, it has mostly focused on documenting sensitive periods. Relatively little research has looked into the factors that might shape those sensitive periods (Granena & Long, 2013a). Effectively, we do not have enough evidence to demonstrate whether maturational processes in the brain account for the SPs.

Of relevance to the current study are the three issues connected to SPs, specified at the beginning of this section. First, research on SPs seems to suggest that there exist windows of time during which learners are more likely to acquire native-like proficiency and that these may be somewhat different for phonology, morphosyntax, lexis and collocations. This point seems especially useful in the context of the current study, in which the participants were aged 7-11. Their sensitive periods for all language domains identified by the research were likely to be ongoing. With that in mind, it seems important that the teaching and assessment methods used in this context should capitalise on that fact by providing ample opportunities for exposure to the FL. Secondly, some studies suggest that individual characteristics and contextual factors may contribute to achieving a NS level of L2 but that they may play a more important role in cases when AO happens after the offset of SPs. However, this is not the case with the participants of the current study. Finally, adopting ultimate attainment as the measure of success does not seem appropriate in TEYL contexts that are characterised by a short length of exposure to a FL. This has implications for the construct of assessment in TEYL contexts.

Although the studies quoted above indicate that AO is an important factor in predicting ultimate attainment, it seems crucial to also consider the context in which AO occurs. Of special interest to the current study is the distinction between naturalistic settings and instructed foreign language contexts. The majority of research on SPs has been conducted in contexts where language is acquired in a naturalistic setting, whereas the

current study has been conducted in an instructional context with limited exposure time. Hence, the findings of studies quoted in the current section cannot be applied directly to the context of the present study. Nonetheless, they are vital to include in the review as they provide useful insights into language learning in childhood. More relevant to the present study, in terms of contextual similarity, seem to be studies that provide evidence for the initial advantage of older beginners in instructional contexts. These are reviewed in the following section.

2.2.1.2.3 Rate of progress – the advantage for older beginners

This section explores studies indicating that older learners have an initial advantage over their younger counterparts in terms of rate of progress. The aim is to tease out what this research reveals about how children learn FLs. The review includes studies which explored:

- The benefits of early start (Burstall, Jamieson, Cohen, Hargreaves, 1974; Muňoz, 2006; Vilke & Vrhovac, 1995)
- Learning the form of a language (Dimroth, 2008; Garcia-Mayo & Garcia-Lecumberri, 2003)
- The relationships between L1 and FLL (Knell, Haiyan, Miao, Yanping, Siegel, Lin & Wei , 2007; Mihaljević-Djigunović, 2010; Wilden & Porch, 2014)

The benefits of starting FL instruction early were investigated in England with children who started learning French at the age of 8 and 11 (Burstall et al., 1974). The measurements of the learners' language proficiency were conducted at the ages of 13 and 16. The results of the first measurement indicated that learners who started earlier outperformed the other group in listening and speaking. At the time of the second measurement, only listening skills were demonstrably higher in the younger beginners. It was concluded that older beginners' rate of progress was faster and hence an earlier start of instruction did not result in higher proficiency at 16; hence there was no argument for funding early language learning provision in schools. However, there was little account in the study of other contextual factors that might have affected the results. Most significantly, the quality of teaching at the primary school and the continuity from primary to secondary school were not evaluated. Moreover, the measurements focused on language skills (listening, speaking, reading and writing) but did not provide insight into other areas such as pronunciation or accent.

A more recent study that explored a similar issue was the Barcelona Age Factor Project (BAF) reported by Muñoz (2006). It was a large scale, longitudinal study that included about 2000 learners of English. A number of measures including speaking, listening, reading and writing were deployed after 200, 400 and 700 hours of FL input. The results were compared for participants who started learning English at the ages of 8 and 11. The findings evidenced that, initially, the older learners outperformed their younger counterparts on all measures. This was attributed to their cognitive development and schooling, especially the development of morphosyntactic ability which was detected at around the age of 12. Based on observations that the younger learners gradually caught up with their older counterparts in tests that measured implicit learning, Muňoz (ibid.) noted that children who started learning English at the age of 8 'seem to favour and be favoured by implicit learning' (p. 32) but this type of learning may take longer. The quality of teaching was not evaluated in the BAF project, so it is not possible to establish if younger children were offered opportunities for implicit learning. Hence, the possibility remains that factors other than age impacted on the rates of progression.

Other studies have investigated how children learn the form of language in childhood. For example, Garcia-Mayo and Garcia-Lecumberri (2003) focused on Spanish-Basque L1 speakers learning English as the third language (L3). The investigation centred on gauging whether the age of starting instruction impacted on the learners' judgements about grammar. The study concluded that participants who started learning English at the age of 11-12 achieved significantly higher scores than those who began at the age of 8-9. The same was reported after the younger group received an additional 198 hours of teaching. The weaker performance of the younger group could indicate that younger children are more focused on meaning rather than form. This may be explained by the findings of the BAF project (Muñoz, 2006), which indicates that children learn languages implicitly and, hence, may not be explicitly aware of form and as such are likely to underperform on grammaticality judgement tests.

Another study investigating how children learn language form compared the acquisition of negation and finiteness in L2 German by two untutored children (aged 8 and 14) with one another and with an adult (Dimroth, 2008). Dimroth concluded that the younger child differed from the adolescent, who was similar to the adult, in that the 8 year old acquired the target structures faster than the 14 year old, in a different order and seemed to do so without analysis. Although this study was not conducted in a school setting, it offered a

useful insight as it explored the implicit acquisition of linguistic form, which was shown to be beneficial for young learners (Muñoz, 2006). Dimroth's (ibid.) results suggest that language form can be learnt implicitly. However, presumably this requires a significant amount of exposure to FL.

The next study investigated the relationship between the age of beginning instruction in a FL and achievement. It analysed the listening and reading skills of over 6500 children in Germany (Wilden & Porsch, 2014). Two measurements were conducted. In 2010, children aged 9/10 who had started receiving FL instruction at the age of 8 were assessed; in 2012, 9/10 year olds who had started instruction at the age of 6.5 were assessed. The results suggested that the development of reading skills in L1 was an important factor in successful FLL. This finding corroborates with the results of the study reported above (Muñoz, 2006) which suggest that schooling (including L1) may have an impact on FLL.

The relationship between L1 and FLL has been explored by other researchers. For example, Mihaljević Djigunović (2010) compared the achievements of learners aged 14 who had begun receiving FL instruction at different ages. She reported interconnections between their development in L1 and FL reading, writing and listening skills. The results indicate that the strongest relationship was found between the reading skills in L1 and FL. Another interesting insight into the relationship between L1 and FLL was found in a study by Knell et al. (2007) who compared achievement in a number of tests measuring vocabulary and phoneme recognition as well as the letter knowledge of children in immersion and non-immersion programmes in China. The 183 participants were 6-9 years old. Knell et al. (ibid.) reported that the phonological awareness and letter recognition in pinyin (a phonetic system of transcribing Mandarin into Latin alphabet) could facilitate the learning of L1 (Chinese) and FL (English). These studies indicate that the development of literacy that occurred as the children matured and were educated could facilitate FLL, suggesting that one of the strongest related factors is reading. Hence, skills and knowledge of L1, especially reading, may impact on a faster rate of progress in FL in older children. This suggests that contextual factors, such as the age of beginning schooling, may be important for successful FLL in childhood and should be considered in assessment.

As indicated by the research reviewed so far, adolescents initially learn language faster than younger children. This advantage could be explained in terms of the development

of cognitive abilities (see also Section 2.2.1.3) alongside L1 literacy skills. Their younger counterparts were shown to benefit from implicit learning. Hence, it seems useful to consider whether assessment practices appropriate for younger children should focus on measuring implicit learning and formulaic language, shifting towards a greater emphasis on analysed language as learners become more mature. This seems to have implications for whether feedback on performance should be implicit or explicit and it raises the question of whether learners should be required to notice if there is a mismatch between their own performance and what is considered correct in developing a FL. These issues are further discussed in Section 2.2.1.3.1.

2.2.1.3 Processing language in TEYL contexts

The research that provides insights into how information is processed by second and foreign language learners lies on the intersection of linguistics and cognitive psychology. A number of IDs that could impact on success in FLL, such as attention, memory, aptitude and metacognition, have been examined. In FL contexts, investigations have been conducted largely with adults, while insights into internal processing in childhood come predominantly from non-TEYL contexts. This section aims to review the research that provides insights into what is known about the way in which children aged 7-11 process foreign languages as they learn them. It is believed that this will inform how feedback on performance can effectively and appropriately be given in TEYL contexts in such a way that supports learning.

2.2.1.3.1. Attention

It is important for the focus of the present study to consider two issues related to *attention*. The first is on how attention develops in childhood. The second is how directing attention or noticing is related to feedback provision and learning. It is hoped that by reviewing the research in these areas it will be possible to identify which considerations related to attention are important in implementing assessment in TEYL contexts.

As children mature they develop their attention capacity. Research in this area confirms that younger children have short attention spans, which become longer as children improve in cognitive inhibition skills (McKay, 2006) and their ability to ignore unnecessary information develops (Ridderinkhof & van der Molen, 1997). In that process, children become able to control their own attention. 'Control refers to the level

of attention and inhibition recruited during cognitive processing' (Bialystok, 2001, p.15). Hence, their ability to sustain attentional focus on the task at hand increases.

Children also develop the skill of managing their own attention as they learn to use strategies. These could include focusing, dividing or switching attention (Gopher, 1993). A study conducted by Vurpillot (1968) demonstrated that children develop strategies to guide their attention at around the age of 6. In that study, learners aged 3-9 were asked to decide if two pictures of a building were identical. Vurpillot (ibid.) observed that older learners were able to compare the windows in the building in a more systematic way than could younger learners. It has also been argued that YLLs devote greater attention to meaning than to form and shift their attention to meaning even when the teacher tried to focus it on form (Bialystok, 2001).

The above research suggests that assessment in TEYL contexts should take into account the attention span of children and ensure that assessment procedures do not require young learners to draw on attention strategies that they may not have developed. For example, 'spot the difference' tasks or problem solving puzzles could be too challenging for younger learners in the early years of primary school (Pinter, 2011). Hence, it seems that understanding how the structure of a task is related to performance is important. This is further discussed in Section 2.2.2.

Pinter (ibid.) comments on another important issue connected with attention in childhood stating that '(b)efore any information can be stored or processed, individuals must notice it or attend to it' (p. 22): thus referring to the Noticing Hypothesis proposed by Schmidt (1992, 2010), who argued 'that input does not become intake for language learning unless it is noticed, that is, consciously registered' (Schmidt, 2010, p.721). In its weaker form, the hypothesis suggests that learners can learn more if they notice specific aspects of what is being taught (Robinson, 1995). For example, to learn how to pronounce, learners need to attend to the sounds of the language as opposed to paying attention to meaning or to other aspects of language. In Schmidt's (2010) opinion noticing is more limited than understanding. The notion of 'noticing the gap' proposed by Schmidt and Frota (1986) is of interest to this study: i.e. 'the idea that in order to overcome errors, learners must make conscious comparisons between their own output and target language input.' (Schmidt, 2010, p. 724).

Empirical studies in adult contexts have examined attention and awareness as two cognitive processes that mediate input and L2 development through interaction. This body of research has provided support for the weaker form of the Noticing Hypothesis. For example, Mackey (2006) found positive relationships between interactional feedback given in class to adult ESL students: L2 learning of questions and learners' reports on noticing the question forms. In a mixed-method study which demonstrated that more awareness could lead to more recognition and accurate written production of the noticed forms, Leow (2000) reported finding evidence for the important role of attention and awareness in mediating L2 processing of targeted irregular verb forms in Spanish. Rosa and O'Neill (1999) demonstrated that the level of awareness of Spanish conditional forms was positively related to the intake of those forms. Although this body of research is relatively well-established very, little insight is available into whether similar claims may be made about young learners.

The relevant research into young learner contexts explored the Noticing Hypothesis by investigating the impact of focus on form (incidental or planned) and implicit vs. explicit corrective feedback. Focus on form is defined as 'briefly drawing students' attention to linguistic elements (words, collocations, grammatical structures, pragmatic patterns, and so on), in context, as they arise incidentally in lessons whose overriding focus is on meaning, or communication' (Long, 1996, p.40). In EFL classes of Spanish adolescents (aged 14-15), Alcón (2007) investigated teachers' incidental focus on form and its effectiveness in noticing and found, from the students' reports, that planned focus on form led to a greater degree of noticing but that both planned and incidental feedback helped learners to learn more vocabulary.

Implicit feedback can be given in the form of either recast 'where a teacher reformulates a learner's non-target-like form' (Oliver & Mackey, 2003, p. 519) or negotiation strategies which include repetitions, clarification requests and comprehension checks. In a French immersion context in Canada, Lyster and Ranta (1997) reported a study of over 18 hours of audio-recorded lessons from the classes of 9-10 year olds. Their findings suggest that teachers used recast to correct phonological and grammatical errors and negotiation of meaning to provide feedback on lexis. They argued convincingly that children were unlikely to notice implicit feedback in the form of recast but that in instances of negotiating meaning learners had to first notice the non-target production and only after that could they modify their output. Similar results were obtained by Tsang

(2004) who analysed 18 lessons with learners aged 12-17 in a secondary school in Hong Kong. These studies provide evidence that feedback giving practices can support noticing the mismatch between the learners' interlanguage and the target language.

Doughty and Varela (1998) investigated the impact that implicit focus on form had on learning in a content-based programme with learners aged 11-14. The study included an experimental group, which received feedback through recast, and a control group. They observed that the experimental group outperformed the control group in their accuracy of past-time reference in spoken and written Science reports. The authors suggested that recasting could draw learners' attention to the linguistic forms and that it could support learning. However, it should be noted that the experimental group was taught by one of the researchers for the whole duration of the study (6-weeks) which may also have impacted on the differences in achievement between the two groups.

Although more research into the importance of noticing in FLL in a TEYL context is needed, the studies quoted above provide some evidence in support of the Noticing Hypothesis. In TEYL contexts, it is important to note that studies that evidenced the initial advantage of older beginners also highlighted the implicit nature of FLL in childhood. The studies quoted above provide conflicting evidence as to whether implicit feedback could lead to noticing and subsequently to L2 development. However, taking into account the context of the current study, instructed TEYL, where exposure is limited and hence not conducive to implicit learning, findings of research that support the applicability of the Noticing Hypothesis to YLLs become very attractive. They pose questions as to whether assessment methods could support noticing and hence contribute to FL development. It is hoped that by analysing the implementation of AfL in TEYL classrooms, the present study can provide some preliminary insights in this area and indicate paths for future inquiry.

In instructed TEYL contexts, learners' ability to 'notice the gap' may be of value but there does not seem to be enough evidence to indicate whether learners aged 7-11 might be able to do so. One area of research that can provide insights into YLLs' ability to notice what is being learnt or taught and how their performance relates to the target language is research on metacognition. Relevant studies in that area are reviewed in Section 2.2.1.3.2.

2.2.1.3.2 Metacognition

Metacognition is understood as 'reflection and evaluation of thinking that may result in making specific changes in how learning is managed, and in the strategies chosen for this purpose' (Anderson, 2005, p. 99). Studies in L2 contexts tend to adopt Flavell's (Flavell, Green & Flavell, 2000) model of metacognition (e.g. Gu, Hu & Zhang, 2005; Vendergrift, 2002), which proposes that metacognition comprises *metacognitive knowledge* and the *executive aspect* (also referred to as *control*). Metacognitive knowledge incorporates the knowledge of a person, task and strategy. Control refers to the use of strategies.

Research in the field of cognitive psychology has shown that young children, aged 3-5, begin to develop metacognitive awareness but are not able to use it effectively until they are approximately 8-10 years old (Flavell, Miller & Miller, 1993; Nisbet & Shucksmith, 1986). Cognitive psychology research has also indicated that training in metacognitive strategy monitoring enhances the effective use of strategies by children (Ghatala, 1986) and that strategy knowledge can facilitate strategy use (Borkowski, 1985). This suggests that in order to be able to use metacognitive strategies, young children may require metacognitive training.

The majority of research on metacognition in FL/L2 settings has been conducted with adult learners and has focused on exploring metacognitive awareness and perceptions of strategies, especially in reading (Zhang, 2001) and listening (Vandergrift & Tafaghodtari, 2010) and on the impact of strategy use on learning. Empirical studies in FL/L2 adult contexts indicate that metacognition is necessary for autonomous learning (Victori & Lockhart, 1995) and underlies the effective use of strategies (Wenden, 1987) so metacognitive strategy use should be included in the teaching (Goh, 1997). They also indicate that good learners make frequent and efficient use of strategies (Griffiths, 2003; O'Malley & Chamot, 1990) while poor language learners fail to select and implement strategies (Vann & Abraham, 1990); that the use of strategies is linked to language proficiency (Zhang, 2001); that monitoring learning by identifying problems and pausing to address them is beneficial to L2 learning (Rubin, 1975); and that activation of background knowledge is beneficial to reading comprehension (Zhang, 2001).

The current chapter does not aim to fully review the research into the use of learning strategies or metacognition in adult context. The focus here is on reviewing what is

known about metacognition in FLL by 7-11 year olds. Metacognition research in young learner contexts is limited. The studies that are available provide insights into types of metacognitive strategies used by YLLs and suggest that metacognitive training can be beneficial for the development of listening (Goh & Taib, 2006; Vandergrift, 2002), reading (Chamot & El-Dinary, 1999) as well as writing (Gu et al., 2005). These studies are summarised in Table 2.2 below.

Table 2.2: Empirical studies on metacognition in TEYL contexts

Authors	Age group	Focus of the study	Context
Vandergrift, 2002	9-12 year olds	Use of reflective activities focusing on prediction and evaluation in listening	French immersion in Canada; beginner level learners
Goh and Taib, 2006	11-12 year olds	Awareness of strategies for listening comprehension tasks	Primary school, Singapore
Gu, Hu and Zhang, 2005	7-9 year olds	Strategy use in reading, writing and listening	Primary school Singapore
Chamot and El-Dinary, 1999	5-12 year olds	Strategy use in reading and writing	French, Spanish and Japanese immersion programmes in the United States

Vandergrift (2002) studied the effect of introducing reflection following a listening comprehension activity in classes of 9-12 year old beginner learners of French in Canada. Following the completion of a reflective task, children were asked to respond to a questionnaire that focused on gathering children's perceptions of the formative quality of reflecting. The results suggest that the reflective tasks helped YLLs become more aware of the strategies involved in effective listening, in particular, predicting and evaluating completion of the task. But there was little evidence of self-knowledge involving strategies for motivating oneself. Vandergrift (ibid.) argues that reflection on listening tasks can facilitate the development of metacognitive strategies and that this could benefit listening comprehension. However, given the qualitative nature of the study design with no control group and the fact that Vendergrift (ibid.) did not include an evaluation of the regular teaching that the children had received, it does not seem possible to attribute the growing metacognitive awareness directly to the use of reflective tasks.

More accurately, it seems that the study provided evidence that the learners believed that the use of reflective tasks helped them develop the use of metacognitive strategies.

Another study that focused on children's awareness of metacognitive strategies was conducted in Singapore (Goh & Taib, 2006). The participants were ten learners aged 11-12. The study aimed to investigate metacognitive strategy use during listening tasks and the impact of raising learners' metacognitive awareness on their listening skills. The learners' listening comprehension skills were tested before and after an eight-lesson intervention programme. Each lesson in the programme followed a three-stage process (listen and answer/individual reflection/self-report and group discussion). The study provided insights into the metacognitive strategies used by the learners and the effectiveness of the three-stage process in teaching listening comprehension. Although the study did not include a control group to investigate the impact of the intervention programme on the development of listening comprehension, the researchers included a post intervention reflective writing task. The learners were asked to write about 'What I think about my listening ability at the end of the eighth lesson'. The results indicated that the children used inference (mostly based on key words in listening), planning (by reviewing questions prior to listening) and directing their attention to what they were listening to. No strategies for monitoring or evaluating the accuracy or completeness of their performance were reported. Furthermore, as in Vendergrift's (2002) account, the researchers noted that no affective strategies for motivating themselves or coping with anxiety were reported.

Some of the studies which focused on investigating language learning strategy use by young learners also provided some interesting insights into the development of metacognition. Gu et al. (2005) used think-aloud one-to-one interviews with children aged 7-9 in Singapore. The prompts included one listening task, one reading task and two writing tasks that were similar to the types of activities that the children routinely engaged in at their school. The findings indicated that the young children were not able to verbalise their strategies and needed probing questions from the researchers, which the team acknowledge might have impacted on the data. The results also suggested that the 9 year olds used a wider range of strategies than the younger children. This finding corroborates with research in cognitive psychology which suggests that although young children develop metacognitive awareness, they are not able to effectively use it until the ages of 8-10, as reported at the beginning of this section. Finally, the discussion of four

case studies of individual learners reported points to some difference in metacognitive strategy use depending on the language skills; more predicting was used in reading and listening and monitoring was more commonly used in writing. However, it should be noted that the one-to-one set up of the interview might have had a negative effect on the quality of the data collected by the research team, as the children might have experienced anxiety in such a context, which could have affected their responses.

A similar research method, a think-aloud protocol, was used by Chamot and El-Dinary (1999) to investigate strategy use by 5-12 year olds in French, Spanish and Japanese immersion programmes in Washington, DC. The results provide an interesting inventory of metacognitive strategies used by children. They included planning (previewing, reading aloud, making predictions); monitoring (checking meaning, self-correction, self-questioning, verifying predictions); and using selective attention (identifying key words, linguistic features, pronunciation and text features such as the title and the pictures). The findings indicate that children report using a greater number of metacognitive strategies in writing tasks than in reading tasks. The results also suggest that good learners are more effective at deploying and monitoring the use of strategies than poor learners. This finding is consistent with research conducted in adult settings (e.g. Griffiths, 2003).

All the above studies used research methods that relied on YLLs ability to describe strategy use orally or in a written form. The results of all the studies point to the conclusion that young learners find this type of reflection difficult. Hence, it seems that children at ages similar to the participants of the current study (7-11) may need support in using and talking about metacognitive strategies that allow them to plan, monitor and evaluate their own learning. Interestingly, although self-evaluation is an important metacognitive skill, TEYL studies have not yet explored that area (with the exception of Butler & Lee, 2006, 2010; see Section 2.3.3.2). It seems that the metacognitive strategies identified by the studies cited above point to the importance of learners noticing what they are learning as well as the gap between their own performance and the target language. Hence, it seems plausible to suggest that noticing is an important component of using metcognitive strategies and that both may have pedagogical merits. Therefore it would be of interest for the current study to investigate whether TEYL teachers' understanding of AfL in TEYL contexts points to the possibility of supporting the learners' development in metacognitive skills. It would also be of interest to investigate whether the implementation of AfL in TEYL classrooms actually requires YLLs to use

metacognitive skills. Furthermore, some of the studies reviewed in the current section point to differences in metacognitive strategy use depending on the language skills. It would be of interest to the current study to investigate if observations of how AfL is implemented in TEYL classrooms indicate similar differences.

Robinson (1995) suggested that pedagogical tasks require learners not only to notice their own mistakes and attempt to rectify them but also to rely on their memory. Similar considerations will be of importance to researching the type of assessment that is implemented alongside teaching and aims to support learning. Hence, a review of relevant studies that provide insights into memory in childhood is reported below.

2.2.1.3.3 Memory

Research in cognitive psychology recognises that *working memory* (WM) is related to the ability to comprehend and process language (e.g. Case, Kurland & Goldberg, 1982; Gathercole & Baddeley, 1993). In the field of foreign language teaching, Wen and Skehan (2011) and Wen (2012) argue that WM is an important component of language aptitude and that WM is the type of memory that is predictive of achievement in FLL. This section reviews what is known about memory, WM in particular, in childhood. But research into aptitude is not included since it is beyond the scope of this study.

Working memory is 'the cognitive capacity to simultaneously store and process information in real time' (Wen & Skehan, 2011, p. 21). The majority of research on WM has been conducted with adults as it relies on the use of complex tasks that require participants to process and store language simultaneously. However, because this research is beyond the scope of this study it is not included here. The focus is on reviewing insights provided by studies with children.

Case (1972) argued that physiological resources needed to process new information effectively develop slowly in 3-8 year olds and faster from the age of 8 onwards. In another study, Case and colleagues studied the use of WM by children aged 2-6 (Case et al., 1982). The children were asked to count the number of objects in a series of pictures and subsequently to recall those numbers in correct order. The researchers observed that the older children were able to recall more numbers and did so faster. The authors argued that processing demands diminish with age, and hence release storage space. The research on WM has since indicated that when the demands of a task are high, there can

be a trade-off between storage and processing; when a lot of storage is required, less processing can occur (Daneman & Carpenter, 1980; Daneman & Merikle, 1996). A more recent study reported the performance on complex listening tasks that required the participants to process syntax and recall linguistic information simultaneously (Siegler, 1994). The results of this large scale study, conducted with over 1200 participants aged 6-49, indicated that learners develop skills in the effective use of WM between the ages of 6-15. These studies suggest that processing in young learners' WM may not be as effective as in the case of adolescents and adults and so there may be little WM space available for storage. This interpretation corroborates the findings of research indicating that children have limited phonological and visual memory spans which increase with age and that adult like capacities are reached at around the age of 10-12 (Wilson, Scott & Power, 1987). This has pedagogical implications for the amount of processing that can be expected from small children. Of importance for the current study is that research on WM in childhood suggests that tasks used for assessment should not require young learners to process large amounts of information.

In the field of foreign and second language learning, Skehan (1996, 1998) proposed the dual processing system theory. It maintains that two systems co-exist in the brain; the memory-based system relies on the ability to recall whole chunks of information, e.g. formulaic phrases, while the rule-based system enables learners to think analytically. Skehan (ibid.) seems to use the term memory-based system (also referred to as the exemplar-based system) to refer to what cognitive psychology research calls storage; whereas the rule-based system seems to refer to processing as discussed in the previous paragraph. The research reviewed so far has indicated that, as they mature, humans develop effective processing i.e. processing requires less working memory and can operate more quickly. This suggests that children rely on the exemplar-based system until the rule-based system develops. As effective processing improves with age, children's reliance on the exemplar-based system decreases. As a result, more storage space (i.e. memory capacity) becomes available.

The dual-mode perspective for language processing (Skehan, 1996) proposes that in second and foreign language contexts learners can move between two possible modes of communication, depending on contextual circumstances. Skehan (ibid.) reports that, if under time pressure, adult learners draw on the rapid exemplar-based system. This enables them to communicate in the lexical mode. If creativity or exactness is important,

learners deploy the analytical, rule-based system. This results in learners focusing attention on language form. It seems plausible to infer that if YLLs rely on the exemplar-based system, they may tend to communicate in the lexical mode, paying less, if any, attention to the form of the language. This could explain the findings of previously reviewed studies indicating that children do not pay attention to form but to meaning (Bialystok, 2001; Gacia-Mayo & Garcia-Lancumberi, 2003). These issues, most importantly, highlight the lexical communication mode is most readily accessible to young learners. This has implications for language assessment. Most importantly, it indicates that the priority in teaching and assessing language in childhood should be on meaning, with form focused instruction being introduced gradually, when learners are developmentally more prepared for it.

The research on committing information to memory provides insights into the development of memory strategies in childhood. Berk (2000) identified organisation, rehearsal and elaboration as examples of short-term memory strategies and recall, reconstruction and recognition as examples of long-term memory strategies. Wood (1998) argued that children can be taught memory strategies but may take some time to master them during the initial years of schooling. Flavell, Beach and Chinsky (1966) observed children who were trying to remember a list of words presented to them as pictures. They concluded that no rehearsal strategies could be observed in children below the age of 7. In a later study, Flavell (1992) suggested that rehearsal strategies can develop between ages 5-10. Smith, Cowie and Blades (1998S) found that children older than 10 developed organisational strategies. According to Schneider and Pressley (2013), the final strategy to develop is elaboration, which rarely occurs before the age of 11. These findings suggest that children's ability to use memory strategies should not be assumed in implementing teaching and assessment practices.

The studies reviewed so far indicate that children rely on storage more than processing and that memory strategies develop during childhood. Importantly, in the field of foreign language learning, the gradual development of the effectiveness of the rule-based system may have implications for the types of communication that children are capable of. This is a very important issue as it has implications for research as well as teaching and assessment. Firstly, research needs to consider how young children use language to communicate with others as this can provide useful insights into the understanding of the dual-mode processing theory in TEYL contexts. This seems to be an especially

interesting research focus given the evidence from SLA research that interactions can contribute positively to language learning (see Section 2.2.2.3). Secondly, teachers and assessors need to be aware of the development of memory capacity and memory strategies during childhood in order to be able to effectively design teaching and assessment methods and accurately interpret assessment outcomes.

The studies reviewed in this section sit within a constructivist tradition, where learning is constructed through the interaction between experience and their own ideas. Significantly, this understanding considers learning to be a largely individual process but it does not account for the social context in which learning happens. This latter perspective is explored in the next section.

2.2.2 The socio-cultural perspective on cognitive development and FLL

2.2.2.1 Introduction

This section reviews research in the socio-cultural tradition, which conceives of learning as socially situated and emphasises the pivotal role of interaction in learning. It begins by exploring the nature of learning that happens though social interaction, as conceptualised by Vygotsky (1987), in Section 2.2.2.2. Section 2.2.2.3 follows with a review of empirical studies investigating the role of interactions in FLL, with a particular focus on childhood and the implications of the findings for assessment in TEYL contexts.

2.2.2.2 Zone of Proximal Development

Social interaction and its impact on learning were studied by the Russian psychologist, Lev Vygotsky (1896-1934). Central to Vygotsky's theory of learning is the notion of *internalisation* i.e. how processes and actions at the *intermental level* (external to the mind) are transferred to the *intramental level* (inside the mind) through *mediation* which uses 'demonstration, leading questions, and by introducing the initial elements of the task's solution' (Vygotsky, 1987, p. 209). Mediation happens during social interaction, often through the use of language between the learner and a *more capable peer* (i.e. a teacher or any other person who can perform the task with a greater degree of expertise than the learner). The distance to which the new skills can extend from the intramental sphere marks the potential capability for learning within the context of the task at hand and was termed by Vygotsky the *zone of proximal development* (ZPD).

The importance of interaction and communication in the learning process is highlighted by the proposition of the *intermental development zone* (IDZ) proposed by Mercer (2000). Mercer (ibid.) argues that while operating within a learner's ZPD, the learner and a teacher can mutually create an IDZ through communication. The IDZ is closely related to the learning aims and it aids the completion of the task at hand. Unlike Piagetan theory, where cognitive development is a pre-requisite for language learning, in Vygotskian understanding, language itself is an integral tool of cognitive development and mediates it. This concept has implications for language classrooms. Most notably, it claims that teaching should enable interaction, offering opportunities to interact with more capable peers.

From this theoretical perspective, assessment practices need to provide information about what the learner can do independently as well as provide insights into what (s)he is capable of doing with support. Additionally, the assessor needs to be able to determine what learners cannot complete even when support is available. That information would enable the teacher to delineate the learner's ZPD. Within the ZPD, learning could be advanced through interaction during which a more capable peer supports a learner in creating an IDZ, thus working towards internalising new skills. Hence, it seems useful to research interactions that happen during lessons and assessment episodes to investigate if and how they contribute to enhancing learning.

2.2.2.3 Interactions in FLL

The role of teacher-learner and learner-learner interactions in FLL has been extensively researched in adult contexts. Many studies have indicated that interaction is an important vehicle of learning, i.e. that FLL happens through interaction (e.g. Gass, 2013; Mackey, 1999; Pica, 1994). The argument is based on the Comprehensible Input Hypothesis, which suggests that FLL is facilitated by the input that is understood by the learners (Krashen, 1985). Of particular interest to the current study is a related hypothesis proposed by Long (1981), which suggests that input could be made comprehensible to learners by their interlocutors who modify what they say when learners have indicated a lack of understanding. This is referred to as the *interactional modification of input*. In other words, the meaning is negotiated in conversation by the interlocutors, who 'express and clarify their intentions, thoughts, opinions etc., in a way which permits them to arrive

at a mutual understanding' (Lightbown & Spada 2008, p. 122): hence the expression *the negotiation of meaning*.

Early studies on the negotiation of meaning focused on investigating whether such negotiation led to better comprehensibility. But conclusive evidence that the comprehensibility does lead to FLL could not be found until 1994, when Ellis, Tanaka and Yamazaki reported two studies with 79 and 127 high school students learning English in Japan. Both studies were based on the same design. The researchers divided the students in two groups; one received pre-modified input and the other interactional modification. Through implementing a pre- and post-test design, the researchers found that the learners' knowledge of new vocabulary items was better in the interactional modification group than in the group that had received pre-modified input, regardless of whether the learners participated in the interactional modification or whether they simply observed it. The studies by Ellis et al. (ibid.) were important as they were the first to provide empirical evidence that interactional modification can lead to FLL. However, they referred only to learning the meaning of concrete nouns. As various aspects of language may be learnt differently, more evidence of the impact of comprehensibility on FLL was needed. The studies by Ellis et al. (ibid.) were conducted with older children; hence the results may not be directly applicable to TEYL classrooms. However, they raise an important question of whether learners younger than those in Ellis et al. (ibid.) studies can negotiate meaning.

Research in adult contexts has demonstrated that opportunities for interlocutors to negotiate meaning can facilitate conditions and processes that are considered significant in learning a second language (Pica, 1994). They can have positive effects on second language production and comprehension (Gass & Varonis, 1994) and in the process of negotiating meaning learners' attention may be on language form (Mackey, 1999) as well as on meaning (Gass, 2013). Similar studies based in TEYL contexts are rare. However, the evidence accumulated in TEYL research suggests that interacting with peers could benefit FLL (Pinter, 2007; Swain, 2000) and that children can and do negotiate for meaning (Oliver, 1998, 2000, 2002; Oliver & Mackey, 2003) although they may not be able to take up their interlocutor's perspective (Butler & Zang, 2014).

Pinter (2007) investigated whether learners benefit from learner-learner (L-L) interaction. Two 10 year old Hungarian learners of English at low levels of FL

proficiency were asked to complete spot-the-difference tasks on three occasions, over three weeks. Each time the same task design was used but the pictures were different. Subsequently, the YLs were asked to watch their video recorded conversations and to comment on their performance. Pinter (ibid.) argues that by implementing the same task design on all three occasions, the study created conditions in which the learners did not have to focus on understanding the requirements of the task but instead were able to use collaboration skills, provide mutual scaffolding and become sensitive to each other's linguistic needs. This interpretation suggests that familiarity with the task type may be important in supporting better language performance.

Additional evidence on how familiarity with the structure of a task can impact on language learning comes from studies conducted with adults (Bygate, 1996; Skehan & Foster, 1999). Bygate (1996) asked 11 participants to retell a video story that they had watched and then asked them to do the same 10 weeks later. He observed a greater complexity of oral production when the task was repeated and concluded that the effects of task repetition may have implications for teaching and learning as well as for assessment. In a slightly different study, Skehan and Foster (1999) explored how the amount of structure in a task impacted on oral production. They assumed that the participants would have general knowledge of a typical conversation in a restaurant and, as a result, judged the retelling of a story from a restaurant scene, sourced from a movie about Mr. Bean, as a more structured task. Another scene (playing golf in an unusual manner) was considered less structured. The results of analysing the retelling of the two stories indicated that more structured activities resulted in greater fluency, but not accuracy or complexity. These two studies provide interesting insights into the significance of familiarity with task type that Pinter's (2007) study indicated as beneficial for oral production in a TEYL context. Nonetheless, more evidence is needed from TEYL contexts to ensure that the validity of claims that familiarity with a task can benefit production.

However, regarding Pinter's (2007) claims that 10 year olds were able to collaborate and that they demonstrated mutuality in interactions, a different study (Butler & Zeng, 2014) demonstrated that 9-10 year olds found it challenging to take their partner's perspective or to mutually develop a topic with them. Instead, the learners relied on formulaic turn taking. The study aimed to investigate holistic interaction patterns and conversation characteristics between children aged 9-10 and 11-12. The authors applied Stroch's

(2002) model to analyse 32 dyadic L-L interactions, recorded during a task-based language assessment procedure. They observed that collaboration occurred more frequently between the older learners (11-12). Butler and Zeng's (ibid.) study seems to offer more detailed insights by reporting a larger number of interactions and by accounting for the age factor. Both studies (Pinter, 2007; Butler & Zeng, 2014) offer interesting insights that have implications for classroom practice and assessment as they indicate that younger learners may be less able to collaborate during language learning or assessment tasks than their older counterparts. To become more competent, they may need support from the teacher or the task design. Pinter (ibid.) suggests that by implementing a similar task type repeatedly, teachers could encourage children to perform tasks with a greater degree of collaboration. Butler and Zeng (ibid.) indicated that this ability develops with age. It would be useful to investigate if the AfL practices implemented in TEYL classes could support the occurrence of collaborative and expert/novice interactions.

Butler and Zeng's (ibid.) study also offers a useful methodological solution for investigating interactions in TEYL classrooms by applying Storch's (2002) model, which allows for classifying classroom interactions into one of four holistic patterns (see Section 3.3.3.2.2), two of which (collaboration and expert/novice) have been shown to be beneficial for FLL. The review will now focus on this area of research.

Swain (2000) has drawn attention to the advantages of collaborative dialogues. She analysed a dialogue between two 13 year old L2 (French) learners from a socio-cultural theory of mind perspective, arguing that the necessity to vocalise helped the learners to focus each other's attention on what they did not know, or were unsure of and that this kind of social interaction enabled them to generate new linguistic knowledge. This is supported by Swain and Lapkin (1998), who analysed the collaborative dialogues of two 13-14 year old early immersion students in Canada, who were completing a jigsaw puzzle. They concluded that learners could co-construct the language that they needed when they collaborated during the task and that they did so in *language-related episodes* (LREs).

LREs are situations in which learners talk about the language they are producing in speaking or writing. The majority of research that has looked into the occurrence and the beneficial role of LREs has come from adult contexts. These studies have demonstrated

that expert/novice (e.g. an advanced student with an intermediate student) and collaborative (e.g. similar proficiency levels) interactions can create conditions that are conducive to learning (Kim & McDonough, 2008; Ohta, 1995; van Lier, 2014; Watanabe & Swain, 2007; Williams, 2001).

For example, Williams (2001) analysed 65 hours of classroom collaborative interactions between learners at varying proficiency levels and concluded that LREs tended to be lexically oriented and that the number of LREs increased as the learners' proficiency level increased. However, no such difference between the number of lexical and grammatical LREs was reported by Kim and McDonough (2008). They investigated dyadic interactions between intermediate level students with other intermediate level students and advanced students in South Korea. They applied Stroch's (2002) model to interactions that took place during a dictogloss activity and concluded that a higher number of LREs occurred when learners interacted with more advanced interlocutors and that the LREs were evenly distributed between lexical and grammatical. They attributed that second finding to the task type (dictogloss). Another study that applied Storch's (2002) model was conducted by Watanabe and Swain (2007) with university learners who engaged in a writing task followed by a stimulated recall interview. 'Stimulated recall is a type of introspective method in which prompts such as videotaped interaction of themselves are used to stimulate the learners' recall of their thoughts at the time the activities originally took place' (Watanabe & Swain, 2007, p.127). The authors concluded that, regardless of proficiency level, collaborative interactions resulted in a high number of LREs. Ohta (1995) analysed the collaborative dialogue of students with minimally differing levels of proficiency in English and also argued that this type of interaction created a positive learning environment; and van Lier (2014) claimed that more advanced students can benefit from collaboration with less proficient learners by teaching them.

The body of research based on analysing interaction patterns demonstrates that collaboration and expert/novice patterns can be beneficial to learning a FL. However, the majority of that research, with the exception of Butler and Zheng (2014) and Swain and Lapking (1998), comes from adult contexts. The other studies conducted in TEYL classrooms have focused on another important aspect of interactions introduced earlier, viz. the negotiation for meaning.

A series of studies in Australia investigated L-L interactions between 8-13 year olds (Oliver, 1998, 2000, 2002). The participants included non-native speakers (NNS) and native speakers (NS) of English. Dyadic interactions were recorded and analysed to gain insights into how children negotiate for meaning. The results indicated that young children can and do negotiate for meaning. The amount of negotiation between children was connected with their language proficiency and their nativeness. Specifically, the children tended to negotiate more in NNS-NNS dyads at low levels of English. Interestingly, in that age group, Oliver (2002) did not report the impact of age or gender on the ability to negotiate and the amount of negotiation.

In a different study, Mackey, Oliver and Leeman (2003) analysed *negative feedback* that was offered by child (8-12 year olds) and adult NS and NNS interlocutors in dyadic interactions. Negative feedback is feedback 'provided in response to learners' non-target-like production' (Oliver & Mackey, 2003, p. 519). Their findings indicated that NNS adults offered the least amount of feedback but no differences in the amount of negative feedback were observed in children with regard to whether they were NSs or NNSs. Mackey et al. (2003) argued that, depending on the age and nativeness of interlocutors, learners encounter different linguistic environments with regard to the amount and quality of negative feedback. Commenting on the important issue of using such feedback to modify output, Mackey et al. (ibid) also noted that the age and nativeness of interlocutors were important factors. However, they indicated that further research was needed to establish the relationships between negative feedback and modification of output. Their study is relevant to the context of the current review as it indicated that, during interactions, YLLs did not only have opportunities for requesting modification of input, but also, importantly, such modification entailed modification of output by their interlocutor. Hence, modification of output could occur as a result of negative feedback. Swain (2000) argued that 'the importance of output for learning could be that output pushes learners to process language more deeply – with more mental effort – than does input' (Swain, 2000, p. 99). Hence, it seems important to consider how negative feedback received through AfL may result in modifications being made to input as well as output (5.4.2.2).

Other studies that have explored classroom interactions have investigated teacher-learner (T-L) exchanges. For example, Ellis and Heimbach (1997) explored the effects of negotiating meaning on the acquisition of lexis in kindergarten FL (English) learners in

Japan. The design used listening tasks and recorded interactions between the children. Their findings indicated that the largest number of interactions in these TEYL classrooms happened between teachers and learners and not among learners.

In an age group similar to the cohort in this study, T-L exchanges between experienced teachers and their learners aged 6-12 in an immersion context were investigated by Oliver and Mackey (2003). The focus of their study was on investigating the feedback that the learners received from their teachers during lessons. The researchers classified the interactions into four categories: content, communication, management and explicit language focus. The results suggested that the teachers were more likely to provide feedback on explicit language and content. However, the learners tended to use the feedback focusing on language more often than that on content. These results suggest that interactions occur in various contexts within the lessons (according to the four different categories) and that the nature of the contexts is an important consideration as to whether the interactions offer conditions for moving learning forward. Hence the analysis of interactions needs to consider the contexts carefully. This has methodological implications for analysing interactions in TEYL classrooms and has been accounted for in the design of the current study. However, in this thesis the contexts are referred to as modes of interactions following Walsh (2006) (see Section 3.3.3.2).

2.2.3 Affective factors

Individual Differences (IDs) of a cognitive nature were discussed in Section 2.2.1. Most recent approaches to considering IDs also include affective factors. Schumann (2001) argues that, as children build up a bank of life experiences, they develop an appraisal system, in which they incorporate their FLL. Each learner has different experiences. Hence the resulting appraisal systems are unique; YLLs bring different levels of motivation and anxiety, and various attitudes and self-concepts to the classroom. This section reviews studies that provide insights into FLL in childhood from the perspective of affective factors and considers their significance for language assessment with young learners.

2.2.3.1 Anxiety

Foreign language anxiety is understood as 'the feeling of tension and apprehension specifically associated with second language contexts, including speaking, listening, and learning' (MacIntyre & Gardner, 1994, p. 284). The notion of *test anxiety*, defined as

'the tendency to become alarmed about the consequences of inadequate performance on a test or other evaluation' (Sarason, 1984), is not important to this study, because, although the focus is on assessment, the assessment techniques that are included in AfL do not comprise testing. Instead, they are task based and incorporated in the teaching. Therefore, it is the insights offered by research into anxiety connected with language teaching and learning that are more relevant here than those of test anxiety research.

There seems to be little research into FL anxiety in TEYL contexts. Nikolov and Mihaljević Djigunović (2011) have suggested that this is because of a popular belief that children are free from anxiety but they note that more recent research has explored this investigation-worthy area. Available insights have come mostly from adult classes, with some studies in adolescent classes also available. Overall, this body of research indicates that there is a relationship between anxiety and a willingness to communicate (MacIntyre, Baker, Clément & Donovan, 2002) and listening comprehension (Mihaljević Djigunović & Legac, 2009). Additionally, research indicates that bilingual learners tend to experience less FL anxiety than others (Legac, 2007).

MacIntyre et al. (2002) investigated the effects of anxiety on communication, alongside other factors, in classes of 12-15 year old learners in a French immersion programme in Canada. They concluded that the level of anxiety was stable across different ages; that L2 anxiety was significantly higher than L1 anxiety and that it negatively correlated with a willingness to communicate (WTC). Importantly, this study pointed out that anxiety was one of many factors that affected interaction in the classroom. They include sex, age, L2 WTC, perceived competence and motivation.

In a more recent study, Mihaljević Djigunović and Legac (2009) compared two groups of learners aged 13-14: 56 monolinguals and 56 bilinguals (Croatian-Albanian, Croatian-Czech, Croatian-Italian) learning English. They explored anxiety levels in three stages of FLL (input-processing-output) in listening achievement. The findings indicated that the bilinguals experienced lower anxiety in all three stages; that they performed better in listening comprehension tasks; that, in both subgroups, listening was negatively correlated with FL anxiety in all three stages; and that all correlations were statistically significant. The authors argued that the linguistic experience of bilinguals facilitated their development of higher self-confidence in FLL, thus lowering anxiety. The results of the analysis using data from the same cohort of students, reported by Legac (2007),

confirmed that bilinguals experienced lower levels of both listening and overall language anxiety.

Oxford (1999) reviewed literature on language anxiety and pointed out that negative relationships have been shown to exist between language anxiety and performance in speaking and writing tasks, grades in language courses, performance in tests as well as in self-confidence and self-esteem. On the other hand, Oxford (ibid.) also noted that a few studies suggest that, in some cases, anxiety can have a facilitating effect, by keeping learners alert. But neither of the studies quoted above provided insights into the causal direction of the relationship between anxiety and language learning. Hence, it is unknown whether a low level of achievement in listening, for example, is caused by FL anxiety or if it is a reason for it.

The research reviewed in this section, indicates that anxiety correlates negatively with language performance, although more evidence for that is still needed from TEYL contexts. The review also suggests that evaluative situations may cause anxiety. Hence, it would be interesting to gain preliminary insights into the possible relationship between implementing AfL and YLLs' FL anxiety.

2.2.3.2 Motivation, attitude and self-concept

This section continues the discussion of affect in the TEYL context by reviewing what is known about *motivation*, *attitude* and *self-concept* in TEYL contexts. The implications for language assessment are also discussed. According to Gardner (2010) motivation to learn a language includes the desire to do so, combined with positive attitudes and with effort devoted to language learning. Gardner and MacIntyre (1993) define attitude as a positive or a negative feeling concerning foreign language learning and what the learner may associate with that language. Attitude and motivation are often considered as interlinked (Mihaljević Djigunović, 2006). As with attitude, self-concept can be positive or negative and refers to 'a person's notion of himself as a FL learner' (Laine, 188, p. 10). According to Laine (ibid.) learners' self-concepts can be considered at three levels: general (related to any FL), specific (related to a particular FL) and task (related to particular language tasks or skills) (p. 20). Three aspects of self-concept have been discussed by other researchers (e.g. Mihaljević Djigunović, 2006, p. 11). These include: the ideal self (the notion of a learner's aspirations), the actual self (the learner's

perception of him/herself) and the social self (the learner's perception of how others see him/her).

Some research into affective factors has suggested that children can adopt the attitudes and motivation of the people who are important to them, such as teachers or parents (Szpotowicz, Mihaljević Djigunović & Enever, 2009) and that, with time, children's motivation may decline (e.g. Mihaljević Djigunović & Lopriore, 2011) in line with their motivation in other school subjects (Matsuzaki Carreira, 2006). However, Cenoz (2003) suggests that a high level of motivation can be sustained but it requires good conditions, some of which may include experiencing success in learning (Cable et al., 2010) or contact with native speakers (Marsholleck, 2002, cited in Mihaljević Djigunović, 2015). Also the starting age of instruction (Kennedy, Nelson, Odell & Austin, 2000; Nikolov, 1999) or increasing the length of FLL (Donato, Tucker, Wudthayagorn & Igarashi, 2000) may contribute to enhancing attitude and motivation. According to Nikolov (1999) a more context sensitive and situated approach to researching affective dispositions indicates that, as they accumulate experience of FLL, children can develop their own attitudes and, according to Mihaljević Djigunović (2015), affective processes follow individual trajectories even in situations where learners experience similar contextual factors. What is of interest here is those studies that have investigated conditions that may support the development of positive affective processes and their relationship to achievement in TEYL contexts (Enever, 2011; Masgoret, Bernaus & Gardner, 2001; Matsuzaki Carreira, 2006; Mihaljević Djigunović, 2006, 2015; Mihaljević Djigunović & Lopriore, 2011; Vilke & Vrhovac, 1995, cited in Mihaljević Djigunović, Nikolov & Ottó, 2008). These are reviewed in more detail below.

Research has indicated that YLLs tend to start FLL with high levels of motivation. A recent longitudinal study, the Early Language Learning in Europe (ELLiE) project (Enever, 2011), which was set up to measure what can realistically be achieved in TEYL contexts in state school settings, has provided interesting insights into motivation, attitudes and self-concept. The study included 1400 children in seven European countries. Data were collected through interviews with principals and teachers, lesson observations, parent questionnaires, children questionnaires, interviews and achievement measures. The analysis of data obtained through smiley face children's questionnaires indicated that in the first year, the majority of learners expressed positive attitudes towards FLL with a quarter of the learners declaring neutral attitudes. After three years,

a significant number of learners remained positive with fewer providing neutral responses. This suggests that the learners expressed more informed opinions. This was interpreted as indicating that changes in attitude can be influenced by the experience of FLL (Mihaljević Djigunović & Lopriore, 2011). These findings corroborate with Cenoz's (2003) suggestion that positive attitudes can be maintained in TEYL contexts given favourable conditions.

Some of the insights reported by the ELLiE team (Enever, 2011) were related to the relationship between affective dispositions and achievement. The data gathered in that study indicated that young learners with a positive self-concept, motivation and attitude were shown to perform better on listening and oral production tests (Mihaljević Djigunović & Lopriore, 2011). Moreover, it was shown that affective factors had a stronger impact on achievement at the age 10/11 rather than in the initial phase of learning at the age 7/8. However, the study did not attribute the impact of individual characteristics on achievement directly to the learners' age as it did not incorporate a control group, for instance, of children who started learning when aged 10/11. Given these insights, it seems plausible to think that assessment providing YLs with information about their FL achievement could play a role in how YLs motivation, attitudes and self-concept change.

A study conducted with 8-9 and 11-12 year olds in Japan, provided insights into change in intrinsic and extrinsic motivation in FLL (Matsuzaki Carreira, 2006). It revealed that motivation declined with time. The author concluded that this was in line with the decrease in motivation towards other school subjects in Japanese primary school students. These results support the findings of the ELLiE study. Both studies suggest that as learners gain experience in FLL (and perhaps across the curriculum), their attitudes and motivation may change.

In another study investigating learners' individual differences (IDs) in motivation, attitudes and self-concept in Croatia, Mihaljević Djigunović (2015) used a mixed-method approach adopting the instruments developed in the ELLiE project. The study was conducted with 284 learners subdivided into younger beginners (age of 6/7 years old) and older beginners (9/10 years old). The aim was to investigate the relationships between the IDs and age, language proficiency and the trajectories of change over time (years 2-4 of studying English). The quantitative results suggest that younger and older

beginners differ in attitude: the younger learners preferring more traditional classrooms as opposed to the older learners preferring the group work arrangement. Self-concept was found to be more positive in the older year group in year 3 and in the younger year group in year 4; and the younger learners were more motivated but both groups experienced a decline in motivation over time. With regards to attitude, Mihaljević Djigunović (ibid.) suggested that the younger learners preferred more teacher controlled environments, while their older counterparts became more aware of the benefits of group work and developed appropriate skills for participating in this type of work. Moreover, the author discussed how more explicit learning outcomes in the group tasks set for the older group could have contributed to providing a more purposeful and motivating environment. Thus, motivating the learners to do their best gives another perspective on how the structure of a task can contribute to learners' performance.

From the qualitative analysis of six longitudinal case studies, Mihaljević Djigunović (2015) arrived at some interesting findings. These studies demonstrated that the trajectories of change in affective processes were idiosyncratic. The author concluded that there is a need for adopting a situated, contextualised approach to researching motivation, attitudes and self-concept. There was also some evidence in four out of the six case studies, but not discussed by the author, that there seemed to be a relationship between the learners' understanding of assessment results and the changes that occurred in their self-concepts and/or motivation. Mihaljević Djigunović (2015) reported that the self-concept of one learner, Zlata, increased when she was awarded grade A; another learner, Maja, believed that she was as good as other learners in the class because she and the majority of the children were awarded grade B; another learner, Stjepan, reported that he was worse than others in English because he had a lower grade; and another learner, Zvonimir, reported that testing impacted negatively on his and on others' motivation towards learning English. These useful insights highlight an area that warrants further research: viz. an investigation into the relationship between assessment practices and changes in affective processes in TEYL classrooms. Although, this is beyond the scope of the current study, hopefully, by analysing data from lesson observations and teacher interviews, the present study can indicate useful paths for future inquiry into the possible relationships between affective IDs (motivation, attitudes and self-concept) and assessment practices in TEYL classrooms.

YLLs' perception of self-achievement has been shown to be related to motivation in another study (Masgoret, et al., 2001). Masgoret et al. (ibid) adapted Gardner's (1985) Attitudes/Motivation Test Battery (AMTB) for use with 10-15 year old Spanish (L1) speakers learning English. The study demonstrated that children who perceive their own achievement in English as good are those who have a positive attitude towards learning English and communicating with English native speakers. Also, Mihaljević Djigunović (2012) found that the relationship between motivation and language achievement 'depended on what kind of measure of achievement was used: stronger correlations were found with self-assessment, course grades and integrative tests than with discrete-point tests.' (p. 161). The studies discussed in this paragraph suggest that the ways in which learners are assessed and the ways in which YLLs perceive their own achievement, perhaps as a consequence of assessment and feedback giving practices, may have implications for the development of their affective dispositions.

Positive affect was also shown to play an important role in learners developing productive language skills (speaking and writing). Mihaljević Djigunović (2006) reported the findings of a large scale study called English in Croatia conducted in 2003 with over 2000 learners in their final years of primary (aged 13/14) and secondary (aged 17/18) education. Positive affect was shown to be correlated with performance in speaking and writing, and to be stronger in the younger age group. Importantly, this study suggests that language achievement is not a general construct. Instead it might be worth considering the relationship between affective dispositions and different language skills separately. Mihaljević Djigunović (ibid.) concludes that '(i)f we interpret the relationships evidenced by the significant correlation coefficients in terms of affect as a cause of success, the teaching implications of these findings are quite apparent: we should help FL learners to create and maintain a positive affective profile' (p. 20). This argument has important implications for the current study as it poses a question of the role of assessment in creating and maintaining a positive affective profile.

Another study that explored the relationship between achievement, motivation and attitude was reported by Vilke and Vrhovac (1995, cited in Mihaljević Djigunović et al., 2008), who suggest that intensive teaching at the beginning of the programme could offer opportunities to feel successful and that such a feeling is vital for motivating future learning. Cable, et al. (2010) and Mihaljević Djigunović and Lopriore (2011) found in their studies that vocabulary learning, specifically, is important for primary aged children

as it can be motivating. These studies all suggest that feelings of success could foster positive motivation.

The studies reviewed in this section indicate that anxiety, motivation, attitude and self-concept are important factors that impact on the learning in TEYL classrooms. The implications for assessment emerging here signal that assessment practices should offer opportunities to foster positive affect as this can support learning. Moreover, an extensive review of TEYL research (Edelenbos, Johnstone & Kubanek, 2006) suggests that the development of affective factors such as motivation and positive attitude may be some of the most valuable outcomes of TEYL programmes. More recently, Mihaljević Djigunović (2015) has commented that this is now a 'widely-acknowledged aim of early FLL' (p. 17). This has direct implications for the current study; if motivation is considered to be an expected outcome, then it becomes a part of the construct of assessment. If so, it would be appropriate for assessment practices used in TEYL contexts to assess and encourage motivation. This suggests that it would be both interesting and important to gain preliminary insights into this area by collecting evidence, from observed lessons and teacher interviews, about the purposes of using AfL. Hopefully, this would indicate whether AfL can contribute to measuring and/or promoting the development of positive dispositions towards FLL in YLLs.

2.2.4 Summary of Section 2.2

This section (2.2) so far has discussed how children differ from adults in terms of cognitive development and FLL processes. It has also explored the importance of affective factors in TEYL classrooms. The aim of the review is to tease out factors that are important to consider in implementing assessment in TEYL contexts. Working on McKay's (2006) proposition that assessment practices in TEYL classrooms should account for how children learn and are taught a FL, this section now examines the findings of the review so far to identify features that the language assessment of YLs should incorporate.

The review in Section 2.2.1 demonstrated that young children initially learn more slowly than older beginners. This is attributed to the implicit nature of FLL in childhood, which requires more time and to the developments in literacy in older learners. The nature of FLL in childhood is further informed by the literature on how children process language. Importantly, the review has suggested that learning in childhood relies on memory and

that attention, noticing and analytical skills develop as learners mature. Furthermore, metacognitive skills, which are vital for conducting self- or peer- assessment, may not be possessed by YLLs but could be taught. These considerations have important implications for assessment in TEYL classrooms. Firstly, it seems necessary that assessment is not delayed in time but happens alongside the teaching and learning, thus being contextualised through on-going classroom practice. Secondly, it seems that the development of metacognitive awareness and strategies that help children notice the gap between their performance and what is expected from them should be incorporated gradually into teaching and assessment in order to enable learners to understand feedback and to self- and peer- assess. Thirdly, assessment practices should also take into account children's short sustained attention span and provide tangible evidence of achievement that they can understand. Fourthly, children's working memory relies to a large extent on the exemplar-based system, resulting in lexical communication being more available to young children. Hence, teaching and assessing the form of language explicitly may not be appropriate in TEYL contexts. Finally, the level of literacy in L1 should also be considered as it may have a direct impact on the choice of assessment tools that can be used in TEYL contexts and on the rate of progress.

The review in Section 2.2.2 suggested that teaching and assessment should enable interactions with a more capable peer in order to provide information about the current level of skill with relation to the task at hand, what can be done with support and what is beyond the learners' ZPD: thus providing formative information about the next steps in the teaching and learning. However, only a limited number of studies in TEYL contexts have explored interactions that occur during assessment and how these can impact on FLL. Given the important role that interactions were shown to play in FLL in adult classrooms, it could be argued that, by facilitating conditions for collaborative dialogues, assessment practices can facilitate learning. Additionally, the review indicated that when learners are familiar with the task type they are able to devote more attention to completing it and collaborating with their interlocutors. These findings seem to have two implications for assessment in TEYL contexts. Firstly, they indicated that assessment aiming to move learning forward should be socially situated. Secondly, by implementing a task that is familiar for the learners, teachers can facilitate the collection of reliable assessment information. Additionally, the review concerning interaction offers important methodological considerations. Firstly, Storch's (2002) model was identified as a useful

tool for analysing the holistic patterns of dyadic interactions in a TEYL context. Secondly, the review indicated that the modes in which interactions occur need careful attention because some modes were shown to offer opportunities for the effective use of feedback by students (Oliver & Mackey, 2003).

The review in Section 2.2.3 explored the importance of affective factors in TEYL contexts. Anxiety, motivation, attitude and self-concept were shown to interact. More importantly, however, the review indicated that achievement and the perception of self-achievement might be important in sustaining learners' motivation to learn and a positive attitude. These, in turn, are closely linked to developing a positive self-concept. Crucially, positive affect was shown to be correlated with performance in productive skills. This has clear implications for the current study. Most significantly, by providing feedback any assessment method seems likely to contribute to learners' building up a perception of their own achievement. Hence, it is important that such feedback provides learners with positive reinforcement as well as with constructive criticism. Additionally, where areas for development are identified through feedback, these need to be communicated in a measurable and achievable way that can be understood by children and that enables YLLs to enhance their achievement: thus contributing to building a positive self-concept and motivation.

Having explored the issues connected with language learning in childhood, the review continues by shifting attention to reviewing the assessment of YLLs.

2.3 Assessment in Teaching English to Young Learners

2.3.1 Introduction

The aim of this section is to discuss the theoretical framework of AfL and review empirical studies from TEYL contexts. This will offer insights into whether AfL could be considered an appropriate method for the assessment of YLLs, as discussed in the previous section. First, in Section 2.3.2 the discussion focuses on the concepts of *assessment, learning, assessment for learning* and *formative assessment*, in order to subsequently engage with the debates about implementation and impact that AfL has been reported to have (Section 2.3.3). Section 2.3.4 reviews relevant studies from contexts closely related to TEYL, including EAL in primary education, and from TEFL contexts with learners older than the cohort in the current study, namely adolescents and adults. The discussion informs the research questions presented in Section 2.4.

2.3.2 The concepts involved in Assessment for Learning

2.3.2.1 Assessment

In language teaching, assessment is a term that encompasses complex concepts related to making quantitative and/or qualitative judgements about learning processes and their outcomes (Scriven, 1967). Drummond (2003) conceptualises the process of assessment in terms of teachers gathering and interpreting evidence of students' learning and using that knowledge to make decisions. It is argued that the purposes of such decisions are of crucial importance in determining the purpose and function of assessment and may impact on the choice of assessment methods. This is supported by James (2013), who emphasises that fitness for purpose is an overarching principle that should guide all assessment practices.

To satisfy the fitness for purpose condition, it is crucial to consider the functions of assessment. A distinction is commonly drawn between summative and formative functions of assessment. Typically, summative assessment (SA) is conducted periodically to measure learners' progress (Stoynoff, 2012). Its outcomes are often reported quantitatively, as a percentage or a grade, with reference to an explicit set of attainment criteria. This form of assessment is frequently contrasted with formative assessment (FA), which is viewed as a less formal, on-going, classroom based process

that seeks to gather data demonstrating students' understanding and gaps in their knowledge and uses those insights to move learning forward (Stoynoff, 2012).

Formative functions of assessment were originally explored in the 1960s by Scriven (1967), whose notion of the formative evaluation of teaching programmes focused on how to improve in the future in contrast to summative assessment which had a perspective on what had passed. Bloom (1984) used the term formative with reference to students. He investigated on-going formative assessment in one-to-one tutoring which allowed the tutor to identify the tutee's errors, inform the learner about them (feedback) and provide immediate intervention to rectify the errors (correctives). His understanding of FA implied a degree of integrating the assessment tasks with the teaching process and emphasised the purpose of formative assessment as furthering learning. The majority of research on formative assessment was conducted in 1980s and 1990s. During that time the term AfL was coined to emphasise the contribution that assessment is expected to make to the learning processes (Gipps, 1994).

More importantly, the emergence of the term AfL demonstrated a shift in thinking about assessment and learning. In language assessment, the mid-1990s witnessed developments in performance-based (Upshur & Turner, 1995) and interactional approaches to assessment (Bachman, 2007). Carless, Joughin & Liu (2006) noticed that there was a shift from understanding assessment as measuring students' performance towards recognising the influence of assessment on teaching and learning. Notably, Prodromou (1995) discussed assessment practices that had negative *backwash* (now more commonly referred to as *washback*) in FLL. These included limiting response time for test takers, assessment methods that often involved single, discrete-point testing with limited context, which valued form over content and accuracy over language development as well as failure to align assessment procedures with teaching pedagogy. These three areas are central to the shift from psychometric testing towards more communicative and context sensitive approaches to assessment. From the 1990s assessment procedures increasingly came to account for the characteristics of learners, tasks, contexts as well as the consequences of assessment and the need to integrate it with the teaching and learning process (Stoynoff, 2012).

It is important to note that the summative and formative functions of assessment can be implemented through a myriad of assessment methods but the functions themselves are

not equivalent to methods. For example, theoretically, summative tests may be used for formative purposes. However, Klenowski (2011) warned against the frequent implementation of summative tests to replace embedded classroom practices arguing that such implementation is not AfL. Harlen (2005) argued convincingly that simplistic interpretations of formative (as 'mini' summative) and summative (as an aggregation of formative) assessment in policy documents in England and Scotland led to teachers' misunderstanding of the formative function of assessment. Harlen (ibid.) claimed that since formative assessment is most commonly conducted by teachers in classrooms, the majority of teacher and classroom-based assessment tends to be erroneously labelled as formative, regardless of its actual function. She argued that, as a result, there is a lack of genuine formative assessment in schools and that this type of assessment is especially important for learning with understanding: i.e. deep learning. A similar critique of (mis)interpreting AfL and a call for a more genuine AfL was presented by Swaffield (2011) (see Section 2.3.2.3).

Different functions of assessment, viz. *bureaucratic*, *pedagogic* and *learning*, were proposed by Rea-Dickins (2001). The bureaucratic function includes externally required assessment used, for example, for reporting purposes. The latter two are internal to the school. The distinction that Rea-Dickins made between pedagogic and learning functions is one that merits further elaboration in the context of the present study. Assessment used for pedagogic purposes informs decisions made by professionals responsible for planning and delivering teaching. The learning function is distinct from the pedagogic one in that it focuses on learning through assessment and on the learner's role in that process. Rea-Dickins (ibid.) argues that the learning function of assessment encourages learners to become engaged in the process of learning and facilitates metacognitive reflection through developing learners' awareness, understanding and knowledge. This resonates with Harlen's (2005) claims about the development of deep learning. Evidently, the conceptualisation of learning seems an important component of the discussion on assessment. Hence, the following section focuses on the relationship between learning and assessment.

2.3.2.2 Learning

The shift in thinking about assessment discussed in the previous section was related to the developments in how learning was conceptualised. Watkins (2003) listed three views

of learning: behaviourist, cognitive constructivist and socio-cultural. The former is concerned with stimulus-response learning and in FL teaching is manifested in audio-lingual approaches. The second perspective focuses on the importance of processing information and on individual characteristics (see also Section 2.2.1). The third perspective understands learning as externally mediated through social interaction. The proponents of this perspective argue that socially mediated interaction, which can be affected by individual and cultural factors, constitutes the means for developing individual knowledge and learning (Shepard, 2006).

In criterion-referenced contexts, where attainment targets for the course/year are provided, the expected outcomes of learning are predefined. However, learners could presumably take different trajectories to achieve those outcomes. Hence, the ongoing learning that happens during the term is viewed as a non-linear process, specific outcomes of which may not be possible to (pre-) determine. Carr (2008) refers to such outcomes as 'fuzzy' (p. 37). This is consistent with the understanding of learning as a dynamic process, not a static performance. However, it seems that the process of learning and performance are not necessarily mutually exclusive terms. For instance, Sadler (2007) considered confident, accurate and independent performance to be a way of demonstrating the outcomes of learning. The important feature of performance that demonstrates learning is its dynamic nature; it can be adapted and improved, hence demonstrating progress in learning. Dweck (2000) distinguished between learning and performance goals. She argued that the former focus on developing one's own skills and understanding, while the latter focus on 'winning positive judgements of your competence and avoiding negative ones' (p. 15). Evidently, the understanding of the term performance by Dweck (ibid.) differs from that of Sadler (ibid.). The understanding of learning adopted in the current study encompasses the setting and achieving of learning goals as defined by Dweck (ibid.) and adopts Sadler's (ibid.) interpretation that the process of learning can be demonstrated through performance. Language assessment that focuses solely on performance with reference to predetermined outcomes is referred to as convergent (Pryor & Crossouard, 2008). This can be contrasted with the type of assessment that promises to move learning forward, which is divergent (Pryor & Crossouard, 2008). It aims to discover what a learner knows and can do to subsequently use that insight to support learning. To sum up, it is argued here that learning can be demonstrated by performance and facilitated by setting learning goals that may be

supported by assessment that is divergent in nature, although presumably constrained by the curriculum.

The notion that learning could be facilitated by assessment has attracted considerable research attention, especially following the claims about the potential of AfL to raise achievement (Black & Wiliam, 1998). Black and Wiliam (ibid.) reviewed 250 articles about studies of formative assessment and concluded that it does improve learning. Their findings were criticised by Bennett (2011) and Dunn and Mulvenon (2009), who questioned the validity of comparing the diverse studies that were included in this meta-analysis or the methodological rigour of those studies. Nevertheless, the Black and Wiliam (ibid.) article inspired further research. The outcomes of that research are complex and although some views are that FA and AfL may benefit the learning process (e.g. Earl, 2012; Ruiz-Primo & Furlak, 2006; Sly, 1999) there is little empirical evidence to support such claims (Bennett, 2011; Dunn & Mulvenon, 2009). Another significant issue was also highlighted, namely that there is not a commonly shared definition of what exactly FA and AfL mean (*definitional issue*, Bennett, 2011). Bennett (ibid.) argued that without a well-established and clear theoretical framework, it is impossible to conduct research that would provide conclusive empirical evidence to support the claims proposing the positive impact of AfL on raising achievement, as suggested by Wiliam (2009) and others. This issue is further complicated by inconsistent uses of the terms FA and AfL. As these are of key importance to the present study, the next section discusses the distinction between AfL and FA.

2.3.2.3 Assessment for Learning and Formative Assessment

This section clarifies the understanding of FA and AfL adopted in the current study. The discussion draws out a number of key differences between FA and AfL with reference to the timing, purpose, participants and beneficiaries of the assessment process.

This study adopts the distinction between the functions and purposes of assessment proposed by Wiliam (2011). In his discussion of the differences between AfL and FA, Wiliam (ibid.) argues that AfL focuses on the purpose of assessment, whereas FA focuses on its function. He quotes Black, Harrison, Lee, Marshall and Wiliam (2003) to demonstrate that AfL may be designed to collect information that can promote learning but it does not become 'formative' until that evidence is actually used to benefit the learning process. However, this reasoning seems strictly theoretical, in that, if an

assessment procedure is designed to benefit learning but consequently the evidence is not used to that end, it raises questions about the quality of the teaching/assessment. Nevertheless, the terminology seems helpful for distinguishing between the terms *formative* and *for learning*.

The notion of outcomes seems of central importance in defining FA and AfL. For example, the Assessment Reform Group in England defined AfL as:

> the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there (ARG, 2002, p. 2–3).

This definition was published together with ten principles (Appendix 1) that further defined AfL as an integral part of teaching and learning and recognised the importance of the affective aspect of assessment. However, neither the definition nor the principles clarified how to set learning goals (sometimes referred to as 'next steps' in learning). Hence, it seems that it was largely down to teachers to decide how to enact these principles. Conceivably, in educational cultures where curriculum objectives are tightly prescribed, teachers tended to interpret AfL as a means to setting attainment targets that were specified in the curriculum. This has led to calls for more genuine AfL (Swaffield, 2011) and for ensuring a better understanding of AfL by teachers (Harlen, 2005).

A more recent definition of AfL emphasises the importance of the timing, and indirectly the beneficiaries, of the assessment process. It states that AfL involves:

> Students and teachers, using evidence of learning to adapt teaching and learning, to meet immediate learning needs, minute-to-minute and day-by-day. (Thompson & Wiliam, 2007, p. 6)

This definition highlights one of the distinctions between FA and AfL, namely, that the formative function of assessment is concerned with using evidence in the future to benefit teaching and learning. This may entail the use of data gathered through assessment to improve the teacher's skills of delivering specific content (i.e. to be formative for the teacher/teaching but not for the learner/learning) or to evaluate and improve the curriculum (i.e. to be formative for the programme). A similar interpretation of the formative function of assessment is evident in the claims made by Broadfoot, Daugherty, Gardner, Gipps, Harlen, James and Stobart (1999) who propose that formative

assessment can help teachers inform future practice while not helping learners further their learning. For assessment to genuinely be *for-learning,* it should benefit the learning process of the learners who are being assessed. This is why the immediate use of assessment evidence to benefit learning seems to be at the heart of AfL. With reference to timing and beneficiaries FA seems to be a broader term than AfL. This understanding agrees with Swaffield's (2011) interpretation of AfL, presented in Fig. 2.1 below.
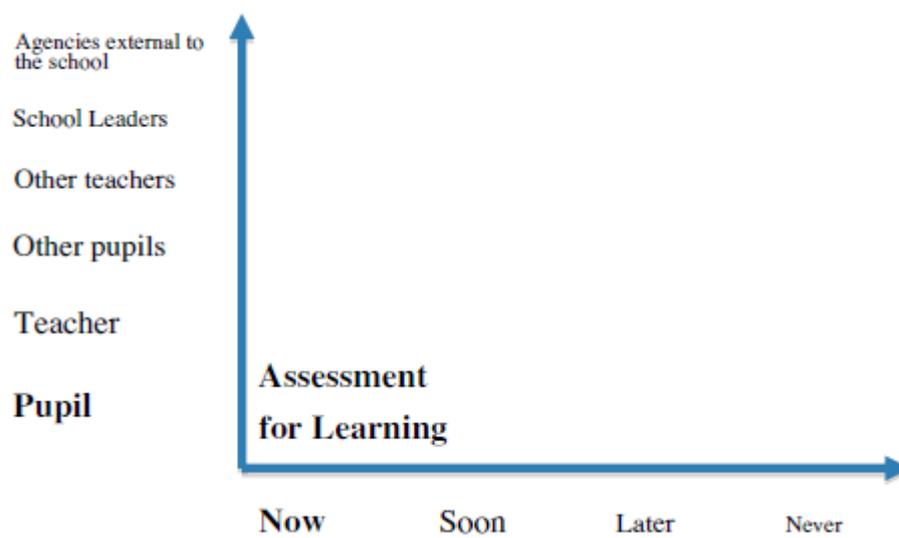
Figure 2.1: Assessment for Learning - beneficiaries and timing, reproduced from Swaffield, 2011

Swaffield (2011) interprets AfL as a form of assessment that has an immediate impact on pupils' learning. The more deferred the impact and/or the further from the pupil it is, the less for learning the assessment becomes. This suggests conceptualising AfL as a continuum and resonates with Black and Wiliam's (2009) definition of formative assessment, which suggests that formative practice is not a nominal category but that classroom practice can be formative to a certain extent:

> Practice in a classroom is formative **to the extent that** evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited. (Black & Wiliam 2009, p. 10) (my emphasis)

This definition highlights the need for assessment-derived evidence to be used to inform decisions about the next steps in learning. As in the case of previously quoted definitions,

'the next steps' are open to various interpretations and contextual factors. Interestingly, Black and Wiliam (2009) do not use the term *assessment* in their definition but replace it with *practice*. This is a welcome development in publications about FA and AfL as it seems to better describe the nature of this form of 'assessment'. As previously discussed (2.3.2.1), the terms seem to have been simplistically interpreted in some educational contexts, which has resulted in their formative function being compromised.

The lack of a commonly accepted theoretical framework and terminology associated with AfL and FA has been referred to as a definitional issue (Bennett, 2011) and was addressed by the Third International Conference on Assessment for Learning in Dunedin, New Zealand in March 2009. A definition was agreed during the event according to which AfL is understood to be:

> part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance on-going learning (Klenowski, 2009, p. 2).

This conceptualisation emphasizes the importance of enhancing on-going learning as the primary purpose of assessment. Unlike the previously quoted definitions, it provides examples of AfL methods. This seems to partially address the problem of misinterpretation while attempting to implement AfL.

The discussion of the theoretical framework of AfL has so far indicated that participants and purposes of using assessment are the main criteria for distinguishing between FA and AfL. Firstly, all the above definitions acknowledge students' (and sometimes their peers') and teachers' agency in the assessment process. Secondly, the purpose of AfL seems to be to advance the learning of learners involved in the learning situation at hand, while FA seems to have a broader sense, in that it can benefit other participants of the learning process (shown by the y axis in Fig.2.1) or have deferred results (shown by the x axis in Fig. 2.1, except for 'never'). Following from Swaffield (2011), AfL in the current study is understood as assessment practice that benefits the learning during which it occurs and engages teachers, learners and/or their peers. This definition of AfL is not dissimilar to Black and Wiliam's (2009) definition of formative practice as quoted above. It is also important to note that the use of the term assessment in AfL has been critiqued as AfL is related predominantly to feedback and learning.

The theoretical framework for AfL adopted in the current study is consistent with the notion of *informal formative assessment* (IFA) proposed by Ruiz-Primo (2011). IFA consists of 'small-scale, frequent opportunities teachers have for collecting information about their students' progress towards the learning goals they have in mind' (p. 16). Ruiz-Primo (ibid.) argues convincingly that such assessment occurs predominantly through *assessment conversations*, i.e. 'dialogues that embed assessment into an activity already occurring in the classroom' (p. 17). The primary aim of assessment conversations is to provide evidence of what the learners are thinking in order to enable the teacher to move the learning forward. Ruiz-Primo (ibid.) proposes that informal formative assessment can be effective in facilitating learning if 'assessment conversations are learning goal-guided' (p. 17); and 'dialogic and interactive in nature' (p. 18); and are used for instructional scaffolding and enculturation: i.e. 'to immerse students into the language, culture, and artefacts of the academic discipline' (p. 18). This view seems to be rooted in the shift towards conceptualising classes as well as teaching and learning processes from a socio-cultural perspective (2.2.2). The conceptualisation proposed by Ruiz-Primo (ibid.) suggests that it is important to examine the interactions that happen during the use of AfL in order to better understand how AfL may impact on learning in the TEYL classroom. Hence, it provides a rationale for the choice to focus on interaction as the feature of a class that can offer insights into the impact of AfL in TEYL classrooms (RQ3).

2.3.2.3.1 Practical examples of AfL in TEYL classes

This section aims to provide the reader with examples of how AfL, as defined above, could be implemented in the classroom. Because little practical guidance for the implementation of AfL in TEYL contexts exist, the examples are based on guides published for the mainstream primary classroom. However, the practical implementation, including types of AfL techniques, has been researched as a part of the current study and the results are reported in Part Two of Chapter 4 with a detailed account of techniques in Appendix 18.

Clarke (2005) describes a number of practical techniques intended to enable teachers to implement AfL in elementary classrooms. These include: separating the learning objectives from the context of learning, using criteria for success, effective questioning and focusing feedback on the learning objectives or criteria for success. All these techniques can be presented visually, e.g. with the use of pictures of simple diagrams that

illustrate for children what the expectations of their performance in each lesson are and/or to what extent they have been met. Some examples of practical implementation could include using an analogy to traffic lights, by indicating with a colour (red, amber or green) to what extend a child has met their learning goal. Another example would be using Success Criteria in the form of a list of items that should be demonstrated by a child while they are completing the task in hand and pointing them out to the child while they work, thus prompting them to monitor their own performance. Peer-assessment could be facilitated by organising children in pairs to form Learning Partners, and providing opportunities for the partners to monitor and evaluate one another's learning. Also, teachers could ask select questions that guide learning and stimulate thinking as opposed to testing questions with right or wrong answers predefined by the teacher.

Having presented the theoretical framework of AfL that is adopted in the present study, and having provided an overview of how AfL could be implemented in classrooms with learners aged 7-11, the discussion continues in the next section by reviewing empirical studies on assessment in TEYL contexts.

2.3.3 Empirical studies on assessment in TEYL contexts

This section reviews empirical research into assessment in TEYL contexts. First, the trends in researching assessment are explored. Subsequently, the attention shifts to reviewing the literature on AfL as defined in the previous section.

It cannot be claimed that a substantial body of research on how language learners at primary age may be assessed exists. It seems unfortunate that the issue of assessment is often overlooked in discussions about TEYL. For example, Copland and Garton (2014) in the editorial to an ELT Journal special issue on Young Learners did not mention assessment in their discussion of key themes and future directions for TEYL. Yet, Nikolov and Mihaljević Djigunović (2011) suggest that the main issue in researching assessment in TEYL contexts is connected to the construct of assessment. The published research on assessment focuses on the following:

- The assessment of language proficiency at the end of primary education (Edelenbos & Vinje, 2000; Johnstone 2000),
- The deployment of the Common European Framework of Reference (CEFR) descriptors (Council of Europe, 1996), as a means of organising portfolio

assessment (Hasselgreen, 2005) or in large scale standardised tests (Bailey, 2005)

- Classroom-based assessment (Hill & McNamara, 2012), including teachers' competence to conduct it (Edelenbos & Kubanek-German, 2004) and teachers' practice (Butler, 2009; Hild & Nikolov, 2010)
- The formative function of assessment (Gattullo, 2000; Hasselgreen, 2000)
- Self-assessment (Butler & Lee, 2006, 2010).

The studies quoted above have investigated summative as well as formative functions of assessment. It is useful to review both perspectives. The studies that focus on summative assessment are reported so as to identify the insights that they provide into language learning and progression in childhood. The studies that investigated classroom-based assessment practices and the formative function of assessment, including self-assessment, are reported in greater detail as they provide insights with direct relevance to the current study.

2.3.3.1 Studies on the Assessment of Learning

This section outlines research on assessment in TEYL contexts. The review aims to provide an overview of research, as the context within which empirical studies on AfL can be located.

Edelenbos and Vinje (2000) reported a comparative analysis of EFL national tests results administered with 12 year old children at the end of primary school in the Netherlands. The tests for English were conducted in 1991 and 1996 and focused on assessing listening, reading, receptive word knowledge, use of a bilingual wordlist, speaking and pronunciation. The assessment procedures included pen and pencil tests for all areas except for speaking, pronunciation and productive word knowledge, which were assessed individually through a discussion with an English speaking partner and involved reading out sentences and naming objects in pictures. The results indicated that the 1996 cohort performed lower than the earlier cohort. The authors argued that such differences might have been due to the amount of time of exposure and institutional characteristics such as school size or teaching practices. Their interpretation that time of exposure was a factor in FL achievement in childhood corroborate with the findings of research suggesting that children learn implicitly and that, therefore, a longer time of exposure may be beneficial (e.g. Muñoz, 2006). Edelenbos and Vinje (ibid.) also found that

students whose teachers used communicative methodology tended to perform better on reading tasks compared with learners who were taught by grammar translation methods. This indicates that teaching methods may impact on the achievement of particular language skills. The authors concluded that 'setting clear goals, sequencing materials, frequent questions to monitor progress in the learning process, opportunity to learn, testing and quality feedback are all important characteristics from which early foreign language learning can benefit tremendously' (p. 160). These findings are important to the current study as they indicate that some of the processes inherent in AfL (setting clear goals, guiding questioning and meaningful feedback) can benefit FLL in childhood. This suggests that classroom teaching and assessment practice are interlinked and of importance for learning (see also the study by Butler (2009) at the end of this section).

Another study that looked at end of primary school achievement was conducted in the context of FL teaching in primary schools in Scotland (Johnstone, 2000). Johnstone's (ibid.) study reported that the procedures addressing the functions of assessment and implemented in the last year of primary education (11 year olds) resembled the bureaucratic, pedagogical and learning functions proposed by Rea-Dickins (2001). The study highlighted that assessment methods can be sensitive to contextual factors and to the stakeholders of assessment. The author also argued for 'more consensus on the aims and intended outcomes' (p. 140) of FL instruction in primary schools and for research that can demonstrate what should constitute various levels of proficiency.

An interesting study that provided some insight into data-driven level descriptors was reported by Huhta et al. (2014). It focused on applying the Common European Framework of Reference (CEFR) criteria to texts written by learners aged 13-16 in Finland. Huhta et al (ibid.) found that raters who were using CEFR descriptors used criteria not mentioned in the scales. This points to issues concerning the reliability of assessment as the subjective human factor might have affected the results. The study included a data set that could be used to analyse linguistic features at different levels of proficiency in adolescents. This research focus seems to be useful since it might address the issue of lack of clarity in the construct of assessment in TEYL contexts. If CEFR were to be adapted to the younger learner context, it seems that similar studies investigate the performance of younger learners.

Attempts to adopt the CEFR descriptors and convert them into assessment portfolios for primary aged children have been made in other European countries: e.g. in Ireland (Little, 2005), France (Debyser & Tagliante, 2001), England and Wales (Cameron, 2003) and for their secondary counterparts (Hasselgreen, 2005; Lam & Lee, 2010). Most of the portfolio assessment tools developed for primary contexts tend to provide descriptors predominantly for levels A1 and A2 of CEFR, except for Norway where descriptors as high as B2+ have been developed (Hasslegreen, 2005).

Hasselgreen (2005) reported a study in which CEFR descriptors were initially translated into Can-do statements for 13-15 year olds. These subsequently served as the basis for devising descriptors to test reading and writing at levels A1-B2+ in primary schools that were included in the National Testing of English Project (partially computer-adaptive) in Norway. The reading tests were computer based, while writing was teacher rated and standardised by expert spot-checks. Hasselgreen (ibid.) highlighted the requirements for good assessment to provide positive feedback (i.e. what learners can do). She argued that tasks appropriate for assessment should captivate attention and interest as well as allow for some support to be provided to YLs.

Given that most assessment frameworks for YLLs have focused on low levels (A1-A2), according to Nikolov and Mihaljević Djigunović (2011) '(l)ow proficiency levels need to be defined and described along a continuum in small steps so that children's relatively slow development can be documented' (p. 109) and these steps should take into account the development of learners' literacy levels. This resonates with Johnstone's (2000) call for consensus on what level of proficiency can be expected from children in primary schools and with Inbar-Laurie and Shohamy's (2009) call for aligning assessment practices with the foci of teaching programmes.

Other attempts to use the CEFR benchmarks include the development of standardized tests for children. Examples of such tests include the University of Cambridge ESOL Young Learners Exams (YLE) suite for primary and lower secondary testing at levels A1-A2, the Pearson Test of English Young Learners for 6-13 year olds at levels pre-A1 to A2, the City and Guilds ESOL Young Learners for 8-13 year olds, at levels A1-A2, and the Trinity Graded Exams in Spoken English (listening and speaking) for ages 5 and older, at levels A1-C1 but available up to level B1 for primary school aged children. Cambridge ESOL also offer a range of 'for schools' exams at levels A2-B2 but do not

specify the age of exam takers other than by stating that the exams are for children who are 'at school'. It is important to consider the washback effect (Prodromou, 1995) that such tests may have on teaching and learning. For example, Choi (2008) reports that standardised EFL tests have a negative impact on teaching and learning in South Korea, putting many primary school as well as older learners under pressure to take and perform well on such tests.

Another interesting discussion concerning the factors that shape classroom assessment practices in primary schools in Japan, South Korea and Taiwan was reported by Butler (2009). The focus was on evaluating the TEYL programmes and assessment methods. Butler (ibid.) analysed government documents and published research relating to teaching and assessment in those three countries to establish a collective knowledge base about how best to assess YLLs. Butler reported that summative tests aimed to measure the progress made by children at primary school were developed by central or local governmental organisations. However, the details of the tests were not released to the public. She also reported that although classroom-based assessment was advocated by the Korean and Taiwanese governments, teachers needed more training in how to conduct assessment and use the gathered data for summative and formative purposes. In Japan, where the government did not require any specific form of assessment, schools used self-assessment techniques. Butler (2009) argued that it was unclear what should be assessed in primary schools, e.g. which skills, to examine the quality of teaching and student achievement. This resonates with Nikolov and Mihaljević Djigunović's (2011) suggestion that clarity of the construct of assessment in TEYL is an important issue. Their discusson emphasised the importance of contextual factors, especially educational policies, in shaping classroom assessment practices and that self- assessment can be an appropriate method for conducting classroom based assessment in TEYL contexts. Insights into the nature of that self- assessment, an integral part of AfL, were reported by Butler and Lee (2006, 2010) and are reviewed in Section 2.3.3.2.

2.3.3.2 Studies on Assessment for Learning

This section reports studies that have focused on investigating different aspects of AfL in TEYL contexts. The available published research is summarised in Table 2.3 and discussed below.

Table 2.3: Empirical studies on AFL in TEYL contexts

	Reference	Focus of the study	Sample	Context
Studies on Assessment for Learning	Hill and McNamara, 2012	Processes of classroom-based assessment	3 teachers and their students aged 11 -13	Primary (Year 6) and Secondary (Year 7) schools in Victoria, Australia
	Edelenbos and Kubanek-German, 2004	teacher's diagnostic competence	49 lessons from 10 schools and the teachers who delivered the lessons	Primary schools, Germany and the Netherlands
	Gattullo, 2000	Use of AfL in TEYL	4 teachers (3 generalists and 1 specialist) and 70 learners aged 8-10	Primary school Italy
	Butler and Lee, 2006	on-task and off-task self-assessment	70 learners aged 9-10 and 81 learners aged 11-12	Primary school in Seoul, South Korea
	Butler and Lee, 2010	effectiveness of self-assessment	254 learners aged 11-12 in 2 schools	Primary schools in Seoul, South Korea
	Hasselgreen, 2000	Measuring strengths and weaknesses in FL performance	1000 learners aged 11-12 in 34 schools	Primary schools in Norway

As Table 2.3 shows, Hill and McNamara (2012) researched the process of classroom based assessment (CBA). CBA is defined as 'any reflection by teachers (and/or learners) on the qualities of a learner's (or group of learners') work and the use of that information by teachers (and/or learners) for teaching, learning (feedback), reporting, management or socialization purposes' (p. 396). This understanding incorporates both the summative and formative functions of assessment. It seems to indicate that in order to move learning forward (i.e. satisfy the formative function) an assessment opportunity must first reflect on what learners can already do (i.e. it must contain a summative component). *Assessment opportunities* are understood by Hill and McNamara (2012) as 'any actions, interactions or artefacts (planned or unplanned, deliberate or unconscious, explicit or embedded) which have the potential to provide information on the qualities of a learner's (or group of learners') performance' (p. 398). Their (ibid.) study was based in classrooms

with 11-13 year olds who studied Indonesian as a foreign language in Australia. The researchers used lesson observations and case studies. They proposed a framework for investigating classroom-based assessment that was sensitive to how teachers plan, enact and follow up on assessment. The framework highlighted four important areas in CBA. These are: what teachers do, what information they collect to inform assessment, and teachers' and learners' theories and beliefs about learning and assessment. These are useful in the context of the current study as they imply that it is important to research teachers' actions together with their understanding of their actions when investigating assessment.

Implementing CBA, as defined above, seems to require a degree of competence in evaluating learning. However, primary language teachers are rarely experts in language assessment (Hasslegreen, 2005; Johnstone, 2000). In the European context, two studies of classroom-based teacher assessment were reported by Edelenbos and Kubanek-German (2004), who used data from 49 lesson observations and from retrospective teacher interviews to discuss the concept of *teachers' diagnostic competence*, i.e. 'the observational and interpretative competence shown by teachers during classroom teaching' (p. 277). Although the authors did not explicitly refer to AfL, they reported 11 'potentially diagnostic' (p. 264) classroom teacher behaviours (Appendix 2) that resemble aspects of AfL with reference to Black & Wiliam (2009) and Black et al. (2003). The study evidenced that diagnostic behaviour accounted for a relatively small amount of classroom time (11.3%). The authors concluded that teachers' diagnostic competence is an important factor in ensuring fairness and the validity of classroom assessment; they recommended ways for developing it through pre- and in-service professional development.

It has been suggested that another way in which diagnostic efforts in TEYL classrooms can be aided is through purpose-designed assessment materials. A case in point is a Norwegian project called EVAluation of English in Schools (EVA), within which assessment materials were developed to evaluate the strengths and weaknesses of YLLs' performance in primary schools (Hasselgreen, 2000). The materials had a formative function in that they informed classroom practices. The materials provided an opportunity for the children to help find a stolen elephant through conducting a series of activities and were implemented over 2 weeks in episodes of about 25 minutes each. The

tasks focused on assessing listening, writing and the use of language (focusing on form) as well as speaking, which was assessed separately in a paired activity where children were asked to play a game using pictures. The testing was trialled on 1000 learners. The results indicated that the children were highly engaged in the activities. This assessment tool also included a self-assessment component; the YLLs were required to indicate whether they could complete the tasks independently on a 4-point scale (yes / mostly / a bit / no) and were also asked to rate the materials. The teachers were provided with scoring sheets to record the children's progress, areas needing development and the results of their self-assessments. This project provided evidence that assessment materials used in TEYL classrooms should be well contextualised and engaging for YLLs. It also seems to indicate that a continuous approach offers an opportunity for gaining reliable insights into the FL achievement of children.

Another insight into how aspects of AfL can be implemented in a TEYL context was reported by Butler and Lee (2006), who examined the validity of on-task and off-task self-assessment in a TEYL context (9-12 year olds) in South Korea. The results of the summative tests and teacher assessment were compared with results of the learners' self-assessment. Butler and Lee (ibid.) concluded that on-task self-assessments, where self-evaluation takes place immediately after a learner has completed a task, are more accurate than off-task self-assessments that are unrelated to a specific task and are less influenced by contextual and individual factors. It was found that older learners (11-12 years old) were able to self-assess more accurately than their younger counterparts (9-10) and that all the children could develop accuracy in their self-assessments over time. This study is especially informative because Butler and Lee (ibid.) discussed self-assessment with a temporal reference to completing classroom tasks. Most importantly it emphasised the importance of integrating self-assessment with teaching and learning: i.e. on-task self-assessment.

Subsequently, Butler and Lee (2010) reported an intervention study of 254 learners aged 11-12 conducted in a TEYL context in two public elementary schools in South Korea. The aims were to investigate whether learners develop accuracy in self-assessment over time and the effectiveness of self-assessment in supporting learning. A series of self-assessment tasks were administered every two weeks over five months. Data were collected from a pre-test, a post-test, a student survey and two teacher interviews. Butler

and Lee (ibid.) observed that the implementation of self-assessment in both schools differed. In one school the focus was on the role of self-assessment in increasing positive feelings, while in the other it was on increasing achievement. They concluded that the differences in how AfL was perceived and implemented were influenced by teacher beliefs. This study provided useful evidence that the primary age children were able to improve the accuracy of their self-assessments over a relatively short period of time but only in the intervention group. The learners in the control group demonstrated declining accuracy in self-assessment. Butler and Lee (ibid.) suggest that this may have been due to the fact that the children started learning with a positive self-concept but that this declined as they accumulated learning experiences (see also Section 2.2.3). This suggests that there is a link between self-assessment in TEYL classrooms and the research on affect as affect may impact on the accuracy of self-assessments. The study also suggests that children could be trained in how to self-assess. Additionally, the same authors reported evidence of self-assessment having a positive, although marginal, effect on performance. It ought to be noted that no delayed post-test was conducted to indicate if the results were sustained over a longer period of time. Butler and Lee (ibid.) acknowledge that the children in one of the sites were receiving extracurricular English lessons but no data were available. Hence some of the improvements might have occurred as a result of factors other than the implementation of self-assessment. The study also revealed that teachers found it challenging to provide feedback to the children because they were concerned that it might increase the already high levels of competitiveness between learners.

A unique study describing the use of AfL in a TEYL context was reported by Gattullo (2000). The participants comprised 70 learners aged 8-10 and four teachers in a primary school in Italy. This was the first and, to the best of my knowledge, the only study that investigated how AfL was implemented in the TEYL classroom up to the time of writing the current thesis. All the previously quoted studies focused on aspects of AfL such as self-assessment, teachers' diagnostic competence or processes involved in CBA. It also seems useful to note that this study included the youngest participants of all the studies reviewed in the current section. Gattullo (ibid.) collected data from 15 hours of audio recorded lessons. She reported nine categories of assessment features that were used in TEYL classrooms: questioning/eliciting, correcting, judging, rewarding, observing process, examining product, clarifying task criteria and metacognitive questioning. Her

findings suggest that the teachers were willing to try implementing AfL and were enthusiastic about doing so. These findings indicate that it was possible to implement AfL techniques in TEYL classes despite the learners' low levels of language proficiency and their young age. However, it should also be noted that the teachers in Gattullo's (ibid.) study tended to use techniques such as questioning and correcting significantly more frequently than techniques which the author considered more beneficial for learning such as 'observing process' or 'metacognitive questioning'. She indicated that the techniques favoured by the teachers were more naturally compatible with the teaching methodology used. She suggested that it may be important for CBA practices to be compatible with the teaching methodology.

Most of the research on assessment reported above was conducted with students aged 8-12. In this age group, in many educational systems, children have already developed some level of literacy. There seems to a gap in the research on assessment that includes age groups younger than 8 years old. This should be addressed in order to investigate FL assessment practice with learners who are only beginning to develop literacy skills.

This section has discussed the published research on AfL in TEYL contexts. The review shows that a limited number of studies have investigated this area. To provide a more comprehensive background to the current study, a number of studies on AfL from the EAL and TEFL literature are reviewed to provide insights into what is already known about AfL in similar educational contexts.

2.3.4 Assessment for Learning in EAL and TEFL

This section reviews the empirical research on AfL in FL teaching contexts to demonstrate the diversity and range of coverage. The studies are summarised in Table 2.4 and discussed in two subsections. Section 2.3.4.1 reports on studies about EAL in primary schools while Section 2.3.4.2 reviews studies about TEFL contexts in classes of adolescents.

2.3.4.1 Empirical studies on AfL in EAL contexts

The studies conducted with learners in primary schools for whom English, the language of instruction, is an additional (second or third) language also provide useful insights in the use of AfL in YLs language classroom. The studies relevant here focused on investigating teacher and learner evaluative competence (Hawe & Dixon, 2014; Rea-

Dickins, 2001; Rea-Dickins & Gardner, 2000), self-assessment (Dann, 2002) and interactions that occur in assessment episodes (Leung & Mohan, 2004; Rea-Dickins, 2006).

The use of AfL and how it informed decisions about the curriculum and language learning were investigated by Rea-Dickins and Gardner (2000). Their analysis of classroom data from nine inner-city schools in England revealed that teacher assessment was generally regarded as low-stakes. They also reported that important decisions regarding the need for focused language support were grounded in the interpretation of learners' classroom performance. However, the authors questioned whether the teachers were adequately trained to make such decisions. This research highlighted the importance of the teachers' interpretation of classroom interactions. It emphasized the teachers' agency in classroom-based assessment, a concept that is further discussed by Rea-Dickins (2001).

Rea-Dickins (2001) discussed the concept of an assessment cycle, i.e. the stages involved in teacher-led formative assessment. She used data from teacher interviews and classroom observations to develop a model for analysing teacher assessment decision-making. She also reported on how the teachers enacted formative assessment through oral interactions with the learners and concluded that there is a need for more research that indicates what constitutes quality in AfL and its impact on facilitating language learning.

Table 2.4: Empirical studies on AfL in primary EAL and TEFL settings.

	Reference	Focus of the study	Participants	Context
AfL in Primary EAL	Rea-Dickins and Gardner, 2000	nature of classroom assessment	9 inner city schools	Primary schools, England
	Rea-Dickins, 2001	assessment cycle	Inner city schools with 98% EAL learners	Primary schools, England
	Rea-Dickins, 2006	teacher-learner interactions	2 language support teachers and 1 mainstream teacher and their learners aged 6-7	Primary schools, England
	Leung and Mohan, 2004	interactions during classroom-based formative assessment	2 Year Four classes (8-9 year olds)	Primary school, England
	Dann, 2002	implementation of self-assessment	Children aged 7-11	Primary schools, England
	Hawe and Dixon, 2014	students' evaluative competence	3 teachers of students aged 9, 11 and 12	New Zealand
AfL in TEFL	Lee and Coniam, 2013	AfL in writing	2 teachers 12 year old students	Secondary School in Hong Kong
	Lee, 2007	feedback in writing, including AfL	26 teachers 174 pieces of feedback, students aged 12-16	Secondary School in Hong Kong
	Sidhu, Fook and Sidhu, 2011	Students opinions about AfL	2684 students	Malaysian secondary schools
	Colby-Kelly and Turner, 2007	AfL in EAP Speaking	9 teachers, 42 students: adults	pre-university course in Canada
	Cheng, Rogers & Hu, 2004	Classroom-based assessment methods and procedures	267 teachers	Canada, China and Hong Kong tertiary education

In addition to exploring teachers' competence in assessment procedures, research has also provided insights into students' evaluative competence (Hawe & Dixon, 2014). Hawe and Dixon (ibid.) used data drawn from the classroom observations of three primary teachers working with 9, 11 and 12 year olds. The teachers in the study all expressed positive attitudes towards assessment practices that allow students to be active participants of the assessment process. Hawe and Dixon (ibid.) reported qualitative differences in implementing the same AfL techniques in a writing class by the three teachers and linked the differences to changing the traditional roles of the students and the teachers. The authors stressed the importance of building evaluative competence through feedback. Although this study was conducted in a language arts context, it focused specifically on developing writing skills, which is often an important focus of TEFL programmes, and was conducted with learners of ages similar to those of the participants in the current study. The findings highlighted the importance of feedback in building the ability of learners to self-assess, hence suggesting a link between feedback and the development of metacognitive strategies.

A study in the use of self-assessment with primary school learners, similar to others carried out in TEYL contexts (e.g. Butler & Lee, 2010), was carried out by Dann (2002) in a language arts context. She discussed a number of crucial considerations for implementing self-assessment, including the need to integrate self-assessment with classroom instruction; to provide opportunities to discuss the criteria and results of assessment with the teachers; and to act on them. This resonates with the discussion in Section 2.3.3.2.

A different focus on exploring the implementation of AfL was taken by Rea-Dickins (2006). From analysing teacher-learner interactions, including feedback and learners' responses to feedback, she observed that both summative and formative functions occurred in assessment episodes. She concluded that limiting the function of assessment to one or the other could lead to minimising its potential as a language learning resource and questioned to what extent learners were aware of the assessment function of the interactions. Her work suggests that language learning can happen through assessment during interactions in the classroom. Therefore, it may be worth analysing interactions that happen during CBA in order to identify whether and how learning can take effect through interaction.

Another study that explored classroom interactions was conducted by Leung and Mohan (2004). They reported a case study conducted in two multiethnic Year 4 classes (8-9 year olds) with EAL learners in England. The authors used Mohan's (2003, cited in Leung & Mohan, 2004) three-part model to analyse extracts of classroom discourse that demonstrated student decision making process during the use of AfL. They concluded that AfL offers opportunities for interaction between students, peers and teachers thus allowing for learner agency in providing feedback. By providing formative feedback through eliciting the reasons for answers, the teachers encouraged learning. This emphasised that the content of interactions that occur during the use of AfL are crucial to consider in order to gain insights into how AfL could facilitate learning.

2.3.4.2 Empirical studies on AfL in TEFL contexts with older learners

Relevant insights into AfL are also offered by studies conducted in TEFL contexts, especially those with participants only marginally older than the learners in the current study. Studies in this category looked at the implementation of AFL in secondary schools (Carless, 2005; Lam & Lee, 2009; Lee, 2007; Lee & Coniam, 2013; Sidhu, et al., 2011) and in adult education (Cheng, Rogers & Hu, 2004; Colby-Kelly & Turner, 2007).

Lee and Coniam (2013) collected data through questionnaires, interviews, pre- and post-tests, and lesson observations to investigate how AfL could be implemented in TEFL writing lessons with 12 year olds in a secondary school in Hong Kong. They also analysed factors that facilitated or inhibited such implementation. The findings suggest that teachers' knowledge and previous experience of using AfL and collaboration between teachers could facilitate the implementation of AfL. Furthermore, two factors were found to inhibit the implementation of AfL: the need to prepare students for external exams and the school's policy of correcting all errors. The authors suggest that students might have ignored formative feedback when presented with summative assessment results. This interpretation is consistent with Butler (1988) who demonstrated that when learners are given feedback comments alone, they show a greater interest in learning than in situations where grades or grades and comments are provided. Regarding the development of writing, despite Lee and Coniam's (ibid.) claims that AfL contributes to improving the quality of students' writing, there seem to be no empirical data in their study that linked the increased level of writing with AfL. Most importantly, no control group was included and comparisons were made between the progress made by the students in the study

(which was set within a Band 1 school, where Band 1 denotes the best achieving schools in the country) and the national average. Nevertheless, the study offers valuable insights into factors that may facilitate or inhibit the implementation of AfL in a FL classroom.

In a similar context, Lee (2007) explored the extent to which feedback on writing was exploited to move learning forward in large classes in Hong Kong. The data included 174 pieces of written feedback on writing performance from 26 secondary school teachers and transcripts of interviews with students (n=18) and teachers (n=6). His findings revealed a limited *for-learning* purpose of feedback and emphasised that the predominantly summative feedback was likely to be the result of factors external to the classroom context (e.g. school policy, external exams) in shaping assessment practice.

Carless (2005) reported difficulties in implementing AfL in Hong Kong that were similar to those found by Lee and Coniam (2013) and Lee (2007), viz. tensions existing between classroom practice and external assessment requirements. In another study, Yung (2002) reported that individual teacher beliefs and attitudes resulted in teachers implementing varied assessment practices, which points to the important role of teachers as mediators of CBA even in an exam culture like that in Hong Kong.

Findings concerning factors that could facilitate or inhibit implementation of AfL have also been reported in studies within tertiary educational contexts. For example, Cheng et al. (2004) conducted a survey of 267 teachers in China, Hong Kong and Canada to investigate the purposes, procedures and methods used in classroom-based assessment in TEFL at tertiary level. They reported that a variety of assessment procedures were being used and noted that some of the teachers explicitly indicated using the assessment data formatively. However, the results seemed to indicate a limited use of assessment to support learning in all three contexts. The analysis revealed a number of possible factors that could account for variety in classroom-based assessment. These were to do with the nature of the course, the teachers' knowledge and experience, the needs and levels of students, external assessment procedures and teacher attitudes and beliefs. These findings corroborate with the other studies that reported factors impacting on the implementation of AfL discussed earlier in the current section.

Colby-Kelly and Turner (2007) discussed the concept of an assessment and learning interface that they termed the *assessment bridge*. They used a mixed-methods approach that included curriculum document analysis, questionnaires, interviews and classroom

observations to investigate feedback practices in an advanced level pre-university EAP course with 42 students and 9 teachers in Canada. Their discussion resonates with Rea-Dickins' (2006) discussion of assessment interactions as a language learning resource. Hence, it provides additional support for the value of considering interactions as a mediator of learning though assessment.

2.3.5 Summary of Section 2.3

This section has reviewed what is known about AfL in TEYL contexts. Having explored the theoretical frameworks available in the literature, the discussion has concluded that an established theoretical framework of AfL shared by many researchers does not exist. For that reason, it was necessary in Section 2.3.2 to provide clarification of the understanding of AfL that is adopted in the present study. This was followed by a review of relevant empirical studies from TEYL contexts and other closely related teaching contexts (EAL and TEFL).

The reviewed empirical research points to the important role of teachers as agents of assessment. The most important considerations included teachers' diagnostic competence and the impact that teachers' judgments have on assessment (Edelenbos & Kubanek-German, 2004; Rea-Dickins & Gardner, 2000; Rea-Dickins, 2001) as well as teachers' understanding of and beliefs about AfL (Hill & McNamara, 2012; Lee & Coniam, 2013). The review also suggests that teachers can employ various behaviours to realise the formative function of assessment but the amount of time which they tend to spend on such activity in classrooms is limited (Edelenbos & Kubanek-German, 2004) and that teachers opt for behaviours that are naturally compatible with their teaching methodology (Gattullo, 2000).

Secondly, the review has revealed a number of factors that can facilitate or inhibit the implementation of AfL. The facilitating factors include teachers' positive attitudes and beliefs (Butler & Lee, 2010; Hill & McNamara, 2012) and their experience of using AfL (Lee & Coniam, 2013). The inhibiting factors comprise externally mandated policies or exams (Butler, 2009; Cheng et al., 2004; Gattullo, 2000; Lee, 2007; Lee & Coniam, 2013).

Thirdly, the review has highlighted those characteristics of assessment that were demonstrated to be appropriate for TEYL contexts. It has indicated that TEYL assessment

procedures should capture the attention and interests of the children (Hasselgreen, 2005), collect information about their progress in a continuous manner (Hasselgreen, 2000), provide clear goals, and enable questioning and effective feedback (Edelenbos & Vinje, 2000; Hasselgreen, 2005; Hawe & Dixon, 2014).

Fourthly, some studies have examined issues connected with implementing specific aspects of AfL, such as self-assessment. The indication from the review is that YLLs can make valid judgments about their own performance and develop that skill over time (Butler & Lee, 2006), that feedback can facilitate the development of learners' evaluative competence (Hawe & Dixon, 2014) and that self-assessment should be integrated into teaching (Butler & Lee, 2006; Dann, 2002).

Fifthly, the review has also drawn attention to the important role of classroom interaction (Leung & Mohan, 2004; Rea-Dickins 2006) in collecting assessment data and moving learning forward. Lastly, a number of authors argued that the construct of assessment in TEYL needs to be clarified (e.g. Johnstone, 2000) and that CEFR may need to be broken down into smaller steps to demonstrate progress made by YLLs (Nikolov & Mihaljević Djigunović, 2011). The review has highlighted a number of gaps in the current understanding of AfL. Most significantly, no studies that link AfL to improved achievement empirically were found. A similar lack of empirical evidence to validate the claims of AfL's efficacy was identified by Bennett (2011) and Dunn & Mulvenon (2009).

The following section introduces the research questions by explaining how they address the gaps identified through the literature review.

2.4 Research Questions

The literature reported in this chapter illustrates a number of useful lines of inquiry that are used in this section to inform the research questions.

The current chapter has discussed issues relating to the lack of a generally accepted theoretical framework and terminology for AfL. Given that there is a need for tightening the theoretical frameworks of both AfL and FA, the first research question seeks to establish how teachers understand AfL in a TEYL context. It is believed that investigating teachers' understanding of AfL is a pre-requisite for conducting research into the implementation of their practice and its impact in TEYL classes.

The review has indicated that little is known about the implementation of AfL in FL classrooms with children of primary school age, specifically 7-11 in the current study. In fact, just one study (Gattullo, 2000) reported on the implementation of AfL in a similar context. Other studies focused on different aspects of AfL; e.g. self-assessment (Butler & Lee, 2006, 2010) and feedback on writing (Lee & Coniam, 2013). Hence, with the exception of Gattullo's (ibid.) study, no descriptive accounts of the use of AfL in TEYL classrooms are available. Hence, the second research question seeks to gain insights into how teachers implement their understanding of AfL in the classroom. This offers an opportunity to report on how teachers enact their understanding of AfL and to evaluate whether AfL can demonstrate the characteristics of assessment appropriate for YLLs.

Finally, the role of interactions that happen during assessment procedures, also referred to as assessment conversations (Ruiz-Primo, 2011), can offer opportunities for moving learning forward. Hence, it seems important to investigate if interactions that occur during the use of AfL demonstrate characteristics that the literature review has shown to be beneficial for learning.

The areas summarised above are similar to those identified by Colby-Kelly and Turner (2007) in an EAP context in Canada (Section 2.3.4.2). Hence the three research questions in the current study are based on those adopted by Colby-Kelly and Turner (ibid.) but adapted to the context of the current study (see also Section 3.2.4). Detailed analysis of the similarities and differences between the two studies is presented in Appendix 7.

The research questions in the current study are:

- **RQ1:** How do teachers understand AfL after receiving a limited amount of training and being encouraged to use AfL techniques for at least one academic year when teaching English to young learners aged 7-11?
- **RQ2:**
    - **2.1:** How do teachers translate their understanding of AfL into classroom practice in a TEYL context with students aged 7-11 in a private language school in Poland?
    - **2.2:** Do teachers report any changes in their practice of using AfL over time?
- **RQ3:** What is the observable impact of AfL on classroom interactions in a TEYL context?

The next chapter reports on the design and implementation of the study used to address the three research questions.

Chapter Three: Methodology

The aim of this chapter is to provide the rationale for the methodological choices in the present study. It is divided into four sections. Section 3.1 describes the research setting, participant selection process and ethical considerations to outline the context of the study. Following that, Section 3.2 reports on how the study design was developed and piloted. Section 3.3 continues this discussion by reporting the data sets and analysis procedures. Finally, Section 3.4 considers the limitations of the methodology.

3.1 The context of the study

This section introduces the context in which the current study was conducted. Section 3.1.1 outlines FL teaching to 7-11 year olds in the Polish educational system. Section 3.1.2 describes the school in which the current study was based. Section 3.1.3 describes the characteristics of the participants included in the sample of the presents study. Section 3.1.4 aims to provide an account of the introduction of AfL which preceded the current study. Finally, Section 3.1.5 discusses the ethical considerations.

3.1.1 The FL curriculum in the Polish educational system

The present study was set within the context of teaching English to young learners aged 7-11 in a private language school in Poland. Although it was not based in a state school, a brief description of foreign language teaching in the national educational system is provided here so that the study can be situated in the broader educational context.

One foreign language (FL) is taught in state schools in Poland from Year 1 (6/7 year olds) onwards. This was introduced by a major educational reform which started in 1998 (Leowiecki, 1999). A second FL is introduced into the curriculum in Year 4 (9/10 year olds). In 2011/12, the FL most commonly taught in primary schools was English (92.5 % of learners), followed by German (7.5%[1]) (Braunek, 2013). Alongside these changes, parents sought to enrol their children on extracurricular English language courses to supplement the state provision (see Point 8 in Appendix 3).

---

[1] Please note that the fact that these two percentages add up to 100% is coincidental because the percentages quoted are for primary schools and include cases in which children learn more than one language. Children start learning a second FL in Year 4 of primary school (9/10 year olds). In 2011/12 14% students in primary schools studied two FLs and 0,08% three FLs. Braunek (2013) reports that 0.5% studied French and 0.3% Russian.

3.1.2 The school setting in which the data were collected

As a part of the process of setting up the study, a meeting was held with the director of the school. Unless referred to otherwise, the information provided in this section is based on this formal discussion in which detailed notes were taken. All the numerical information refers to the beginning of the academic year of the data collection stage (2011/12).

The school was set up in Poland in 1996 as one of the first English language providers in central Poland after the fall of communism. Initially it mainly catered for the adult and teenage markets. Courses for students aged 7 – 11 were introduced in 2004. The curriculum for this age group was organised in a set of 'Can Do' statements based on course books. There was no explicit guidance about assessment in the curriculum documents. Teachers were expected to use their professional judgment as to what assessment techniques they deployed. The school's policy was to report summative grades to parents twice a semester. This was done in a form of report cards with numerical grades from 1 (fail) to 6 (exceeds expectations) awarded in the following categories: speaking, listening, reading, writing, effort, homework and behaviour. This system was consistent with the state school grading system used from Year 4 (9/10 year olds) onwards.

For the purposes of the present study, two branches of the school were selected. This was based on two criteria: 1) the schools were close enough to allow the researcher to make frequent visits, and 2) the researcher did not teach in those branches. It is believed that applying the second criterion contributed towards minimising the impact that the researcher could have had on the study. For details of the researcher's role at the school, see Section 3.1.4.

3.1.3 The research participants

The sample consisted of all the learners aged 7-11 and their teachers in the two branches of the school selected for the current study. This was a convenience sample (Cohen, Manion & Morrison, 2007), i.e. one to which the researcher had access. It consisted of eight teachers and 148 learners. The sample of learners consisted of 75 children aged 7-9 and 73 aged 10-11. There were 68 males and 80 females: similar numbers of each gender

in both age groups. More detailed characteristics of these participants are described in the following sections.

In order to test the methodology a pilot study was conducted before the main study. The sample included in the pilot study consisted of 2 teachers and 20 learners of whom 10 were aged 7-9 and 10 were aged 10-11. There were 12 males and 8 females in that sample. The pilot study is described and discussed separately in Section 3.2.5.

3.1.3.1 A profile of school learners aged 7-11 in Poland.

To provide the reader with a wider context, Table 3.1 below illustrates how the number of participants compared with the populations of school learners aged 7-11 in Poland. The learners who participated in the study shared many characteristics, such as socio-economic status or parental support for language education, with the larger populations in columns C and D. Hence, the insights provided by this study could be useful for teachers and researchers working in similar contexts.

Table 3.1: Number of learners aged 7-11 in the study, the country and the school contexts

Column A Studying in state education	Column B Studying English in state education	Column C Studying FL(English) in language schools	Column D Studying English in the study school	Column E Participant learners in the study
1 710 266 (GUS, 2012)	1 584 210[2] (Braunek, 2013)	474 598[3] (CBOS, 2011)	279	148

A typical participant learner had the following characteristics:

1. Attended day-time primary education where English was taught as a subject. Depending on the age and school the provision was between 90-270 min per week;
2. Learnt English as a foreign language at the language school where the study was conducted in addition to the provision described in Point 1;

---

[2] This percentage does not include Grade 6 of primary school, i.e. 12 year olds.
[3] Based on the only estimate available. No newer data were available. I contacted the Central Statistical Office of Poland (Glowny Urzad Statystyczny, GUS) to enquire about newer data but was informed that no data about the number of children in private language courses were held by them.

3. Came from a socio-economic background which enabled participation in a fee paying extracurricular English course;

4. Had parents/carers who supported them in language education, as evidenced by arranging for the child to participate in the language course;

5. Were undergoing cognitive, social and literacy development while simultaneously studying English as a foreign language.

The ethnicity of the learners who participated in the study was Polish (73%), Vietnamese (17%), French (6.7%), Russian (1.3%), Ukrainian (1.3%) and Portuguese (0.7%). All the Vietnamese students spoke both Polish and Vietnamese as L1. However, none of the remaining students spoke Polish. The school's policy was to use English only as the medium of instruction in all groups. The learners were divided by the school into groups of eight to twelve members, according to their English proficiency level and age (subgroups: 7-9 and 10-11). Table 3.2 below shows the details of the groups that participated in the study.

Table 3.2: The number and level of the participant learners

Group Code	No. of students	Age group	Level of English	
			CEFR descriptor level	No. of academic years at this level[4]
G1	12	7 - 9	A1	1
G2	9	7 - 9	A1	1
G3	11	7 - 9	A1	2
G4	12	7 - 9	A1	2
G5	11	7 - 9	A1	3
G6	12	7 - 9	A1	3
G7	8	7 - 9	A2	1
Total no. of 7-9 year olds	75	7 - 9		
G8	12	10 - 11	A1	3
G9	10	10 - 11	A2	1
G10	11	10 - 11	A2	2
G11	11	10 - 11	A2	1
G12	12	10 - 11	A2	2
G13	12	10 - 11	A2	2
G14	6	10 - 11	B1	1
Total no. of 10-11 year olds	73	10 - 11		
Total no. of learners	148			

Each group was taught by the same teacher for two 60-minute lessons per week for the whole duration of the cross-sectional phase of the study (October 2011- May 2012). The data were collected in two phases: a cross-sectional phase and a longitudinal phase (see Section 3.3). The cross-sectional data collection period lasted for twenty-eight teaching weeks, i.e. fifty-six lessons. The longitudinal phase was conducted sixteen months after the cross-sectional phase.

3.1.3.2 The participant teachers

All eight teachers who taught the groups of learners aged 7-11 in the two branches where data collection took place agreed to participate in the study. Each of them had a native-level command of English and only three spoke Polish. A typical participant teacher had the following characteristics:

1. Was a university graduate;

---

[4] One CEFR descriptor level is used by the school to describe more than one course lasting one academic year. A1.1 is the lowest level; A1.2 is the following year of study and A1.3 signifies the third year. Not all children start by taking the A1.1 course. Some children join the school with a level of English that allows them to join a higher level group. This is established during initial placement testing at registration.

2. Had professional qualifications which included at least a certificate level teaching qualification and a qualification to teach children;

3. Was likely to have a post-graduate diploma in teaching;

4. Had at least three years of experience of English language teaching and at least two years of teaching children;

5. Was likely to have significantly more experience in teaching but not necessarily in the target age group 7-11;

6. Had a minimum of one year's experience of using AfL in the target setting (TEYL with ages 7-11);

7. Taught other age groups in addition to 7-11 year olds in the school.

The characteristics of individual teachers at the beginning of the data collection (Oct 2011) are presented in Table 3.3 below.

Table 3.3: Participants: Teachers

| Teacher code | Experience (in years) | | | Age | Gender | Role at the school | First Language |
	of teaching	of TEYL	of using AfL				
T1	16	3	1	41-50	M	Teacher and coordinator	English
T2	14	7	1	41-50	M	Teacher	English
T3	5	2	1	31-40	F	Teacher	English
T4	6	4	1	21-30	M	Teacher	English
T5	18	12	1	41-50	F	Teacher and teacher trainer	Polish
T6	7	4	1	31-40	F	Teacher and coordinator	English
T7	13	7	1	31-40	M	Teacher	Polish
T8	14	8	1	61-70	M	Teacher	English

Having reported the characteristics of research participants, this chapter continues by reporting on AfL within the school where the study was conducted.

3.1.4 Assessment for Learning within the school where the study was situated

Assessment for Learning was introduced in the school thirteen months before data collection commenced. Initially, all the teachers in the school participated in three 90-minute workshops co-delivered by an established TEYL researcher, a teacher trainer[5]

---

[5] My thanks to dr Magda Szpotowicz from Warsaw University for her time and effort in preparing and delivering the AfL training sessions in September 2010.

from Warsaw University and myself. This constituted a part of the training and development program of the school and aimed to introduce teachers to AfL and equip them with a basic toolkit, which would allow them to embark on the implementation of AfL in their teaching practice. This initial training was followed by the teachers creating and implementing their individual continuous professional development (CPD) action plans and by a 90-minute swap-shop style training session at the end of the first term of implementation. Table 3.4 specifies the stages included in the process of introducing AfL at the school and, importantly, quantifies my role in that process.

As a mainstream primary specialist with experience of using AfL, I co-delivered the initial training sessions. My contribution, as a peer with experience of using AfL in a different context, was to supplement the input of the external teacher trainer. My involvement in the remaining components of the introduction of AfL was limited to instances when teachers approached me for informal advice related to their individual action plans (Table 3.4). All administrative and managerial functions relating to training and development activities and to the timetabling of teachers were fulfilled by the deputy director of the school. This did not involve me. All teachers at the school (n=32) participated in the introduction but thirteen months later only the eight, timetabled to teach 7-11 year olds, were included in the study.

Table 3.4: The process of introducing AfL into the school and the researcher's role within it (quantified).

Stage of AfL introduction	Time (per teacher)	Aim and procedure	Led by	Researcher's input
In-service training	3x 90min workshops Sept '10	To provide practical implementation ideas and a brief rationale	An experienced researcher from Warsaw University and myself	2h out of 4.5hrs
Individual action plans	Estimated: 88hrs Oct '10 – June '11 32 weeks x (2hrs teaching plus 45min preparation)	Teachers were advised to identify one or two classes to implement AfL in (2hrs per week, 31 weeks in one academic year) but not to choose the final term of exam preparation courses. This advice was similar to advice given in KMOFAP[6] (Black et al., 2003).	Teachers	ca 4hrs of informal discussion/ advice during the whole year for all 32 teachers
Follow-up workshop	1x 90min Feb '11	A swap shop of ideas and questions at the end of first term of implementation; included reflection on the individual action plans.	Teachers	No input
CPD	30min – 1hr Jul '11	The action plans constituted a part of the teachers' CPD in 2010/11	Teachers and CPD managers	No input
TOTAL Sep '10 - Jul '11	ca 95 hrs per each teacher	4 workshops (6hrs), action plans plus own teaching (88hrs) and one CPD meeting (1hr)	Mostly led by teachers themselves	ca 6hrs per 32 teachers per year

It is necessary to emphasize that the goal of this study was <u>not</u> to evaluate the impact of that training. Instead, the study aimed to investigate how the teachers understood AfL in a TEYL context having received limited training (6hrs of input). Furthermore, the current study aimed to understand how teachers implemented their understanding in practice. Finally, it intended to investigate what impact of such implementation could be observed

---

[6] King's Medway Oxfordshire Formative Assessment Project. For details please see Black et al. 2003.

on interactions that occurred in TEYL classrooms. The above account of how AfL was introduced in the study school is reported in this section to provide the reader with a comprehensive understanding of the context of the study, hence contributing to this study's validity. Furthermore, the account aims to highlight the limited role that the researcher played in the introduction of AfL at the school [ca 6/3040hrs (3040hrs = 32 teachers x 95hrs each)].

It was the introduction of AfL into a context where it had not been researched previously (TEYL) that inspired me to design and conduct the current study. Calls for studies of AfL in young learners' language classrooms have previously been made (2.3). Hence, having access to the context reported in this section presented an opportunity to design a study which would contribute new and useful knowledge to the field.

3.1.5. Ethical considerations

This study involved human participants, many of whom were children. This required the researcher to comply with a number of ethical considerations based on BERA best practice guidelines (Hammersley & Traianou, 2012). These included: obtaining informed consent, ensuring anonymity, applying ethical sensitivity and not influencing behaviours through data collection. This section outlines how each of these requirements was addressed.

Prior to the commencement of the study, written consent was sought from teachers (Appendix 4), learners' legal guardians (Appendix 5) and the school (Appendix 6). Considerations were taken to ensure that this consent was *informed* (BERA, 2011) i.e. that all participants or their legal guardians were aware of the aims of the research activities that they would be asked to participate in; the time required to complete these; the possibility to withdraw from the research at any time; that they would have access to the data; how the data would be stored; and how findings would be disseminated (Gray, 2004). The consent forms for parents were translated and distributed in English and the parents' L1s.

The second important ethical consideration was the need to ensure the anonymity of the participants. In order to do so, teachers, students and groups of students were allocated alphanumerical codes, following the pattern: Student 1 = S1, Student 2 = S2, Group 1 = G1, Teacher 1 = T1 etc. The names of the respondents and consent forms were kept

separately from the findings as per Data Protection Act 1998. This ensured that the data were secure whilst in transit. Digital data were stored on password protected hardware. Hard copies of data were kept in a locked cabinet and on completion of the project all were shredded. Obsolete back-up recordings were password protected and contained within a locked cabinet in the researcher's flat. These procedures ensured that participants could not be identified.

Finally, it was recognised that the need for demonstrating ethical sensitivity may arise in the course of conducting a study, especially where there is a conflict of interests or a dilemma to be solved (Cohen et al., 2000). Participants must not experience any type of inconvenience, harm or other negative effect during the research process (e.g. Gray, 2004). This was ensured by developing, piloting and fully reporting on the design of the study, and by behaving sensitively during interactions with the participants.

The study received ethics clearance at London Metropolitan University (where it was initially based for two years before moving to University of Reading) prior to the start of the data collection stage.

3.2 The design of the study

The research design adopted in the present study was initially based on the Colby-Kelly and Turner's (2007) study (Section 2.3.4.2) because their research questions were similar (see Appendix 7 for details of similarities and differences). This section reports on how the study design was developed. It starts with an outline of a classroom as a complex research context. This is followed by an elaboration on the methodological choices adopted to study such a context. Naturally building on the discussion of methodological considerations, the chapter continues with a discussion of how research methods for collecting data were selected. *Research method* is understood as a strategy adopted by the researcher to collect data, for example an interview. Finally, the section moves on to report how research tools for collecting data were developed and piloted. A *research tool* is understood as the prompt used for collecting data, for example, the interview schedule.

3.2.1 Classrooms as a complex research context

The social world of a classroom in this study was complex and multi-faceted with many, often uncontrollable, variables simultaneously at play. Hence, knowledge about the social

world could not be generated by arriving at testable generalisations based on empirical evidence. Instead, it had to be gained by investigating the participants' experiences in their socio-cultural context. Explanations of classroom realities required collecting *emic data*, i.e. from the people operating within the culture/school. This understanding impacted on the methodology adopted in the study.

3.2.2 Methodological considerations

Research *methodology* is understood here as 'the philosophy or the general principle which will guide research' (Dawson, 2007, p. 15). This study was of an exploratory and descriptive nature, as it aimed to collect and interpret data about the use of AfL in a context where this type of assessment had not previously been researched. Similar to other studies that have investigated YL classrooms (e.g. Enever, 2011), this research lent itself to a mixed-method design driven by a pragmatic approach (Creswell, 2007).

Interpretive rigour of the design was ensured by developing a cogent and coherent *interpretive framework*, which Guba (1990) defines as a 'basic set of beliefs that guides action' (p. 17). This was important in order to warrant the research's claims of contribution to knowledge as legitimate. Furthermore, it ensured that the findings represented the participants' experience of the studied phenomena as fully as possible. In the framework adopted here, different parts of the study fit together well. This was a warrant of the study's internal validity which enabled the researcher to collect data that represented the phenomena under study (Punch, 2004). The following section discusses how the interpretive framework of the current study ensured its validity and reliability.

As Colby-Kelly and Turner's (2007) study used content analysis, the same approach was adopted in the current study. However, it should be noted that grounded theory was also considered. The term *grounded theory* refers to a qualitative research methodology, proposed by Glaser and Strauss (1967), which aims to uncover social realities as seen by those who live them by generating theory from the data. Such theory can be further elaborated and modified through constant comparative analysis (Glaser, 1992). The methodology relies on identifying themes and collecting data until the saturation point is reached; that is, until no more new themes can be identified in the data (Strauss & Corbin, 1990). The necessity for the researcher to collect data until the saturation point is reached was an important consideration in deciding not to select grounded theory for the current study. For example, it was not possible to conduct multiple interviews with participating

teachers. Hence, the dataset which could feasibly be collected might not be appropriate for the requirements of conducting grounded theory research.

3.2.3 Validity and Reliability Considerations

Ensuring the validity of the adopted mixed-method design, with a larger qualitative component, was a central consideration in the methodology. As Maxwell and Mittapalli (2010) rightly noticed, while quantitative and qualitative traditions tend to be in disagreement about what constitutes validity, they agree that it is connected with the procedures of collecting data and drawing inferences from the data. More specifically, *internal validity* 'seeks to demonstrate that the explanations of a particular event, issue or set of data which a piece of research provides can actually be sustained by the data' (Cohen et al. 2007, p. 135) while *external validity* 'refers to the degree to which the results can be generalised to the wider population, cases or situations' (p. 137). As Somekh and Lewin (2005) argue, mixed-method research offers unique opportunities for increasing the validity of a study through:

1. *triangulation* (seeking corroboration and convergence of results),
2. *complementarity* (gaining insights into different facets of phenomena),
3. *development* (using the outcomes of one method to inform another),
4. *expansion* (adding depth and scope to the study).

'Triangulation is characterised by a multi-method approach to a problem in contrast to a single-method approach' (Cohen et al., 2007, p. 142). This understanding of triangulation is consistent with what Denzin (1996) refers to as *methodological triangulation*. In the current study data were triangulated through different sources and from different participants to validate the findings and interpretations. To ensure *time triangulation* by considering change over time (Denzin, 1996), a longitudinal aspect was incorporated in this largely cross-sectional study. This was accomplished by conducting a delayed teacher questionnaire, (sixteen months after completion of the data collection stage). Complementarity was warranted by studying the same phenomenon through quantitative as well as qualitative data. For example, data on the use of AfL comprised sections of lesson observations (qualitative) as well as school documents called Records of Work Done (ROWDs; quantitative). With reference to development, the mixed-method design added coherence to the study design by allowing the use of draft findings from one

research method (for example, teacher interviews) to designing another (for example, teacher focus group protocol). Finally, expansion was ensured. This included, for instance, the quantitative analysis of lesson observation data to draw inferences about the amount and type of classroom interactions across twenty-eight lessons (scope), while qualitative discourse analysis provided in-depth insights into the content of classroom interactions (depth).

Miles and Huberman (1994) argue that in qualitative research the researcher is the key instrument in collecting and interpreting data. In order to ensure the validity of the qualitative component, it was important to deploy a reflective approach to the role of the researcher in the study. This required careful consideration of the impact which the researcher had on the research participants and findings. Attention was paid during the data collection stage not to change or in any other way influence teachers' opinions or behaviours connected with AfL. This was addressed by:

- developing open-ended, non-leading questions;
- piloting and refining the interview and focus group protocols;
- not providing comments to participants by the researcher after the interviews, focus group or lesson observations.

The programme of events may also have posed a threat to the internal validity of the study. The data collection stage lasted for eight months (excluding the delayed questionnaire). For that reason, the impact that the events which occurred in that time might have had on the validity of the study was another factor carefully taken into account. The two threats to the internal validity of the study posted by time events and actions taken to minimise them are summarised in Table 3.5.

Table 3.5: Threats to internal validity and action taken to minimise them.

Threat to validity	Reason why this was considered a threat	Action taken to minimise the threat
Professional development activities	The teachers may have participated in various forms of professional development during the duration of the study.	To include that as a possible factor in the data analysis all teachers were asked to complete the last row on the demographic form in June 2012 (see Appendix 8).
Learning from research tools	The teachers were observed by the researcher more than once during the study (2-4 times each). This posed the risk that they might adapt their practice as an outcome of lesson observations.	To prevent that, no feedback, formal or informal, was given to teachers following the lesson observations. Neither were they given access to the video recordings of the observed lessons.

In addition to ensuring validity, the development of the study design also considered issues connected to ensuring the reliability of the findings. *Reliability* is a term associated with the quantitative tradition. For a research study to be reliable 'it must demonstrate that if it were to be carried out on a similar group in a similar context (however defined), then similar results would be found' (Cohen et al. 2007, p. 146). In the qualitative tradition, the term is often contested (e.g. Winter, 2000) and replaced with terms such as dependability (Lincoln & Guba, 1985) or trustworthiness (Shenton, 2004). In the current study, care was taken to ensure reliability through piloting and comprehensively reporting on the processes of data collection, coding and analysis. Trustworthiness and dependability of the current study were ensured by reporting thoroughly on the context of the study, characteristics of the participants as well as by acknowledging the limitations of the design (see Section 3.4). Having discussed the interpretive framework, this chapter continues by outlining how that framework was implemented in practice.

3.2.4 Data collection methods and tools

This section reports on the process of selecting research methods and developing research tools. Decisions connected with qualitative methods initially posed a challenge because, as Denzin (1996) rightly notices, 'an embarrassment of choices now characterizes the field of qualitative research. Researchers have never before had so many paradigms, strategies of inquiry, and methods of analysis to draw upon and utilize' (p. 135). In order to remain loyal to the pragmatic methodological approach of this study, the main focus was on selecting research methods which would provide the most useful insights in the

areas explored by the research questions. As the RQs were largely based on the Colby-Kelly and Turner (2007) study, similar methods of data analysis were selected. However, given two major differences between the two studies, i.e. the educational contexts and the learners' age group (Appendix 7), research tools had to be developed for the purpose of this particular study. Table 3.6 provides an overview of the research methods and data collection tools used in the study for each of the three RQs. The following section describes how each data collection tool was developed.

Table 3.6: Research questions, methods and data collection tools

Research question	Method	Data collection tool
1: How do teachers understand AfL after receiving a limited amount of training and being encouraged to use AfL techniques for at least one academic year when teaching English to young learners aged 7-11?	Semi-structured Interview	Teacher interview schedule (Appendix 9)
	Focus Group	Focus group prompts (Appendix 10)
2.1: How do teachers translate their understanding of AfL into classroom practice in a TEYL context with students aged 7-11 in a private language school in Poland?	Scrutiny of Curriculum Documents	Records of Work Done[7] i.e. ROWDs (Appendix 11)
	Lesson Observations	Lesson Observation Schedule and observation notes, Part One (Appendix 12)
2.2: Do teachers report any changes in their practice of using AfL over time?	Delayed Teacher Questionnaire	Questionnaire design (Appendix 13)
3: What is the observable impact of AfL on classroom interactions in a TEYL context?	Lesson Observations	Lesson Observation Schedule and observation notes, Part Two (Appendix 12)

3.2.4.1 Collecting Data from Research Participants

As in the Colby-Kelly and Turner (2007) study, the current study used interviews to collect data from the teachers. Additionally, the current study incorporated a focus group discussion and a delayed questionnaire. As discussed earlier, all research tools were developed for the purpose of this study. The subsections below report the rationale for

---

[7] Record of Work Done (ROWD) is a name used for the documents by the school. For the description of ROWDs see Section 3.3.2.

selecting each research method and the purpose for deploying it as well as the processes of developing the research tools.

3.2.4.1.1 Teacher Interviews

Interview as a research method enables participants 'to discuss their interpretations of the world in which they live and to express how they regard situations from their point of view' (Cohen et al., 2007, p. 349). For that reasons, interviews were identified as a useful research method to investigate the teachers' own understanding of AfL. This choice was similar to other studies which have investigated teachers' beliefs about self- and peer-assessment (Dixon, Hawe & Parr, 2011) or teachers' interpretations of the effectiveness of assessment (Butler & Lee, 2010) and the implementation of AfL (Colby-Kelly & Turner, 2007; Lee & Coniam, 2013, Rea-Dickins, 2001) and assessment in TEYL contexts (Edelenbos & Kubanek-German, 2004). Table 3.7 below summarises studies which used teacher interviews for purposes similar to those in the current study.

Table 3.7: Studies which used teacher interviews

Study	Focus of the study	Focus of teacher interviews
Edelenbos and Kubanek-German, 2004	Classroom assessment of young language learners	To investigate 'diagnostic activities' (p. 264) used by YL teachers
Rea-Dickins, 2001	Formative assessment in an EAL context	To 'provide evidence for a range of strategies in relation to the implementation of classroom assessment and, in particular, formative assessment' (p. 432)
Lee and Coniam, 2013	Assessment for Learning in Secondary School L2 (English)	To investigate 'the implementation of AfL in writing' (p. 38)
Butler and Lee, 2010	Self-assessment of L2 (English) in a Primary School	To 'understand the teachers' observations and perceptions towards the effectiveness of self-assessment and to elicit the teachers' insights regarding how self-assessment may be implemented effectively in their respective teaching environments.'(p. 14)
Dixon, Hawe and Parr, 2011 and Hawe and Dixon, 2014	'Teachers' espoused beliefs about self- and peer- assessment and their congruence with practice; (p. 365) in New Zealand state schools	To investigate teachers' beliefs about learners conducting peer and self-assessment
Colby-Kelly and Turner, 2007	Assessment for Learning in an EAP context	To investigate teachers' views on using AfL in speaking activities

Deploying the interview as a research method had some limitations. First, data which were collected were teacher's subjective interpretations of classroom realities. Hence, that data provided insights into teachers' understanding about the implementation and impact of AfL. Secondly, the one-to-one nature of the interviews may have posed a risk of some teachers saying what they thought was expected of them rather than sharing their beliefs. To minimise that risk, as the researcher I ensured that I built a positive relationship with each teacher before the interviews were scheduled. It is believed that this contributed to teachers feeling more relaxed and willing to share their honest opinions with the researcher during the interviews. Moreover, at the beginning of each interview, I briefed

each interviewee about their anonymity and the purposes of the interviews, drawing attention to the fact that the research was focusing on gathering data about the teachers' understanding of AfL and was not testing their knowledge about it (Appendix 9). Finally, if teachers were not able to readily respond to the prompts included in the schedule, allowances were made for them to come back to any topic at a later time during the interview.

The process of designing a protocol entailed a number of steps specified by Cohen et al. (2007). First, the purpose for conducting interviews was defined. In the current study, the purpose was to gather information with a direct bearing on the research questions (Tuckman & Harper, 2012). This purpose was then translated into the specific aims of the interviews, which was to collect data illustrating the teachers' beliefs about:

- what AfL means in a TEYL context;
- how they implemented AfL in TEYL classrooms;
- the impact of using AfL on interactions within TEYL classrooms.

The next steps involved designing a semi-structured schedule using a standardised open-ended approach (Patton, 1990) by summarising the outcomes of the literature review, the preliminary outcomes of lesson observations from the pilot study (n=4) and mapping them out against the interview objectives. This resulted in preparing thirteen open-ended questions. The wording of the prompts took into account the profiles of the teachers. The interviews took place between January and March 2012. The procedures used to analyse the data obtained through the interviews are discussed in Section 3.3.1. The following section reports on how the focus group schedule was developed.

3.2.4.1.2 Focus Groups

Focus groups are a useful research method which can be used to triangulate findings from other sources (Somekh & Lewin, 2005). In the current study, a focus group with eight teachers was organised to verify the draft findings from individual interviews. The aim was to validate the researcher's interpretations of the interview data and to gain more in-depth insights into the teachers' understanding of AfL. Moreover, the purpose of using a focus group as a research method was to elicit collective (not individual) beliefs through the data that emerged from the interactions within the group (Morgan, 1988).

As Cohen et al. (2007) rightly notice, despite the claims that 'its potential is considerable, the focus group, as a particular kind of group interviewing, still has to find its way into educational circles to the extent that it has in other areas of life' (p. 377). Adopting the focus group as a method to verify and expand findings from individual interviews contributed to the innovative nature of this study. It was influenced by Lee and Coniam's (2013) study that used this method with adolescents to investigate-learners' perspectives on how implementing AfL impacted on their motivation and writing performance.

Prior to data collection, the researcher had been trained in how to conduct focus groups and was therefore aware of the challenges involved in facilitating a focus group, such as managing the group dynamics in order to ensure participation and the coverage of relevant topics.

The schedule for the focus group (Appendix 10) was developed through a process similar to that adopted for developing the interview schedule. The formulation of the questions was informed by draft findings from the analysis of the interview transcripts. The focus group discussion with all eight teachers took place in May 2012. The process of the analysis of the data collected through the focus group is reported in Section 3.3.1.

3.2.4.1.3 Teacher Questionnaire

Questionnaires as a research tool are relatively easy and time efficient to administer and allow for collecting quantitative as well as qualitative data (Somekh & Lewin 2005). Questionnaires are often used in educational research: most often to investigate attitudes, motivation and perceptions. Studies of AfL, which deployed questionnaires for such purposes, including those based in TEYL contexts, collected data from learners (Butler & Lee, 2006; Enever, 2011; Gattullo, 2000, Lee & Coniam, 2013,) and teachers (Colby-Kelly & Turner, 2007, Rea-Dickins & Gardner, 2000). Table 3.8 below provides details of the purposes and designs of questionnaires used to collect data from teachers. As evident from Table 3.8, in studies of classroom assessment, questionnaires were deployed with small samples of teachers, similar to the size of the sample in the current study.

Table 3.8: Studies about AfL, which gather data from teachers through questionnaires

Study	Focus of the questionnaire	Number of respondents	Design
Colby-Kelly & Turner, 2007	'Teachers' perceptions about classroom based assessment' (p.19)	9	4 point Likert-type scale
Rea-Dickins & Gardner, 2000	Teachers' perceptions of the assessment of EAL students	9	Open-ended

In the current study, a questionnaire was selected as a method which allowed obtaining a longitudinal perspective and ensuring time triangulation. A delayed teacher questionnaire (Appendix 13) was administered sixteen months after the cross-sectional data collection stage ended. The aim of the questionnaire was to gather data about changes in the implementation of AfL over time.

The design incorporated a Likert-type scale and open-ended questions. The scale, ranging from 1 (*never*) or (*almost never*) to *5* (*every lesson*) or (*almost every lesson*), was deployed. A list of all AfL techniques identified in the data thus far was provided. The teachers were asked to indicate how often they used each of the techniques. Additionally, space was provided for recording techniques that teachers used but which had not been included in the list. This was followed by a number of open-ended questions which aimed to gain more nuanced insights into the developments that occurred over time. The delayed questionnaire was piloted in September 2013 and administered in October 2013.

This section has discussed the research methods used for collecting data from teachers. The rationale for selecting an interview, a focus group and a delayed questionnaire as data collection methods has been provided. This section has also reported how the respective research tools were developed. The next section reports on how the lesson observation schedule was developed.

3.2.4.2 Lesson Observations

Lesson observation is a useful research method that allows the complexities involved in teaching and learning to be captured. Cohen et al. (2007) argued that 'the distinctive feature of observation as a research process is that it offers the investigator an opportunity to gather 'live' data from naturally occurring situations' (p. 396). Deploying this method

enabled the researcher to gather authentic data in situ, with direct relevance to RQs 2.1 and 3. It is believed that collecting such *authentic data*, i.e. data not mediated by others (Cohen et al., 2007), contributed to ensuring that the findings of the study reliably represented the phenomena studied.

In the current study, lesson observations were used primarily to collect evidence about how teachers implemented their understanding of AfL (similar to Dixon et al., 2011). The lessons were either video or audio recorded and the transcripts enabled the researcher to document examples of classroom interactions, which was similar to the studies conducted by Gattullo (2000) and Rea-Dickins (2006). Table 3.9 below, reports studies that used lesson observation as a research method in similar contexts and/or for similar research foci.

Table 3.9: Studies which used lesson observation as a research method

Study	Context of the study	Focus of classroom observations
Enever, 2011	Young language learners' progression in L2 (mostly English, also French, Spanish) over three years: 7-11 year olds	'linguistic and non-linguistic behaviour during FL lessons focusing on learners' attention, participation, relationship to teacher and classmates, language comprehension, and production.' (Nikolov & Mihaljević Djigunović, 2011, p. 99)
Rea-Dickins, 2006	Formative assessment in an EAL context in a primary school: 6-7 year olds	Formative and summative assessment and feedback evident through teacher and learner interactions during the activities.
Rea-Dickins and Gardner, 2000	Formative assessment in an EAL context of 5-7 year olds	'To probe in depth key issues of English language assessment' (p. 219)
Gattullo, 2000	Formative assessment in TEYL classrooms with 8-10 year olds	To record assessment events during lessons and transcribe examples of interaction during those events.
Dixon, Hawe and Parr, 2011	'Teachers' espoused beliefs about self- and peer- assessment and their congruence with practice' (p. 365) in state schools in New Zealand	To collect evidence about how teachers' beliefs were enacted in practice through teachers' feedback practices in AfL in writing oriented Literacy lessons
Colby-Kelly and Turner, 2007	Assessment for Learning in EAP context	'To catalogue assessment episodes' (p. 19), including their origin, focus, duration and language skills.

While offering the advantages of gathering authentic empirical data, classroom observations are sensitive to the researcher as a mediator of the data (Cohen et al., 2007). This was carefully considered in the process of designing the tool (described below) and analysing the data. Field notes were made during the observed lessons. Subsequently, more detail was added when the video recordings were reviewed. This ensured the reliability of observations and enabled the researcher to observe *parallel forms* (events that happened simultaneously) hence addressing the issue of the selective attention of the observer (Cohen et al., 2007, p. 410).

While developing the observation protocol, particular attention was paid to ensuring that it gathered data relevant to the research questions. Having reviewed observation schedules reported in the literature, it was decided that there was not one that would satisfy the above criterion of relevance. For that reason, a lesson observation protocol was developed for the purposes of this study. The design process was informed by similar choices made by other researchers, as discussed below, and by the outcomes of the literature review (Chapter 2).

The observation schedule was divided into two parts. The first one was designed to catalogue the use of AfL techniques. This was achieved by recording duration and the brief description and, where appropriate, names of consecutive tasks in each lesson, annotating them with the language skills that they focused on and AfL techniques used. Additional detail, related to the perceived purposes for deploying each AfL technique, was added when the researcher reviewed the video and audio recordings[8]. Part Two of the lessons observation schedule aimed to collect data about the impact of AfL on interactions. This entailed documenting behaviours similar to those investigated in the ELLiE study (Nikolov & Mihaljević Djigunović, 2011). Hence, the design of the schedule was based on the ELLiE classroom observations schedule (Enever, 2014, personal communication). This included using similar codes to record the type of interactions as in the ELLiE schedule: T-C (teacher – whole class), L-C (learner – whole class), T-xLL (teacher – group of students, e.g. T-4LL = teacher – group of 4 learners), T-1L (teacher – individual learner), IND (individual work), L-L (Pairs), and LL (groups).

---

[8] Out of the total of 28 lesson observations, 26 were video recorded. The remaining two were voice recorded due to parental preference not to video record their children as expressed in the process of gaining informed consent.

The adopted unit for recording data during observations was *a task*. There have been multiple discussions of the definition of a task (e.g. Ellis, 2000; Skehan, 1996, 1998). Ellis (ibid.) provides a useful argument that a task may entail different concepts when defined either from a research or pedagogical perspective: i.e. '(r)esearchers, for example, may view a task in terms of a set of variables that impact on performance and language acquisition whereas teachers see it as a unit of work in an overall scheme of work' (p. 194). Skehan (1998) lists four characteristics of a task: '1. meaning is primary; 2. there is a goal which needs to be worked towards; 3. the activity is outcome-evaluated; 4. there is a real world relationship' (p. 268).

However, it should be noted that in the complex context of a classroom, these clear-cut characteristics can, at best, account only for how tasks are planned. They do not take into account individual characteristic, interests and abilities which learners bring to the process of performing the task. Furthermore, it is important to note at this point that the phrase 'real world relationship', which is frequently used in connection with communicative teaching, is problematic. It indicates that tasks in language lessons should attempt to somehow recreate situations, which learners can encounter in the world outside the classroom (Nunan, 1987). Thus, a classroom is not considered a real world situation. However, Walsh (2006) has convincingly argued that participating in a lesson is a real world situation for both learners and teachers. This seems even more applicable to children for whom going to school constitutes a significant part of their lives.

Adopting Walsh's (ibid.) view of a classroom as a real world situation has implications for the current study. Specifically, it means that any task will have a relationship with the real world; for instance, controlled practice of a grammar point could be an authentic activity in a lesson. Hence, Skehan's (1998) Criterion 4 becomes unhelpful and as such was not included in the analysis. However, the first three of Skehan's (1998) criteria were regarded as useful and, consequently, were adopted in the definition of 'task' in the current study. In order to be true to the context of the study, it seemed necessary to supplement this largely context-free definition with context specific considerations. Therefore, lesson observation notes were compared with the accounts of the same lessons collected from teachers through ROWDs (3.3.2). This allowed me as the researcher to compare my judgements of what constituted as a task in the observed lessons with the

records made by teachers who planned and delivered the lessons. This approach ensured that systematic data were collected.

The adopted semi-structured approach ensured a clear description of the types of events to record while maintaining the flexibility to gather a rich collection of data. This contributed to ensuring internal validity of the study. For an example of a completed lesson observation schedule, please see Appendix 14.

All data collection tools described above were piloted. The next section reports on the pilot study.

3.2.5 Pilot study

The purpose of the pilot study was to trial the research tools and methods of analysis. The aim was to refine them where necessary in order to ensure a good fit between the study design and the research questions.

The pilot study lasted for twelve weeks (March –June 2011) and was conducted in a branch of the school which was not planned for inclusion in the main study. The data collected in the pilot study included: audio recordings of two teacher interviews, ROWDs from forty-eight lessons, field notes and video recordings from four lesson observations. The participants were twenty students and two teachers, none of whom were included in the main study. This decision was made in order to minimise the possibility of them learning from the research tools and thus helped to ensure greater validity of the findings (Black, 2005). The delayed teacher questionnaire was piloted in September 2013 with the same two teachers.

Piloting the study design resulted in making five refinements. These are summarised in Table 3.10 below.

Table 3.10: Refinements to the research design following the piloting stage

Refinements regarding:	Details of the refinement	Reasons
Research questions	RQ2 (What are parents' and students' beliefs about AfL?) was dropped	Following the piloting stage and the confirmation process at the university, this question was dropped due the restrictions of the scope of a PhD research and the size of the thesis.
Research questions	RQ1 (What is the nature of AfL in a TEYL context?) was divided into two RQs: RQ1 focused on the teachers' understanding of AfL in a TEYL context and RQ2 (2.1 and 2.2) on how they implemented that understanding in the classroom.	Following the piloting stage and the confirmation process at the university, RQ1 was adapted to provide a more specific focus.
Interview schedule	The wording of the introduction to the teacher interviews was expanded to include the information that the draft interpretations would be shared with the teachers during the focus group discussion.	These changes were introduced in response to the teachers' questions for clarification.
Lesson Observations	The decision was made to video record the observed lessons.	To enable more in-depth analysis and contribute to ensuring the reliability of the observations.
Consent forms	The consent form for teachers, parents/guardians and the school were adapted to include information about the video recording.	As a result of the decision to video record the observed lessons.

The aim of the data analysis in the pilot study was to trial the methods of data collection, coding and analysis which were planned for the main study. The same data analysis methods were used in the main study. These are reported in the following section.

3.3 Data sets and analysis

The process of reaching the interpretive conclusions reported in Chapter 4 included analysing a substantial set of qualitative and quantitative data. The data were collected in two phases. Phase One was cross-sectional and included data from teacher interviews,

lesson observations and the focus group discussion in academic year 2011/12. Phase Two was longitudinal. It was completed in October 2013. The overview of data collection is provided in Table 3.11 below.

Table 3.11: Timeline of the data collection process

Research method	Phase One 2011/12								Phase Two 2013/14
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Oct
RoWDs	X	X	X	X	X	X	X	X	
Interviews				X n=3	X n=4	X n=1			
Lesson observations			X n=6	X n=6	X n=5	X n=6	X n=5		
Focus group								X n=1	
Teacher questionnaire									X n=8

Before discussing the procedures used for the analysis, a summary of all the data sets is provided in Table 3.12 below.

Table 3.12: Data sets in the main study

Research question	Method	Data sets obtained
1: How do teachers understand AfL after receiving a limited amount of training and being encouraged to use AfL techniques for at least one academic year when teaching English to young learners aged 7-11?	Semi-structured Interview	8 audio recordings 8 transcripts
	Focus Group	1 audio recording 1 transcript
2.1: How do teachers translate their understanding of AfL into classroom practice in a TEYL context with students aged 7-11 in a private language school in Poland?	Curriculum Documents	448 lessons reported by teachers in ROWDs
	Lesson Observations	28 lesson observation field notes (Part One of the schedule) 26 video recordings of observed lessons 2 audio recordings of observed lessons (both in T8's classes)
2.2: Do teachers report any changes in their practice of using AfL over time?	Delayed Teacher Questionnaire	8 completed teacher questionnaires
3: What is the observable impact of AfL on classroom interactions in a TEYL context?	Lesson Observations	28 lesson observation field notes (Part Two of the schedule) 26 video recordings of observed lessons 2 audio recordings of observed lessons (both in T8's classes)

The sections below report how each data set was analysed.

3.3.1 The procedures for the analysis of the data from Teacher Interviews and the Focus Group

Content analysis was used to analyse the data obtained from eight teacher interviews and one focus group. Hsieh and Shannon (2005) define content analysis as 'a research method for the subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns' (p.1278). Applying this method entailed transcribing and coding the spoken discourse from the teacher

interviews and the focus group. The methodological choices related to the processes of transcribing and coding are discussed below.

3.3.1.1 Transcribing

Transcribing the interview and the focus group data was an important stage in the data analysis process as it resulted in obtaining a written record, which was subsequently coded to identify themes. It should be noted that transcribing the spoken text resulted in the loss of some data (Cohen et al., 2007) because the transcripts did not include non-verbal communication and other contextual factors. The data which were recorded were verbatim records of what the teachers said. The transcripts included the following:

1. What was being said
2. What was emphasised, which was denoted by capitalising the emphasised phrases, e.g. *I REALLY think that*
3. The occurrence of hesitation according to the number of seconds, denoted in parenthesis, e.g. *I think that (2) that students find it helpful.*
4. Unintelligible speech by writing down *<unintelligible>*

The transcribing convention was based on Walsh (2006) and can be reviewed in Appendix 15. The resulting transcripts were coded following the procedure reported in Section 3.3.1.2 below.

3.3.1.2 The coding process

The coding process was completed in two stages. The initial coding included proofreading the text and highlighting key phrases 'because they make some as yet inchoate sense' (Sandelowski, 1995, p. 373). The technique adopted for Stage One was to highlight recurring text. This choice was justified by D'Andrade's (1991) claim that, when interviewed about a topic, 'frequently people circle through the same network of ideas' (p. 287) and by Ryan and Bernard's (2003) claims that 'themes are only visible (and thus discoverable) through the manifestation of expressions in data' (p. 86). The term 'expressions' is referred to here as 'thematic units' in accordance with Krippendorf's (1980) terminology. Such thematic units were initially coded descriptively and subsequently, as the analysis progressed, some codes merged or were changed (Cohen et al., 2007) and alphanumerical acronyms were developed. Careful attention was paid to ensure that the coding categories proved exhaustive and mutually exclusive (Cohen et al.

2007). This process entailed multiple reading, following Bogdan and Biklen's (1992) suggestion that a text should be read a minimum of two times.

The second phase of coding was completed using a cutting and sorting procedure (Ryan & Bernard, 2003). This entailed printing the coded text obtained from Stage One and organising it into piles: each one relating to a similar theme. Six themes were identified. Once organised, the units relating to one theme were glued onto an A3 sheet of paper (Appendix 16) to facilitate storing and summarising. The findings from this analysis are reported in Chapter 4, Part One.

The validity and reliability of the coding process were carefully considered. Ryan and Bernard (2003) suggest the following strategies for ensuring the validity and reliability of coding:

- making judgements explicit in reporting the study;
- ensuring agreement across the coders;
- verifying this with the respondents.

All the above strategies were deployed in the current study. To satisfy the first criterion, the definitions of all emergent themes identified in the study are provided in Table 3.13 below.

Table 3.13: Emergent themes in coding the interview and focus group data

Theme code	Definition of the theme
INTAFLTL	Data concerning the teachers' beliefs about the integration of AfL with teaching and learning processes in TEYL classrooms.
COMPTM	Data concerning the teachers' descriptions of whether and how AfL was compatible with TEYL methods.
AFLTECH	Data concerning accounts of how the teachers implemented AfL techniques.
PURUSE	Data concerning descriptions of the purposes of using AfL in TEYL classrooms.
IMPIPG	Data concerning accounts of the teachers' beliefs about the impact of AfL on sustaining individual, pair and group work
IMPINTER	Data concerning the teachers' reports of the impact of AfL on interactions in TEYL classrooms

Secondly, an inter-rater was invited to take part in the coding. This was an experienced teacher, who also had some research experience, and coded one of the interview transcripts so as to test the researcher's coding by comparison. (Based on the word count, this amounted to 10% of the data.) The teacher had an MRes degree from a UK university. As a part of his/her studies for that degree, the teacher had completed the university's Researcher Training Programme and had conducted a small scale study in which (s)he analysed data from semi-structured interviews with parents. Following Cohen et al. (2007) an inter-rater reliability coefficient was calculated using SPSS v19. Codes were used as cases and researcher and inter-rater as variables. The inter-rater coefficient calculated in this study was .918**, which was statistically significant at the level of 0.01. Shaughnessy, Zechmeister and Zechmeister (2003) suggest that an inter-rater coefficient of > 0.9 should be obtained. That requirement was satisfied.

The third strategy for ensuring the validity and reliability of the interpretations entailed verifying draft findings from interviews through the focus group with teachers. It is believed that applying all three strategies resulted in a set of themes which reliably represented the phenomena studied.

3.3.2 The procedures used to analyse the data from Records of Work Done

Records of Work Done was the name of the documents used by the study school to record what was actually done in lessons. They were not prescriptive or planning documents. Towards the end of each lesson or soon afterwards, the teacher recorded the learning objectives, activities and assessment methods used in that lesson in a ROWD. Completing ROWDs was compulsory for all teachers in the school. As such, the documents provided a valuable source of data about what teachers believed had happened in the classroom. The data set included ROWDs from fifty-six lessons per teacher: a total of 448 lessons. For an example of a ROWD, see Appendix 11.

While the Colby-Kelly and Turner (2007) study included an analysis of school documentation, it drew on different types of documents to those in the current research. Colby-Kelly and Turner (ibid.) used curriculum documents, which contained guidance on AfL. In the current study, AfL was not explicitly mentioned in curriculum documents at the school. Instead, ROWDs (Appendix 11) were included in the data set.

ROWDs were analysed largely quantitatively by calculating frequency counts of techniques and deploying descriptive statistics in SPSS v19, including mean, median, standard deviation, maximum and minimum values. The aim was to investigate how often teachers used AfL across and within lessons and how many different types of techniques they implemented. The frequency counts were grouped by age (7-9 and 10-11) and by teacher to enable comparisons. These findings were visually presented as a sliding scale showing the use of AfL.

The findings from analysing ROWDs were triangulated with the findings from lesson observations. The procedure deployed to analyse field notes from lesson observations is reported in the following section.

3.3.3 The procedures used to analyse the data from Lesson Observations

The data set included field notes from twenty-eight lesson observations. Twenty-six of them were video recorded. But the other two lessons (both in the younger age group taught by T8) were audio recorded due to the lack of parental permission for video recording for one child in that group. Permission for audio recording was fully granted. The method of analysing the field notes is reported in the two sections below. In the first section the analysis of the data relating to RQ2.1 (use of AfL) is reported. The second section explains how the data relating to the impact of using AfL on classroom interactions were analysed.

3.3.3.1 Use of AfL techniques in TEYL classrooms

The data obtained from Part One of the lesson observation schedule included a chronological list of tasks annotated with AfL techniques and language skills. They were also analysed with SPSS v19. Descriptive statistics were used to calculate the median, mean and standard deviation of the number of AfL techniques used to investigate the frequency of using AfL in the lessons. The analysis of this data was used to generate findings for RQ2.1 (use of AfL). The findings were compared with the similar analyses carried out on ROWDs. An inventory of AfL techniques observed in the lessons was also created (Appendix 18). The analysis proceeded with annotating each technique with the purpose, language skills and the time of the lesson at which it was used. It was hoped that the analysis would provide some insights into the frequency (how often) and diversity (range of technique type) of using AfL techniques in TEYL classrooms.

3.3.3.2 The impact of AfL on interactions

Field notes recorded in Part Two of the observation schedule provided data about the types of interactions. This section reports how the interactions were analysed first quantitatively (3.3.3.2.1) and then qualitatively, by applying the Storch's (2002) model and deploying the Variable Approach to FL classroom interactions (3.3.3.2.2).

3.3.3.2.1 The procedures of quantitative data analysis from classroom interactions

Data representing the types of interactions that occurred in the lessons were first used to create scatter plots and subsequently to calculate a set of bivariate correlations in SPSS v19. There were fifty-six cases (1 case = 1 lesson). The two variables included the number of AfL techniques used in the lesson (code: NAFL) and the number of interactions of each type (e.g. code: NT1L – the number of T-1L interactions). In total twenty-one correlations were calculated. This included correlations between *NAFL* and each of the seven modes of interaction (3.2.4.3) for the whole cohort and separately for the two age groups of students (7-9 and 10-11). The purpose of this was to provide empirical evidence about the relationship between the frequency of using AfL and the types of interactions that occurred in the lessons. The statistically significant findings are reported in Chapter 4.

3.3.3.2.2 The procedures of analysing classroom discourse

A number of extracts of classroom discourse were transcribed from the video recorded lessons to provide a qualitative insight into conversations that took place between teachers, learners and peers while using AfL techniques. The transcribed extracts were the ones that were intelligible in the video recordings. Nine of them were selected for reporting in this thesis to represent the two interaction patterns (L-L and T-1L) which the outcomes of the quantitative analysis of lesson observation showed to be correlated with the use of AfL. Five of the extracts were sourced from the younger age group (7-9) and four from the older one (10-11). The aim was to gain insights into whether the interactions could facilitate learning. This was done by deploying two types of qualitative analysis. First, holistic interaction patterns were analysed by applying Storch's (2002) model. Subsequently, following Walsh (2006), classroom interactions were analysed using the Variable Approach. The processes involved in conducting each type of analysis are reported in the following two subsections.

A. **Applying Storch's (2002) model**

According to Butler and Zeng (2014), Storch's (2002) model of dyadic interaction is most frequently used to analyse interactions in SLA. It is concerned with the levels of *equality* and *mutuality* between the interlocutors. Mutuality is defined by Storch (2002) as 'the level of engagement with each other's contribution' (p. 127) while equality is defined as 'the degree of control or authority over the direction of the task' (p. 127). Based on these binary features, the interactions are classified in four quadrants. The model is presented in Figure 3.1 and discussed below.

	Low ← EQUALITY → High	
High ↑ **MUTUALITY** ↓ **Low**	Quadrant 4: Expert/novice	Quadrant 1: Collaborative
	Quadrant 3: Dominant/passive	Quadrant 2: Cooperative or dominant/dominant

Figure 3.1: Storch's (2002) model of dyadic interaction patterns

Quadrant 1 (*collaborative*) is characterised by medium to high levels of equality and mutuality. This means that both interlocutors equally control what happens during the interaction and respond to each other by asking questions, acknowledging contributions and negotiating solutions. Interactions in Quadrant 2 also have a medium to high level of equality but a low level of mutuality. Storch (2002) subdivided this quadrant into two patterns. In the first one, both interlocutors try to gain control of the interaction (*dominant/dominant*). This can lead to a break down in the interaction. In the second pattern, called by Storch (ibid.) *cooperative*, 'parallel participation [can be] observed during the task' (Butler & Zeng, 2014, p. 49). Its name seems similar to the one used for Quadrant 1. However, there is a notable difference the collaborative (Quadrant 1) and cooperative (Quadrant 2) patterns. During collaborative interactions, interlocutors work towards a solution together, i.e. 'alternative views are presented and discussed, and agreed-upon solutions are reached through negotiation' (Butler & Zeng, 2014, p. 48) while in the cooperative pattern 'both parties contribute equally but without engaging with each other's contribution' (p. 49). As Butler and Zeng (ibid.) convincingly argue, the two names can easily be confused. Hence, following Butler and Zeng's (2014)

suggestion, in the current study the term *passive parallel* (p. 55) has been adopted instead of *cooperative*.

The two patterns described so far, are characterised by medium to high equality. In the remaining two quadrants the interlocutors are not equal. In Quadrant 3 (*dominant/passive*), one adopts an 'authoritarian' role while the other adopts a 'subservient' role (Stroch, 2002, p. 129). In that pattern, most of the discourse is produced by the dominant person who does not tend to ask questions or acknowledge the views of their interlocutor. Finally, in Quadrant 4 (*expert/novice*), despite low equality, interlocutors contribute to the conversation, negotiate solutions and respond to each other. This means that despite being in control of the conversation, the 'expert' tries to engage the 'novice' and facilitates their contribution.

While offering an informative and efficient way to categorise dyadic interactions, Storch (2009) recognises that this model is not ideal. This is largely because both bilateral values (equality and mutuality) are continuous and somewhat artificially split into two sections of the continuum. Nevertheless, it has been used by other researchers to analyse holistic patterns of dyadic interactions (e.g. Butler and Zeng, 2014). Importantly, patterns in Quadrants 1 and 4 have been shown by other researchers to facilitate learning, especially by increasing uptake (Kim & McDonough, 2008; Ohta, 1995). As the current study was interested in gaining insights into how AfL could contribute to raising achievement, the Storch's (2002) model was considered a helpful framework for the analysis of classroom interaction patterns.

Once the holistic patterns were investigated, the analysis continued by investigating how the interactions that occurred during the use of AfL facilitated meeting the pedagogical aims of the lessons. The details of that process are described in the next section.

B. **Variable Approach to investigating L2 classroom interactions**

The Variable Approach to investigating L2 classroom interactions (Walsh, 2006) was applied to extracts of transcribed classroom discourse recorded while AfL was being used. Walsh (2006) argues that the notion of *authenticity* of interactions understood as the types of interactions that could happen in everyday situations outside the classroom fails to acknowledge that a classroom is an authentic situation itself. This study adopted the view that, for learners and teachers, a classroom provides a context for authentic

interactions. Hence, it is believed that a relationship exists between the use of language and the pedagogical aims of lessons. By analysing language and pedagogical aims together, Walsh (2006) argues, researchers are able to analyse language in a manner that is more sensitive to the complex context of a classroom. Such a context is created by the interlocutors in the process of interacting.

This type of analysis is a departure from the more traditional, three-part frameworks for analysing classroom discourse used in SLA. These include the well-established initiation-response-feedback/follow up (IRF) or that developed through a study conducted in a TEYL context by Jarvis and Robinson (1997): focus, build, summarise. Walsh's (2006) approach was adopted here because the focus of the current study was to investigate if and how the discourse used during the use of AfL could facilitate learning. Effectively, the analysis aimed to provide insights into the relationship between using language and achieving pedagogic goals and/or creating learning opportunities. It is recognised here that this innovative approach to analysing classroom discourse is characterised by the absence of shared terminology to describe the findings. Terms such as *contexts* (Oliver & Mackey, 2003) or *pedagogic functions* (Jarvis & Robinson, 1997) or *modes* (Walsh, 2006) have been used in other studies. This study uses Walsh's terminology.

The extracts of classroom discourse were analysed using Walsh's (2006) framework, which identifies four modes of interaction: *managerial, materials, skills and systems* and *classroom context*. The managerial mode is characterised by long teacher sequences, confirmation checks (e.g. Is it clear?), transitional markers (e.g. right, so) and sometimes by the absence of learner turns. This mode aims to signpost activities, organise the space or provide classroom management information. In the materials mode 'pedagogic goals and language use centre on the materials being used' (p. 70). Interactions are often characterised by very little input from the teacher and the main aim is to complete the tasks at hand. The interactions in the skills and systems mode often follow the IRF pattern. The aim is to foster the development of accurate language forms and sub-skills (e.g. skimming). Finally, in the classroom context mode, interactions are less constrained and allow for the development of topics of interest to learners. The aims of such interactions are to develop fluency and enable learners to talk about their own experiences and feelings in the lessons.

Walsh (ibid.) acknowledged that these modes are not exhaustive of what happens in a language classroom and that some conversations cannot be categorised within a single mode. He identified mode side sequences (i.e. when a conversation starts with a main mode, temporarily moves to a secondary mode and then returns to the initial, main mode). Walsh (ibid.) reported the following patterns of mode side sequences:

• skills and systems – classroom context – skills and systems

• classroom context – skills and systems – classroom context

• materials – skills and systems – materials

• materials – managerial – materials

• materials – classroom context – materials

• managerial – skills and systems – managerial

Difficulties in identifying the mode of a conversation can also arise when it is uncertain from the teacher's use of language what the pedagogical aim for the conversation is. Walsh (ibid.) observed that when 'teacher talk and learning objectives are incongruent – the teacher's use of language actually appears to hinder rather than facilitate learning opportunity' (Walsh, 2006, p. 88).

The aims of the analysis were to:

- categorize the interactions which occurred during the use of AfL into the modes proposed by Walsh (2006);
- investigate whether teachers aligned their use of language with the pedagogical aims for conversations when they were using AfL;
- evaluate whether similar observations could be made in different modes;
- evaluate whether conversations which occurred during the use of AfL offered opportunities to contribute to achieving the learning aims;
- if so, whether those opportunities were effectively used;
- understand whether, during the use of AfL, learners and teachers entered into collaborative dialogues with LREs;
- understand the results of such LREs, especially in terms of input and output modifications.

Finally, it seems important to recall that Walsh (ibid.) developed his framework in a TEFL context and subsequently deployed it with primary school aged children. Hence, it was considered appropriate for the context of the current study.

This section has presented the processes of analysis that were applied to data collected in the cross-sectional phase. The analysis used in the longitudinal phase is described in the following section.

3.3.4 The procedure of analysing data from the Delayed Teacher Questionnaire

The delayed teacher questionnaire was returned by all eight teachers. Data collected using a Likert-type scale (Q2) were collated to create a sliding scale of use after sixteen months. This outcome was compared with a similar scale obtained from analysing the ROWDs and lesson observations in the cross-sectional phase. This analysis provided insights into changes that occurred over time. Open-ended questions generated qualitative data that were coded following the same procedure as that used in analysing the interview transcripts. The findings offered a more nuanced insight into changes in the use of AfL over time. The longitudinal findings are reported in Part Two of Chapter 4.

3.4 The limitations of the design

This section discusses the limitations of the study design. First, it is acknowledged that in the complex context of a classroom, judgements of what constituted a task were ultimately subjective. Although the lesson observations used a clearly defined unit of measurement, i.e. a task, the subjective nature of observation field notes implied that what the researcher considered as one task may have been interpreted differently by a different researcher. However, it is believed that, by comparing the researcher's interpretations with the teachers' records in ROWDs, a sufficient level of reliability was ensured.

Secondly, quantifying interactions, especially the L-L type, was challenging. Due to the number of learners who participated in the lessons, the number of interactions could fluctuate during one task. For example, within a class of ten learners, four dyads could be interacting while the remaining two would not be. Also more than one mode of interaction could be happening simultaneously. These were all recorded in the field notes and included in the analysis. These challenges reflected the non-participatory and exploratory

nature of the study within the real classroom. The study design accounted for this limitation by triangulating findings from quantitative analysis with qualitative one.

Another limitation of the study design was connected with recording the classroom discourse. These recordings were made in conditions similar to those described by Walsh (2006), i.e. 'under normal classroom conditions with no specialist equipment' (p. 165). This meant that the amount of available equipment and the methodological choice to limit the impact of the researcher on the context made it impossible to record every single conversation among teachers, learners and peers. The classroom discourse reported in Chapter 4 provided examples that would enable more in-depth analysis. However, the process was limited in that it did not offer the possibility of analysing all the conversations occurring in the lessons. Hence, it offered only a partial insight into the potential immense richness of classroom discourse. This type of limitation is inherent in research based in settings such as the one in the current study.

The above limitations are acknowledged in order to fully report on the study design.

3.5 Summary of the chapter

This chapter has provided an overview of the contextual framework for the study, including the educational context of the country, of the school, and of the sampling and the process of introducing AfL into the school that preceded this study. It has been highlighted that this study did not set out to evaluate the training process of teachers in AfL but focused on exploring and describing how they understood AfL, how they implemented it and on the observable impact of AfL on interactions in TEYL classrooms. These research foci offered an opportunity for a useful contribution to knowledge by exploring the concept of AfL within a context where it is largely under-researched and by offering empirical data linking AfL to facilitating learning: a call repeatedly made by other researchers (e.g. Bennett, 2011).

This chapter has also reported on how data collection tools were designed, piloted and implemented. It has also provided an overview of resulting datasets and how they were analysed to address the three research questions. By providing this full account supplemented by several appendices, the researcher ensured the interpretive rigour of the study.

The chapter has concluded by discussing the limitations of the design. The next chapter discusses the findings obtained by applying the study design reported in the current chapter.

Chapter Four: Findings

4.1 Introduction to Chapter Four

The data analysis process reported in Chapter 3 resulted in a set of findings for each research question in the current study. This chapter discusses the findings in three parts, each one corresponding to one research question.

Part One (Section 4.2) addresses RQ1. It discusses teachers' understanding of AfL in a TEYL context. The findings reported in Section 4.2.2 indicate what teachers understood by the concept of AfL. The remaining sections in Part One present teachers' reports about their own use of AfL and the impact it had in their classrooms. This connects the discussion about teachers' understanding with the findings based on empirical data about the use and impact of AfL in TEYL classrooms which are presented in Parts Two and Three of this chapter. Parts Two and Three address RQ2 (2.1 and 2.2) and RQ3, respectively.

The data related to the use of AfL were analysed in terms of its *frequency*, *diversity* and *richness*. Frequency is related to how often AfL was deployed across and within lessons. Findings about frequency were obtained through applying descriptive statistics to data obtained from 448 lessons self-reported by teachers and 28 lesson observations. Diversity refers to the range of AfL technique type, purposes and timing of the use of AfL. The initial sections of Part Two report on the cross-sectional perspective obtained through analysing the data from Phase One of the current study (RQ2.1). Subsequently, the longitudinal insights are provided by reporting on the findings from Phase Two in Section 3.3 (RQ2.2). Examining frequency and diversity allowed the researcher to gain insights into how the teachers implemented AfL in TEYL classrooms, having received limited training. Richness is understood as a qualitative dimension of AfL. It was evaluated by analysing the contribution that the use of AfL made to facilitating learning. The findings were obtained through applying the Storch (2002) model to classroom discourse. It is believed that the richness of AfL could contribute to supporting learning by facilitating the type of interaction patterns and conversations that have been shown by research (Butler & Zeng, 2014; Swain, 2000) to facilitate language learning. Thus, investigating the richness of AfL in TEYL classrooms allowed for making inferences about the impact that this type of assessment had on facilitating learning. Furthermore, the analysis explored the relationship between classroom discourse which occurred during the use of

AfL and meeting the pedagogical aims for the lessons. This was achieved through deploying the Variable Approach (Walsh, 2006) to analysing L2 classroom interactions.

The findings presented for each RQ are summarised in the final section of each of the three parts in the current chapter. The next chapter (5) follows with a discussion of the findings for each RQ in the context of the whole study in the light of the literature.

4.2 Part One: Teachers' understanding of AfL in a TEYL context (RQ1)

This part presents the findings for Research Question One: How do teachers understand AfL after receiving a limited amount of training and using AfL techniques for at least one academic year when teaching English to young learners aged 7-11?

4.2.1 Introduction to Part One

RQ1 aimed to investigate teachers' understanding of AfL. It was a particularly valuable research focus as it provided insights into the nature of AfL in TEYL classrooms: an insight currently not available in the literature, as indicated by the review in Chapter 2. Hence, by addressing RQ1, this study could contribute to the discussion about the theoretical framework of AfL. Furthermore, establishing what teachers understood as AfL was considered a pre-requisite to investigating the foci of RQ2.1 (use of AfL), RQ 2.2 (changes in use over time) and RQ3 (impact on interactions). The data that provided insights into this area were drawn from eight teacher interviews and one focus group. The development of themes described in Section 3.3, was fundamental to gaining insights into the teachers' understanding.

The themes identified in the process of analysing the data are reported in three sections. Section 4.2.2 focuses on what the teachers perceived as AfL. Section 4.2.3 focuses on the themes related to the teachers' reports of how they implemented their understanding in practice. Finally, Section 4.2.4 presents the emergent themes that provided insights into the teachers' understanding of the impact of AfL on interactions in TEYL classrooms. By reporting the findings in these three categories, this section provides links with Sections 4.3 and 4.4, which present empirical findings about the use (RQ2: 2.1 and 2.2) and impact (RQ3) of AfL, respectively.

4.2.2 What do teachers understand as Assessment for Learning in a TEYL context?

This section reports on the findings about what teachers perceived as AfL. Each theme is discussed in a separate subsection.

4.2.2.1 The integration of assessment with teaching and learning

The most frequently mentioned theme identified in the process of analysing the content was related to how teachers positioned AfL in relation to their teaching practice. The overall belief was that AfL implied integrating assessment into classroom practice. The findings related to this theme are summarised in six statements, reported in Points A-F below. Each point is accompanied by quotes from the interview and focus group transcripts, which exemplify the teachers' comments.

Theme 1: AfL implies a degree of integrating assessment with teaching and learning by:

  A. making the learners more aware of what they are learning

> *'you're getting them to try and think a little bit what the aim means and whether they've achieved it so any activity that's related to this is I guess AfL.'* (T6/INT)

> *'it's all about sort of cluing them into the learning process.'* (T2/INT)

> *'it focuses on the whole process of learning and the stages and makes the kids more aware of this process and they can consciously participate in it'* (T5/INT)

  B. focusing the learners on achieving the learning objectives

> *'it's like setting up goals and then kind of being accountable for those goals all the time with them and they know what they should achieve'* (T8/INT)

> *'it's a way of students and teacher establishing clear goals and working* together to achieve those goals' (T1/INT)

  C. continuously building AfL opportunities into lessons

> *'since it happens more regularly because I do it practically every lesson or almost every lesson I think you think paradoxically about assessment more often than with a test every month or something like that and the students' progress*

and their needs what they need actually and you keep putting this in your teaching' (T5/INT)

'my students have the assessment like AfL built into what they do in the lessons so they know that it's a part of the lesson and it's also important I also think that it's continuous and it's more meaningful because they know it's about the things we learnt to do that lesson and not something that they did a while ago and now perhaps cannot remember' (T1/INT)

D. providing a framework for giving meaningful feedback (from teacher, peers and own reflection) that helps to move learning forward:

'it's a framework and it defines (1)[9] *lets students know what they have to do and makes it easy for me or them to mark because everyone knows what the criteria are and they get useful feedback out of that and acting on it'* (T3/INT)

'you're not just handing them a mark and telling them you're poor at this or good at this but you're getting them to try and think about the aim and whether they've achieved it and if not then help them work out exactly how to achieve this' (T6/INT)

E. motivating the learners to learn and helping them to enjoy the process

'training learners to learn and helping them see the advantages of learning helping them enjoy learning motivating them (1) *this is what I think it could be'* (T7/INT)

'it's sort of like you were trying to put the assessment and the teaching together (1) *it's not like giving them a test every so often* (1) *so I think that perhaps it motivates them a bit more to be more focused or attentive the whole time and not just sort of revise vocabulary for a test if you see what I mean.'* (T2/INT)

F. providing a structure and focus to the teachers' lesson planning

'it's definitely helped with lesson planning and semester planning it's given me clear targets (1) *within the lesson it provides a REALLY good structure of how*

---

[9] Number of seconds of a pause. For details of transcribing convention, please see Appendix 15.

things are organised and I think the students have benefitted from that because they can have more input in the lesson themselves' (T1/INT)

'my planning now includes this so definitely the planning is more thorough and I use it to set time aside for this type of thing [AfL]*'* (T7/INT)

The findings reported above suggest that AfL was understood as a part of the teaching that took place in TEYL classrooms. Teachers reported that AfL allowed them to make learners aware of what the learning aims for each lesson were. It could be inferred from the comments quoted in point B that when learners became aware of the learning aims, teachers could use AfL to focus their and the learners' efforts on achieving those aims. It seems that one way in which this could be achieved was to build opportunities for continuous assessment and feedback into lessons (Points C and D). Notably, teachers themselves seemed to draw a distinction between this type of assessment and summative tests. They indicated that AfL occurred simultaneously with the teaching and learning processes, as opposed to testing which tended to be deferred in time. This characteristic implied that AfL practices were contextualised by the teaching and learning that was taking place.

Teachers' comments also indicated that the feedback that the learners received through AfL could move learning forward (Point D). Most significantly, it facilitated ongoing evaluation of where the learners had arrived with reference to achieving the learning objectives for the lessons. Assessing the students' performance while they were learning was a distinctive characteristic of AfL in its formative function. Additionally, it is important to note that the ongoing teaching and learning served as a context for feedback. In practice, the children were given opportunities to relate feedback to the aims and criteria for success that had been shared at the beginning of the lesson. Those who implemented AfL in their lessons evaluated such integration of AfL with teaching positively. Attention was especially drawn to the fact that, because it happened in a context of teaching and learning in a continuous manner, AfL motivated the learners to be attentive and systematic (Point E). It was also interesting to note that the teachers indicated that the learners played an active role in the assessment process. The learners' main responsibility was not only to provide feedback, but more importantly to take control of their learning.

It was also indicated that the process of integrating assessment with teaching and learning was connected with lesson planning (Point F). Most notably, time allocated for providing feedback or self-reflection was allocated in the lesson plan. Teachers also reported that when they used AfL, their lesson plans became more focused on learners achieving learning goals within a lesson.

The teachers' understanding of AfL reported above indicated that AfL might be considered to be a way of integrating assessment practices into teaching and learning. The emergence of that theme was an important finding, especially when compared with the content of the training sessions (which took place 14-16 months before the interviews), in which the teachers were introduced to a variety of AfL techniques and little attention was devoted to the underlying principles. As previously explained, this study does not aim to evaluate that training. However, it seems useful and justifiable to compare the teachers' understanding of what AfL is with the input of the training. Notably, there were differences in how AfL was presented to this group of teachers (i.e. as set of techniques) and the way they conceptualised it (i.e. as a way of integrating assessment into teaching and learning). This suggested that those teachers reflected on AfL and were sharing their own understanding, rather than reproducing what had been presented to them. Thus, it seems reasonable to claim that the analysis reported here provides an insight into the understanding of AfL developed by the teachers specifically in a TEYL context.

The finding that AfL may facilitate the integration of assessment with teaching and learning raised the question of compatibility of AfL with the teaching methodology used in the context of the study. The teachers' reports on that issue are presented in the following section.

4.2.2.2 AfL in a TEYL classroom

In their accounts of AfL, some teachers also discussed an issue of broader methodological nature: the compatibility of AfL with their teaching methodology. This view is summarised in Theme 2 below.

Theme 2: AfL is compatible with the teaching methodology used by the teachers in the study but tensions exist between AfL and summative reporting.

The compatibility of AfL with the teaching methodology adopted by the teachers was expressed in two different ways:

A. AfL was considered a new name for a concept that was already present in the teachers' TEYL practice

'initially I felt like maybe I already did some elements of this especially with good and bad models' (T1/INT)

'it's always kind of existed like TEFL would call it concept checking or checking instructions having students tell you the instructions back' (T2/INT)

B. AfL was considered easy to implement in TEYL classrooms

'to a certain extent it [AfL] *is the means to an end just a part of what you're doing anyway with them just teaching them in slightly different way and it is helpful'* (T6/INT)

'it [a lesson] *has a much clearer purpose they* [learners] *know what is happening and when it is seamless like thumbs up and down with the primaries for example this is all seamless and normal and very informative to show what they have learnt'* (T2/INT)

Points A and B above explain how some teachers understood the compatibility of AfL with their teaching methodology[10]. AfL was considered easy to implement as a part of the activities that teachers used in their young learners' classrooms. Additionally, some teachers recognised that some of the practices that they had used for several years were very close in nature to AfL, although the terminology used to describe them differed. However, the comments do not indicate whether the teachers considered such practices to be a component of assessment before they became familiar with AfL.

However, it should also be noted that there was one area where a number of teachers considered AfL less compatible with their practice. They indicated that parents had expectations of numeric grades to be included in the reporting, suggesting that this

---

[10] It is useful to note at this point that the teaching at the school was based on a series of coursebooks (Appendix 3) designed for school children. From the lesson observation data, it was evident that lessons were designed to include activities which provided opportunities to practice speaking, listening, reading and writing as well as vocabulary with some focus on form; and the lessons tended to follow a version of the PPP (presentation, practice, production) format, with the final P often absent. The methods used included songs, role plays, elements of TPR, drills, games, coursebook based listening and reading comprehension as well as word, sentence and text level writing activities.

might be due to the characteristics of the wider educational context. This is summarised in Point C below.

C. There existed some tension between AfL and summative assessment in a TEYL context, which seemed to be connected to reporting to parents

> *'you have parents and they want you to be able to give them word lists to learn and numbers percentages to show how much the child can do what level they are and they seem to be very much focused on the grades and numbers (1) I'm not sure if that's because this is what the Polish schools do but for parents a teacher's and child's opinion of progress does not seem to be enough'* (T1/INT)

> *'I think that we'd need to sell it to parents because Polish parents are very much you know the exact grades they expect numbers out of five or percentages'* (T3/INT)

Two observations could be made about the finding that the summative reporting practice was considered not compatible with the use of AfL. Firstly, the necessity of obtaining numerical grades for written reports seemed to have a washback effect on classroom practice. Specifically, it seemed to inhibit the implementation of AfL. This interpretation finds confirmation in the empirical data as careful scrutiny of the ROWDs evidenced that teachers recorded less use of AfL at times preceding the reporting period (Appendix 17). Secondly, it indicated that parental expectations may impact on the teaching and assessment practices in a TEYL context. This may be a factor specific to teaching children that would not occur as often, if at all, in adult contexts. This finding highlighted that the teachers' classroom practice was influenced by certain contextual factors: the school policy of summative reporting, parental expectations and reporting practices in a wider educational system.

The next section discusses further the teachers' understanding. The focus shifts to how teachers reported their own implementation of AfL in lessons. It is important to note that all the findings for RQ1, discussed in Part One of this chapter, provide insights into the teachers' understanding of how they implemented AfL. Hence, what is actually reported is the teachers' interpretations of what they did. Empirical data from the lesson observations that offer insights into the use of AfL are reported in Part Two.

4.2.3 The teachers' beliefs about using AfL

This section reports on two themes related to using AfL that were identified through analysing the interviews and the focus group transcripts. Each of them is discussed in a separate section below.

4.2. 3.1 AfL techniques

The most frequently recorded theme related to the use of AfL was that teachers recognised that *'tactics, tools and techniques'* (T2/INT) were used to implement AfL in the classroom. This is captured in Theme 3, supported with examples of how the teachers described the implementation of AfL.

Theme 3: AfL could be implemented in a TEYL context using a number of different AfL techniques:

> *'AfL creates a kind of a circle when everybody knows where everybody is (1) and then what it involves is a series of tactics tools techniques and things that you can do to facilitate that'* (T2/INT)

> *'depending on the lesson plan we use the* [AfL] *techniques at different parts throughout the lesson to help them get there and be able to do the can[11] (1) every single lesson they have at least one element of AfL'* (T7/INT)

> *'obviously the can dos traffic lights success criteria quite a bit and thumbs up thumbs down to check understanding that's what I can think of off the top of my head'* (T6/INT)

> *'I used some of the techniques for marking writing such as two stars and a wish and perfect purple and green for growth'* (T3/INT)

Theme 3 suggested that several AfL technique types could be used in TEYL. The full list of AfL techniques identified in the study, with descriptions, can be found in Appendix 18. This finding was important for the current study as it provided a rationale for choosing 'an instance of implementation of an AfL technique' as a unit of the observable use of AfL. This allowed the researcher to report on the use of AfL that was consistent with the

---

[11] Saying 'the can', T7 is referring to 'Can I ......?' type questions which (s)he often used to start the lessons, e.g. 'Can I tell my partner about ten things that I did yesterday?'

understanding discussed in this chapter; hence the interpretations remained faithful to the emic perspective adopted in this study.

Additionally, it should be noted that as the teachers reported which technique types they implemented, they indicated for which purposes the techniques were employed. This suggests that when selecting technique types for use, teachers considered the purpose of use. This area is discussed in more detail in the following section.

4.2.3.2 The purposes for using AfL

The second theme related to the use of AfL examines the purposes for using AfL. In the interviews, the teachers indicated seven; these are shown in Table 4.1 below. The most frequently cited purpose was that AfL aided teachers in giving and clarifying instructions. The second purpose cited was to share learning aims. The third purpose cited was to provide feedback on students' learning. The fourth purpose was related to evaluating how confident learners felt about their own progress towards achieving the learning aims. Table 4.1 indicates the frequency of the occurrence of comments about each purpose in the teacher interview transcripts. The other purposes cited were to keep records, to set homework and to communicate with parents, though these were less frequent.

Table 4.1: Purposes of using AfL in a TEYL context as reported by teachers

Purpose for using AfL	T1	T2	T3	T4	T5	T6	T7	T8	Total
To give and clarify instructions	1	2	0	3	4	0	3	2	15
To share learning aims and criteria for success	2	0	1	0	3	1	3	0	10
To provide feedback	2	0	1	0	0	2	2	0	6
To evaluate learners' confidence	3	0	0	0	1	0	1	0	5
To keep records	0	0	0	0	0	2	0	0	2
To set homework	0	0	2	0	0	0	0	0	2
To communicate with parents	0	0	1	0	0	0	0	0	1

An interesting insight into the relationship between the purposes for using AfL and the use of particular techniques was gained through analysing the focus group data. In Extract 4.1 below, the teachers were discussing how they implemented AfL in their practice. It was inferred from this conversation that they recognised AfL as a way of raising their students' awareness of what the learning aims were (Turn 1) and focusing the learners' efforts on achieving those aims through ongoing reflection and by making them aware of what and how to improve. The discussion suggested that some teachers implemented AfL

through a number of techniques e.g. traffic lights. However, it seems particularly noteworthy that the teachers considered the pedagogical aims for which they used the AfL techniques to be more important than the technique type (Turn 5). The discussion also suggested that it was possible to meet the same aim through using different techniques (Turn 3). Furthermore, the focus group data indicated that diversity in technique type was considered useful in TEYL (Turn 4) in order to maintain the students' interest or engagement. An extension of Extract 4.1 can be found in Appendix 19.

EXTRACT 4.1

(FG discussion on what AfL is, May 2012, for a longer extract see Appendix 19)

[1]   **T3:** *the whole thing is about that; getting them to understand what to do and*

[2]   *then to figure out how well they have done it and it doesn't matter which*

[3]   *techniques you use, right?* (...)

[4]   **T8:** *I'll support that actually both of you I think (1) some variety is needed but*

[5]   *in fact it's the purposes that make it all meaningful and worthwhile*

This section has discussed the teachers' beliefs relating to how they implemented their understanding of AfL in TEYL. The findings indicated that AfL could be implemented through a variety of techniques that were used mainly to communicate instructions, learning aims and feedback to the students and in some cases to gauge the learners' confidence about their performance and learning. The next section continues the report of findings for RQ1 by presenting teachers' beliefs relating to the impact that the implementation of AfL had on interactions in the TEYL classrooms that featured in this study.

4.2.4 The teachers' reports on the impact of AfL on interactions in the TEYL classrooms in this study

From the content analysis of the teacher interviews and focus group transcripts, two themes emerged: the increased independence of learners and the increased number of one-to-one interactions occurring in the lessons. They indicated that there might be a relationship between using AfL in TEYL and interactions that occur in lessons. Both are discussed in this section.

Theme 5: When AfL was used, learners were able to sustain independent and pair work for longer, without the need for support from the teacher by:

A. providing scaffolding

'they get used to this that they're given steps in what to do and they need this and they want this and they really (1) and it makes them actually complete the whole task rather than shout out finished (1) because they know exactly what they need to do' (T5/INT)

B. encouraging the learners to become more responsible for own learning

'first of all they are more responsible for their learning and this is something very very important (1) this is something that when you use AfL really small kids start feeling' (T3/INT)

'it was amazing how quickly kids got used to taking some responsibility and when you think about young learners they still don't have that concept of taking responsibility for their learning until you start using AfL and then somehow it clicks for them so for me that does it (1) I no longer have to be responsible for everything in the classroom when I use AfL' (T4/INT)

Many teachers reported that by using AfL techniques they were able to create conditions in which students became aware of how to proceed with task completion independently and that this was facilitated in two ways. First, the teachers were able to set tasks which better scaffolded the learners' work. This meant that the AfL techniques provided the support necessary for students to complete the task at hand. In doing so, it seems that AfL might have facilitated the process of learners acting as instructional mediators for themselves or their peers. The effectiveness of this approach may be explained by Vygotsky's theory of the zone of proximal development (Section 2.2.2.2). The second process, which teachers identified as an impact of AfL in their lessons, was creating conditions in which even the youngest learners were able to take responsibility for their own learning. This demonstrates how AfL helped to scaffold learning. It also seems to indicate that the roles that teachers and learners played in the lessons changed in the classrooms where AfL was used. In other words, the learners took ownership of their learning such that the teachers were not solely sources of language but became facilitators of language learning. However, fully examining the roles of the teachers and the learners and changes that occurred is beyond the scope of this study but would undoubtedly constitute a worthwhile focus for future research.

Theme 5 also resonates with an extract from the focus group (Extract 4.2, below). The focus group data indicated that when learners required less assistance to complete the tasks in hand, the teachers could spend more time interacting on a one-to-one basis with the learners and in doing so could focus on facilitating leaning. This is summarised in Theme 6 below.

Theme 6: Using AfL allowed for a larger number of one-to-one interactions between teachers, learners and peers.

    A. When the students were able to work without a teacher's help, the teacher could spend time on monitoring work more effectively and on providing individual support (T-1L interactions).

Extract 4.2

(Focus Group discussion, May 2012, for an extended version see Appendix 27)

[1]   **T2:** *with time input from teacher should be smaller* (…) *because they know their*

[2]   *success criteria how to do it or they should be aware of what is expected but I think*

[3]   *at that point monitoring becomes more effective to make sure that they're actually*

[4]   *doing it properly*

[5]   **T4:** *I think this is a very good point=*

[6]   **T1:** *=yes more time to monitor better* (…)

[7]   **T2:** *yes and how you would do that is an important element of this discussion*

[8]   *when you monitor and how you monitor and what you say to individual students*

    B. Using AfL enabled the teachers to introduce more pair work in the lessons (L-L interactions).

        'the one that I also found the most useful has been the learning partners idea (1) obviously there is benefit to the students in taking a bit of the control themselves and in doing that helping each other' (T4/INT)

        'I used this with primaries mainly to involve more pair work and it sort of helps them get what pair work is about.' (T7/INT)

    C. The students collaborated rather than competed when working together (L-L interactions).

'I think the biggest benefit for my groups has been the peer learning working together and not being competitive in their English but being supportive of each other's learning' (T1/INT)

Theme 6A seemed to complement Theme 5 by explaining what teachers tended to do when students were working independently. The discussion during the focus group indicated that, in the lessons when AfL was used, teachers had more time to monitor the students' work (Turn 6). The teachers who participated in the discussion believed that the time could be spent on T-1L types of interaction, offering individual support to the learners (Turns 7-8), thus recognising that the quality of such interactions was important (Turn 8).

Themes 6B and 6C offered insights into the teachers' reflections about the relationship between L-L interactions and the use of AfL. Theme 6B suggested that using AfL techniques made it more possible, or perhaps easier, for teachers to introduce pair work activities in TEYL classrooms hence naturally increasing the number of L-L interactions. Whereas Theme 6C indicated that what happened during L-L interactions was perceived by teachers as collaborating to complete a task together and not competing with each other. This finding may also provide some explanation as to why the learners were able to sustain pair work for longer (Theme 5). They did so by collaborating and supporting each other in completing tasks.

The findings reported in the current section have pointed to a possible relationship between the use of AfL and interactions in TEYL classrooms. Empirical findings about the impact of AfL on the interactions are reported in Part Three of the current chapter.

4.2.5 Summary of findings for Research Question One

The six themes that emerged from the data collected to address RQ1 provide an insight into teachers' understanding of AfL. They indicate that AfL was considered to be an approach that helped to integrate assessment into teaching and learning in a TEYL classroom. This was considered to be compatible with communicative classroom practice but not with summative reporting. Most teachers believed that AfL was implemented in lessons through a variety of AfL techniques that were used for purposes linked to ensuring that the students understood what they were learning, how they could achieve their aims, where they were in relation to those aims and what to do to achieve them. Finally, it was

also reported that using AfL created conditions in the classroom that facilitated one-to-one interactions between the teachers and learners and enabled the learners to sustain individual and pair work independently of any teacher support.

The next part of the current chapter reports the findings about the use of AfL in TEYL.

4.3 Part Two: The use of Assessment for Learning in a TEYL context

4.3.1 Introduction to Part Two

Building on the discussion of what the teachers understood as AfL in a TEYL context, this part presents the findings that address RQ2.1: How do teachers' translate their understanding of AfL into classroom practice in a TEYL context with students aged 7-11 in a private language school in Poland?; and RQ 2.2: Do teachers report any changes in their practice of using AfL over time?

The data sets included ROWDs from 448 lessons and field notes from 28 lesson observations. The findings are presented in two sections. Section 4.3.2 focuses on AfL technique types which the teachers used, including when and for what purposes these were implemented. This is followed by Section 4.3.3, which presents the findings about the frequency and diversity of use of AfL techniques. Comparisons are made between individual teachers and between the two age groups of learners (7-9 and 10-11 year olds). These provide insights into between-teacher variance in using AfL and differences in implementation across the age groups. The final part of Section 4.3.3 presents findings about changes in the frequency and diversity of use over time. The data offering that longitudinal perspective came from the delayed teacher questionnaire.

4.3.2 The use of AfL techniques in TEYL

This section provides an inventory of the AfL techniques recorded in the present study. As discussed in Part One, the teachers believed that AfL could be implemented with AfL techniques. For that reason, 'an instance of the use of an AfL technique' was adopted here as an observable representation of the use of AfL in the observed lessons. This choice provided a pragmatically efficient way of recording and analysing the implementation of AfL.

4.3.2.1 AfL techniques

The inventory of AfL techniques was based on data drawn from ROWDs and lesson observations. Eighteen AfL techniques were identified in all from both sources. The techniques were ordered from the most to the least commonly used one. Because the data came from 28 lesson observations and 448 lessons recorded on ROWDs (8 teachers x 56 lessons each), the following procedure was adopted to ensure the equal weighting of the sources:

1. AfL techniques were identified and separate frequency counts based on ROWDs and lesson observations were created.

2. Within each of the data sources, the identified AfL techniques were ordered ranging from the most to the least frequently used one.

3. As eighteen AfL techniques were identified, points from '18 – most frequently occurring' to '1- least frequently occurred' were awarded to each list. Those techniques not found in a given data source were assigned 0 points. The scoring can be reviewed in Appendix 20.

4. Both lists of AfL techniques were combined by adding the points assigned to each technique and the results are collated in Table 4.2.

The field notes from lesson observations were scrutinised to identify the language skills and types of activities where AfL was used (Table 4.2, Column C) and the timing of implementation (Column D). Detailed descriptions of all AfL techniques used by the teachers in the study are provided in Appendix 18.

Table 4.2: An inventory of AfL techniques

Column A.	Column B. Name of AfL technique	Column C. Types of tasks and skills the technique was used with	Column D. When it was used within a lesson.
1	Success Criteria (SC)	Writing, Arts and crafts, Classroom instructions	Before and during a task
2	Learning Partners (LP)	Speaking; Vocabulary learning	Throughout lessons and tasks; Often in preparation for learners to work independently
3	What are we learning today? - type questions (WALT)	Learning objectives for the lesson	At the beginning of a lesson; Often referred to throughout the lesson
4	Traffic Lights (TL)	Speaking; Vocabulary learning	At the end of the lesson or a task; Sometimes followed directly by NST
5	Thumbs up or down (THUD)	Classroom instructions; Arts and crafts	Throughout lessons; Often during and after giving instructions;
6	Two stars and a wish (TSAW)	Writing	After an activity; Often when SC were used before the activity
7	Sharing a good and a bad model (SGBM)	Writing	At the beginning of the lesson; Sometimes with reference to ICS or WALT
8	Smiley faces (SF)	Writing; Speaking	After a task; At the end of the lesson
9	'I can' statements (ICS)	Learning objectives for the lesson	At the end of the lesson
10	Perfect Purple and red to Remember (PPRR)	Grammar Writing (often with grammar focus)	After an activity has been completed to be revisited by a student; Often followed by NST; Often recorded in notebooks
11	Next steps (NST)	Writing; Vocabulary learning; Grammar	Often after a different AfL technique; Most commonly after TL and PPRR
12	Colour coding (CC)	Writing	After an activity has been completed to be revisited by a student; Often followed by NST or combined with SC; Often recorded in notebooks
13	Find the Fib (FTF)	Speaking; Vocabulary; Grammar	After a new grammar rule or set of vocabulary has been introduced and practiced
14	Increased thinking time (ITT)	Speaking; Reading comprehension tasks	Throughout lessons or activities; Often after asking a question
15	Star charts (SCH)	Learning objectives for the lesson	At the end of a lesson
16	Indicate mistakes without explanations (IMWE)	Grammar; Spelling; Writing; Reading comprehension tasks	Recorded in notebooks; While a task is being completed; Often with reference to ICS or SC
17	Sheriff's star (SST)	Learning objectives for the lesson	At the end of the lesson or an activity
18	Mind maps (MM)	Projects	At the beginning of a project or lesson; Referred to at the end

The data presented in Column C of Table 4.2 indicate that teachers were observed to use AfL largely for teaching productive skills: speaking and writing. There was also some indication in the data that AfL was used while teaching vocabulary and grammar, conducting arts and crafts activities or providing feedback on reading comprehension tasks. This finding suggests that AfL was helpful in teaching productive skills. It could be inferred that a possible reason for that finding is connected with a difference in the degree of scaffolding inherently provided by the task design of the productive and receptive activities. In the tasks aiming to develop receptive skills, especially listening (often implemented through songs or listening comprehension activities), the duration was determined by the task, not the learner. In the tasks aiming to develop productive skills, such as describing something, the learners were expected to use prompts to complete the task. Because of these considerations, it seems that productive tasks inherently offered less scaffolding. Hence, more support was needed from outside the task itself. The findings presented in this chapter suggested that AfL techniques could offer that type of support. T6 explained the use of AfL for writing tasks by saying:

> *'it's a really good way to make them focus on a few different things in their writing (...) and the students can then focus on just a few key things and get a good grade whereas before it was not up to standard and you just had to fail them based on that but then here you can actually say* (1) *well you did this this and this* (1) *this was what was the most important and that's why they got a bad grade and they understand that and I think that's good that's probably the most positive thing'* (T6/INT)

The above quote confirms that AfL techniques were believed to provide support for completing writing tasks by indicating what constituted good performance. This has clear implications for understanding AfL in the TEYL context. Namely, by specifying the criteria for success, teachers were able to focus their feedback on specific areas which were considered important within that task. The finding that AfL was used largely for teaching productive skills supports teacher beliefs that AfL helped to develop learner independence in sustaining a task at hand, as reported in Part One of the current chapter.

4.3.2.2 The timing in the use of AfL

This section reports data pertaining to the timing of using AfL within lessons.

In Table 4.2 each technique was reported with information about the timing of its use (Column D). The analysis reveals that there were three categories of 'time of use'. In the first category, the same technique types were used at the beginning of the lesson and at the beginning of a task. The second category includes technique types that were employed towards the end of a task. The same technique types were observed in the final phases of lessons. The third category of technique types were used continually during the activities. Figure 4.1 demonstrates graphically which AfL techniques were observed to be used within each of the 'time of use' categories. These observations suggested that there were reasons why some AfL techniques were selected for use at the beginning of an activity or a lesson and others towards the end or throughout.

As Figure 4.1 below shows, a wider range of AfL technique types could be observed to be used towards the end of tasks or lessons. This finding supports the findings referring to RQ1 that teachers believed that variety was needed when implementing AfL to ensure that students remained interested. However, this finding provided a more nuanced insight into the need for variety; it seems that a diversity of technique type was especially needed. Another observation, according to the data presented in Column B of Table 4.2, was that techniques which were used in the initial stages of tasks and lessons were the ones most frequently used (i.e. SC, LP and WALT). This suggests that the teachers used a smaller number of technique types for setting up the tasks but that they tended to opt for a greater diversity of AfL techniques towards the end of a task when facilitating feedback provision and working towards raising the learners' awareness of achievement.

Figure 4.1: Time of use of AfL techniques in the TEYL classrooms

The findings reported above suggest that there was a link between the timing of use of AfL techniques and the purpose(s) that they served. The nature of that link could be inferred by considering the lesson as a context for using AfL. As the consecutive stages tended to serve specific pedagogical aims, it seems reasonable to infer that the teachers used AfL techniques that served the purpose of meeting those aims. For instance, when explaining what learners were expected to demonstrate in a task, most teachers opted for AfL techniques that allowed clarification of the expected outcome (e.g. SGBM[12]) or the process of achieving it (e.g. SC). To provide further insight into the accuracy of that inference, video recordings of observations were reviewed and each use of AfL was assigned to one of the seven purposes for using AfL identified in the analysis of the data addressing RQ1.

---

[12] For detailed description of all AfL techniques which were identified in the current study, please see Appendix 18.

4.3.2.3 The purposes of using AfL

This section reports findings about the purposes of using AfL techniques in TEYL lessons. In the process of reviewing the lesson observations, the technique types were recorded against the relevant purpose(s) in three mutually exclusive categories. These were techniques used: 1) in both age groups, 2) with 7-9 year olds only, 3) with 10-11 year olds only. That analysis is summarised in Appendix 28. The outcomes demonstrated that the teachers' who used AfL selected different technique types to serve different purposes. This suggested that fitness for purpose might be a consideration in implementing AfL in TEYL classrooms. Comparing the AfL techniques which were used for each purpose between the two age groups in the study allowed for gaining insights into the differences in implementation of AfL in the two age groups. These findings are summarised in Table 4.3 and discussed below.

The findings presented in Column A of Table 4.3 indicate which purposes for using AfL in TEYL that were identified in teachers' interviews and the focus group discussion (Table 4.1, page 144) could be confirmed through empirical data from lesson observations. It is necessary to note that the data obtained through lesson observations suggested that two of the purposes which were identified in the teachers' reports (giving/clarifying instructions and measuring learners' confidence) seemed to be enacted as a part of other purposes (sharing learning objectives and expectations or providing feedback, respectively). As a result of this analysis, new terms for three categories of purposes which AfL was observed to serve in TEYL classrooms have been proposed. These are: *setting objectives and expectations*, *monitoring performance* and *checking achievement*. The descriptions of each of the purposes are provided in Column A.

The findings presented in Table 4.3 illustrate how AfL was practically implemented for each of the three categories of purposes and demonstrate the differences in implementation between the two age groups in the study. These are interesting insights, as the learners were at low levels of language proficiency and were developing their literacy skills. In such a context, sharing learning aims and feedback seems inherently difficult, while monitoring own or a peer's performance, or responding to monitoring conducted by a teacher may be metacognitively too challenging for the younger age group (7-9 year olds).

Furthermore, the empirical findings presented in Column A also point towards a complex nature of feedback provision through AfL. There was an indication in the data that the technique types used throughout the lessons or tasks were employed with the aim of helping the learners monitor their own performance (mostly in the older age group) or enable the teachers to monitor learner performance. This type of ongoing feedback was based on short fragments of the learners' work, e.g. one sentence from a longer piece of writing. In contrast, the feedback provided towards the end of a lesson or a task tended to be based on the entire performance during the given task or lesson that was being assessed, e.g. on the whole piece of writing. In both cases, the feedback aimed to indicate what positive elements the learners were able to demonstrate and/or consider how to improve. However, in the former case, the emphasis was on improving performance during the given task. Whereas in the latter case, the points for improvement could be addressed either through Next Steps (NST) techniques or later in future tasks or lessons (i.e. improvement can happen 'later'). This seems to be an important characteristic which warrants making a distinction between the purposes for providing feedback in Table 4.3.

Additionally, the findings that refer to feedback provision suggest that the same AfL techniques were used for providing feedback by teachers and by peers. A difference was observed between this group of techniques and the ones used to facilitate self-assessment and reflection. This suggests that the process of providing feedback might have differed depending on who the feedback provider was, i.e. the learners themselves or somebody else. This was an interesting insight as it indicates that feedback in a TEYL classroom could be a very complex phenomenon and would warrant detailed research in the future.

Table 4.3: Purposes for using AfL techniques based on data from lesson observations

Purposes for use of AfL confirmed through lesson observations				Column E:
Column A: **Purpose for using AfL techniques observed in lessons** **Description of the purpose** **(Purpose reported by teachers)**	Column B: **The relationship between the timing and the purpose for using AfL**	Column C: **Specific for 7-9 year olds**	Column D: **Specific for 10-11 year olds**	**Purposes for use of AfL not confirmed through lesson observations**
Setting objectives and expectations When teachers used AfL techniques to share learning objectives, they focused on clarifying expectations of outcomes or explicitly raising students' awareness of what they were learning. (Purpose reported by teachers: To share learning objectives. To give and clarify instructions)	The AfL techniques used at the beginning of tasks or lessons were used predominantly to ensure that the students understood the expectations of good performance, the learning objectives and their teacher's instructions.	The purpose of giving and clarifying instructions, reported by teachers, was observed only in the 7-9 year group. AfL techniques were less reliant on literacy skills than those used in the older age group.	A wider range of techniques was used in the older age group to explicitly raise the students' awareness of what they were learning.	There was no confirmation in the data from lesson observations that AfL was used for record keeping or setting up homework, in either of the age groups. Using AfL for the purposes of communication with parents was not observed in the 10-11 year olds' classes and only one instance was recorded in the younger age group.
Monitoring performance This was done by referring to the expectations set earlier in that lesson or by providing feedback on short fragments of the learners' performance. (Purpose reported by teachers: To provide feedback, To measure learners' confidence)	The techniques used throughout the lessons seemed to be used mostly for teacher or peer feedback on ongoing performance.	A greater reliance on monitoring by the teacher. There was a wider range of technique types used to measure students' confidence.	A wider range of technique types used to encourage self-monitoring,	
Checking achievement This included self-reflection or peer- and/or teacher feedback. The AfL techniques which were deployed for teacher feedback were also used for peer feedback while those used for self-assessment were different from those used for the first group. (Purpose reported by teachers: To provide feedback,).	The techniques used towards the end of a lesson or a task served the purpose of checking if students knew what they had learnt, and what the areas for improvement were.	There was a wider range of technique types used to provide teacher and peer feedback AfL techniques less reliant on literacy skills.	There was a wider range of technique types used to facilitate self-assessment and reflection on learning and areas for improvement.	

The findings presented in Column B above indicate that timing of use was indeed related to the purpose of using AfL techniques. This is useful as it confirms that there were techniques that were better suited to the initial stages of the activities or lessons when the aims were shared and the instructions given. Other AfL techniques were more appropriate towards the end of the activities when feedback was provided and areas for improvement explored. This provides some suggestion as to why the findings addressing RQ1 indicate that the majority of the teachers believed that AfL techniques were compatible with the teaching methods used at the study school. In other words, AfL techniques were used in the parts of the lessons that aimed to address a purpose that was effectively served by a specific set of techniques. This has clear pedagogical implications, as it suggests that how AfL is used and whether it can be successfully implemented is connected to the teaching methodology.

The findings in Columns C and D provide more detailed insights into the differences in how AfL was implemented in the two age groups. First, they indicate that the purpose of using AfL to give and clarify instructions expressed by the largest number of teachers in the teacher interviews (Table 4.1, p. 144), was only observed in the classes of the younger age group. This may suggest that the younger learners needed the type of scaffolding that AfL was believed to provide not only for sustaining work but also for comprehending instructions, or perhaps remembering them for long enough to actually complete the task. This resonates with the finding that teachers tended to deploy techniques that were visually attractive (e.g. SF, SCH, SS, see Appendix 28) and less dependent on literacy skills with the younger learners (7-9 year olds). The AfL techniques used with learners aged 10-11 (e.g. TSAW, IMWE, NST, see Appendix 28) tended to rely more on the students' ability to read and write.

The findings presented in Table 4.3 indicate how sharing learning objectives was enacted in TEYL. In both age groups, teachers used AfL techniques to help their learners understand the objectives in two ways. One focused on clarifying the expectations of the required standard of performance, e.g. by demonstrating examples of the outcomes that students were expected to complete by the end of a lesson. The other method entailed explicitly informing the students about what they were going to learn, often through *I can* statements. These were frequently discussed at the beginning of the lesson and returned to towards the end of the session, sometimes to facilitate self-assessment. Not surprisingly, there was more evidence of using the second method in the older age group,

where children's language and literacy levels tended to be higher. It also seems that by explicitly raising the learners' awareness of what they were learning, the teachers effectively contributed to raising their metacognitive awareness: especially of the requirements of the task at hand (see Section 2.2.1.3.2).

Furthermore, although AfL was used in both age groups to provide feedback, teachers were observed using a wider range of AfL technique types which facilitated self-assessment and reflection on own learning with the 10-11 year olds. In the classes of the 7-9 year olds, a greater reliance on feedback from teachers and peers was observed. This might be interpreted as follows: as learners developed their metacognition, they were able to become more effective at self-assessing and reflecting on their own progress. Another difference, between the two age groups in the study was evident in the use of technique types for judging how confident learners were about their learning. Two AfL techniques were used for that purpose exclusively in lessons with 7-9 year olds, a different one was used with both age groups. This might be due to differences in classroom pedagogy in the two year groups. Specifically, it could indicate that most teachers paid more attention to fostering a positive effect in the younger age group. An alternative interpretation could be that the teachers needed more tangible tools, like AfL techniques, to foster positive feelings, while the same may have been possible to achieve in the older age group without AfL, e.g. by using praise.

The differences in how AfL was implemented in the two age groups are summarised in Table 4.4 below.

Table 4.4: Between-age group differences in the use of AfL

7-9 year olds	10-11 year olds

1. Visual, picture based techniques
2. Wider diversity of technique type used for:
  - Measuring learners' confidence
  - Peer assessment
  - Teacher assessment
3. AfL frequently used for instruction giving
4. AfL also used for sharing learning aims and feedback
	1. AfL techniques reliant on literacy skills
	2. Wider diversity of technique type used for:
	  - Self-assessment and reflection on progress
	3. AfL used predominantly for sharing learning aims and feedback

The findings in Column E of Table 4.3 (p. 159) suggest that the purposes not strongly represented in interview data could not be confirmed through lesson observations. There was no evidence of using AfL for keeping records and setting homework in video recorded lessons. The one instance of using AfL to communicate with parents was observed solely in the younger age group. This indicates that some differences in the use of AfL might occur between different age groups.

4.3.2.4 Summary of findings about the use of AfL techniques

This section has presented findings about how AfL techniques were implemented in TEYL. It has reported that eighteen AfL techniques were identified in TEYL classrooms. These tended to be used when teaching productive skills, grammar and vocabulary. The findings confirm that teachers used AfL for the purposes of giving feedback, sharing learning aims and gauging learners' confidence with both age groups (7-9 and 10-11 year olds) as well as for giving and clarifying instructions with the younger group. The findings also suggest that while the timing in the use of AfL techniques was related to the purpose(s) for using them, a greater variety of techniques was observed at the end of the lessons and tasks, perhaps to maintain students' interest levels. Furthermore, depending on the age of the learners, the majority of the teachers employed different techniques that seemed appropriate to the learners' growing literacy and language levels. These findings indicate what the teachers did. However, they did not provide insights into how frequently or diversely AfL was used in TEYL classrooms. To better understand how the teachers put into practice their understanding of AfL in TEYL classrooms, it is of interest to this study to gain insights into the frequency and diversity in the use of AfL. The findings from that analysis are presented in the next section.

4.3.3 Frequency and diversity in the use of AfL and changes over time

This section reports on the findings about the frequency and diversity in the use of AfL techniques in the study classrooms. Specifically, it addresses the following questions:

Q1: How frequently did teachers use AfL in TEYL lessons? Was there a between-teacher variance in terms of frequency of using AfL?

Q2: How many types of AfL techniques did each teacher use (diversity of use)?

Q3: What, if any, were the changes in the frequency and diversity in the use of AfL over time?

Eight measures were employed to investigate the foci of Questions 1-3 above. The frequency of using AfL (Q1) was investigated with four quantitative measures. A further two measures provided information about the diversity of use by individual teachers (Q2). The remaining two measures offered insights into the changes which occurred in the frequency and diversity over time (Q3). The descriptions of all measures and a list of data sources, on which they were based, are reported in Table 4.5 below.

Table 4.5: Measures used for the quantitative analysis of the frequency and diversity in the use of AfL

What is measured?	Measure data source	Description
	CROSS-SECTIONAL PERSPECTIVE	
Inter-lesson frequency	**InterLFRC** ROWDs	Frequency count of AfL techniques self-reported in all lessons (n=448[13])
Intra-lesson frequency	**IntraLFRC** ROWDs	Average number of self-reported AfL techniques used per lesson in all lessons (n=448)
	InraLFRCEx0 ROWDs	Average number of self-reported AfL techniques used per lesson, excluding lessons where no AfL techniques were used (n lessons =194 out of 448)
	InraLFLOC Lesson Observations	Average number of AfL techniques in lesson observations (n=28, a subset of the 448 lessons)
Diversity	**DivRC** ROWDs	The number of different AfL techniques self-reported by each teacher in all lessons (n=448: 56 lessons per teacher)
	DivLOC Lesson Observations	The number of different AfL techniques used by each teacher observed in the lessons (n=28)
	LONGITUDINAL PERSPECTIVE	
Inter-lesson frequency	**InterLFDTQL** Delayed Teacher Questionnaire	Self-reported frequency of use of AfL (delayed, October 2013)
Diversity	**DivDTQL** Delayed Teacher Questionnaire	Self-reported diversity of use of AfL (delayed, October 2013)

<u>**Key to codes:**</u>
InterLFRC – inter-lesson frequency of use of AfL reported in ROWDs, cross-sectional perspective
IntraLFRC – intra-lesson frequency based of use of AfL reported in ROWDs, cross-sectional perspective
InraLFRCEx0 – intra-lesson frequency of use of AfL, excluding lessons when AfL was not reported in ROWDs, cross-sectional perspective
InraLFLOC – intra-lesson frequency based on use of AfL observed in lessons, cross-sectional perspective
DivRC – diversity in use of AfL reported in ROWDs, cross-sectional perspective
DivLOC - diversity in use of AfL observed in lessons, cross-sectional perspective
InterLFDTQL – inter-lesson frequency of use of AfL, reported in Delayed Teacher Questionnaires, longitudinal perspective
DivDTQL – diversity in use of AfL, reported in Delayed Teacher Questionnaires, longitudinal perspective

---

[13] 448 lessons were reported in school documents called Records of Work Done (ROWDs). These documents were included in the data set of this study. There were equal numbers of lessons reported in ROWDs from both age groups. A subset of 28 lessons (14 in each age group) out of the 448 lessons were observed.

The terms *frequency* and *diversity* of use are central to presenting the findings in this section. *Frequency* is understood as a measure of how often AfL techniques were used in TEYL classrooms. It was investigated on two levels:

1- In how many lessons, out of the total 448, was AfL used? (*the inter-lesson frequency*)

2- How many AfL techniques were used on average in one lesson? (*the intra-lesson frequency*)

Both inter- and intra-lesson frequency will be presented for the whole sample and for individual teachers.

The term *diversity* is used with reference to the number of different AfL technique types. For example, a teacher who implemented four different technique types within a lesson was considered here to have used AfL in a more diverse way than a teacher who used the same AfL techniques four times. In this example, the *frequency* of use would be the same but the *diversity* would differ. Given the two criteria (frequency and diversity), four possible types of implementation were theoretically possible: high frequency/high diversity, low frequency/low diversity, high frequency/low diversity and low frequency/high diversity. Employing these two criteria as measures facilitated the potential for gaining more nuanced insights into the implementation of AfL in TEYL.

The following section reports on the findings about frequency and diversity from Phase One of the data collection: the cross-sectional phase.

4.3.3.1 Frequency and diversity in the use of AfL – a cross-sectional perspective

The aim of this section is to report on the findings about the frequency and diversity in the use of AfL gained through applying descriptive statistics to the data from 448 lessons recorded in ROWDs and, separately, to 28 lesson observations. This section reports on the findings gained from the six cross-sectional measures presented in Table 4.5. First, the findings about inter-lesson frequency are discussed, then those about intra-lesson frequency, and finally those about diversity.

4.3.3.1.1 Inter-lesson frequency in the use of AfL

The inter-lesson frequency measure indicates in how many out of the total 448 lessons the teachers reported the use of AfL. It was investigated with a frequency function on SPSS v19. The result is reported in Table 4.6.

Table 4.6: AfL - inter-lesson frequency

Number of AfL techniques used within a lesson	Frequency (number of lessons)	Percent (%) (100% = 448 lessons)
0	254	56.7
1	49	10.9
2	59	13.2
3	69	15.4
4	17	3.8
Total	448	100.0

The main purpose for investigating inter-lesson frequency was to establish in what proportion of the total number of lessons the teachers used AfL. As is evident from the frequency count in Table 4.6 above, no use of AfL was reported in 56.7% of the lessons. Another observation indicated that one or two AfL techniques were used in a little over 24% of the lessons. Three or more AfL techniques were used in just under 20% of the lessons. Overall, this cohort of teachers used AfL in less than one out of every two lessons on average. It was interesting to see if the same was true for individual teachers. Hence, frequency counts were calculated for all eight teachers in the study individually (see Appendix 21).

The individual frequency counts indicate that the inter-lesson frequency in the use of AfL differed between individual teachers in the study. The teachers could roughly be divided into those who used AfL frequently (T1, T5), moderately (T3, T6, T7) and rarely (T2, T4, T8). The frequency counts for individual teachers (Appendix 21) reveal that T1 and T5 used AfL in almost every lesson, 89% and 91% respectively; that T7, T3 and T6, used AfL in 50%, 38% and 31% of their lessons, respectively; and that T4, T2 and T8 recorded the use of AfL in 23%, 18% and 5% of their lessons, respectively. This finding indicated that there existed between-teacher variance in the inter-lesson frequency of using AfL. Having examined the frequency of use across lessons, the analysis focused on investigating how frequently the teachers implemented AfL within lessons: intra-lesson frequency. This is examined in the following section.

4.3.3.1.2 Intra-lesson frequency in the use of AfL

The descriptive statistics applied to data from ROWDs to investigate the intra-lesson frequency included median, mean and standard deviation. The values were calculated for the whole cohort and for individual teachers. Each teacher reported on fifty-six lessons so 448 lessons in total were analysed. The numerical outcomes are presented in Table 4.7 and discussed below.

Table 4.7: AfL - intra-lesson frequency (all lessons)

Teacher	Number of lessons	Minimum	Maximum	Median	Mean	Std. Deviation
T1	56	0	4	3.00	2.68	1.177
T2	56	0	2	.00	.21	.494
T3	56	0	2	.00	.66	.900
T4	56	0	2	.00	.30	.601
T5	56	0	4	3.00	2.62	1.019
T6	56	0	3	.00	.73	1.168
T7	56	0	2	.50	.63	.702
T8	56	0	1	.00	.05	.227
All lessons in ROWDs	448	0	4	.00	.99	1.288

The median values were larger than zero for only three teachers. The value of 3.00 for T1 and T5, meant that they implemented 3 or more AfL techniques in at least half of their lessons. A median of .50 for T7 indicated that the teacher implemented one AfL technique in at least half of her/his lessons. The medians for all the remaining teachers confirmed that there was a large percentage of lessons in which no AfL was reported but did not offer insights into the intra-lesson frequency of use.

The values which provided insights into intra-lesson frequency for all the teachers were the means and standard deviations. The highest means were obtained for T1 and T5, confirming that they used 3 AfL techniques in the lesson on average. These were accompanied by standard deviation values of around one. This meant that in the majority of lessons the difference between the mean value and the actual recorded value was close to one. Hence, it can be concluded that T1 and T5 reported between 2-4 AfL techniques in each of their lessons. Lower mean values were obtained for the remaining teachers. The mid range values of .73 (T6), .66 (T3) and .63 (T7) indicated that these teachers used AfL more often on average than the remaining three. The lowest values were .30 (T4), .21 (T2) and .05 (T8). For those six teachers (T2, 3, 4, 6, 7, 8), the standard deviation

values were larger than the means. This indicated that many cases fell far from the mean. The same was true for the values obtained for the whole sample of 448 lessons. The mean of .99 was accompanied by a standard deviation of 1.288. Hence, the mean values calculated above were not very informative. A different analysis was needed to supplement those insights and provide a more detailed understanding of intra-lesson frequency.

The subsequent analysis aimed to investigate the intra-lesson frequency of AfL in the lessons in which AfL was used. As was evident from the analysis of the inter-lesson frequency in Section 4.3.3.1.1, six out of eight teachers used AfL in 5-50% of their lessons. Hence, the mean values calculated for all fifty-six lessons taught by any of the teachers in that group included a large number of lessons in which AfL was not used. While this is informative in itself, it does not provide an insight into how many AfL techniques were used in the lessons where AfL was actually implemented. It is believed here that understanding how AfL is implemented is possible by analysing the lessons in which it is in fact used. For that reason, the lessons with 'zero' AfL techniques were removed from the data set and the median, mean and standard deviation values were calculated for the new data set. This choice seems further justified by the standard deviation values being larger than the means in the findings from all fifty-six lessons per teacher. This indicated that the mean values did not represent cases well. The outcomes of the second quantitative analysis are presented in Table 4.8 and discussed below.

Table 4.8: AfL - intra-lesson frequency (excluding lessons where no AfL was used)

Teacher	Number of lessons in which AfL techniques were recorded	Minimum	Maximum	Median	Mean	Std. Deviation
T1	50	1	4	3.00	3.00	.756
T2	10	1	2	1.00	1.20	.422
T3	21	1	2	2.00	1.76	.436
T4	13	1	2	1.00	1.31	.480
T5	52	1	4	3.00	2.83	.585
T6	17	1	3	2.00	2.41	.618
T7	28	1	2	1.00	1.25	.441
T8	3	1	1	1.00	1.00	.000
All lessons in which AfL was used	194	1	4	2.00	2.28	.941

The number of lessons in which AfL was used for each teacher was different in this analysis and varied between three and fifty-two out of the total fifty-six. As is evident from Table 4.8, T8 reported using AfL in only three out of the fifty-six lessons, and used one technique in each of those lessons. This was an extremely limited use indeed. In fact, it would be challenging to argue that T8 used AfL at all. The median values for individual teachers indicated that in 50% or more of their lessons, three teachers (T2, T4, T7) used one AfL technique, two teachers (T3, T6) used two and the remaining teachers (T1, T5) used three AfL techniques. The mean values for T2, T4 and T7 suggest that a little more than one technique per lesson was used on average by those teachers. T3 reported using just under two techniques on average, T5 and T6, over two on average, and T1, three techniques on average per lesson in which they implemented AfL.

Overall, the data indicate that two of the teachers (T1, T5) recorded high inter- and intra-lesson frequency and three (T2, T4, T8) low inter- and intra-lesson frequencies. One teacher reported medium (50%) inter- but low intra-lesson frequency (T7). The remaining two teachers (T3, T6) reported low inter- lesson frequency 30-37% but relatively high intra-lesson frequency of using AfL. These outcomes confirmed between-teacher variance in the frequency of using AfL. The findings from the whole sample of lessons are illustrated with box graphs in Figure 4.2 below. These demonstrate that T1 and T5

used AfL techniques most frequently: the interquartile range between 2 and 3 (highest of all teachers) indicates that 50% of the cases fell within that range and the box graphs are placed higher on the y-axis compared to all others. The box charts illustrate visually what statistical data showed above, i.e. that T7 used AfL in a greater number of lessons than T3 and T6 but used a smaller number of techniques within each lesson. Notably, the small number of cases reported by T2, T4 and T8 did not allow for box graphs to be created.
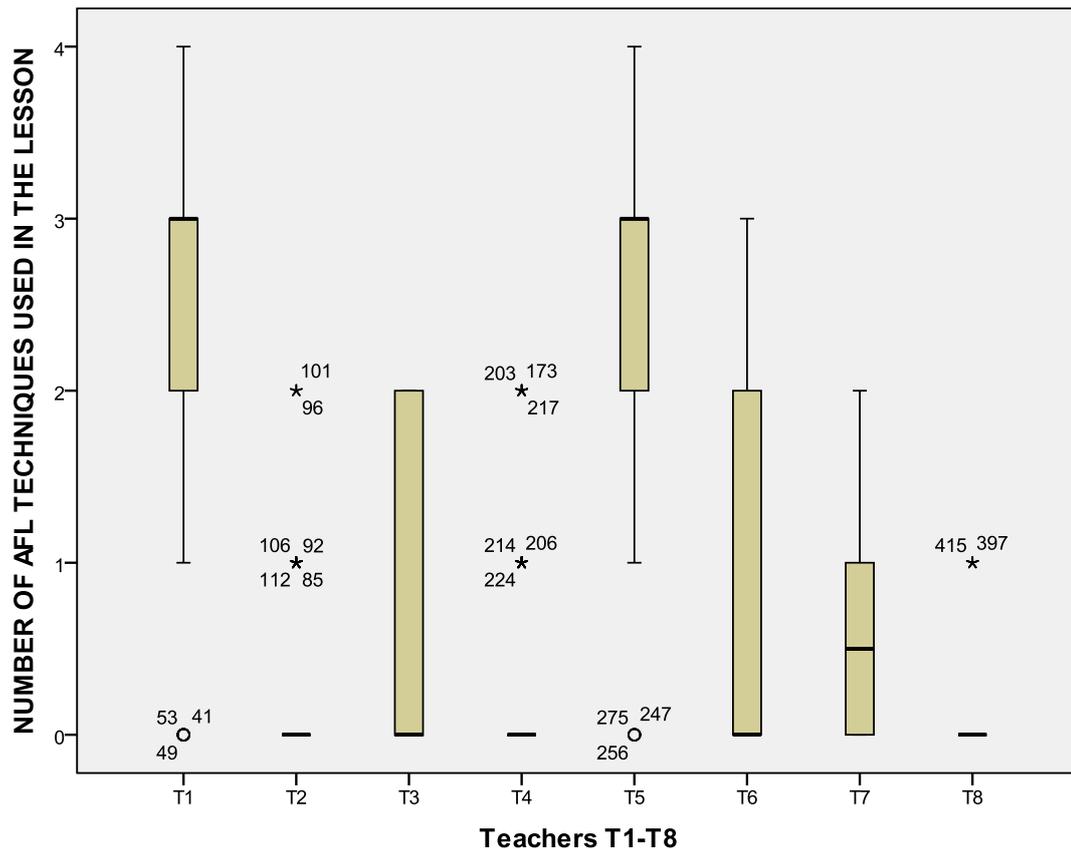
Figure 4.2: Box chart of the frequency of using AfL

The findings discussed so far are based on the data collected from ROWDs, which were accounts of lessons self-reported by the teachers. As such, the ROWDs were the teachers interpretations of what happened in the lessons. In order to gain empirical data about the use of AfL, the study design included lesson observations.

The same descriptive statistics were used to analyse the data from the lesson observations. It is recognised here that the quantitative analysis from the lesson observations may not be very informative on their own, given the relatively small sample of 28 lessons in total. However, it is believed that reporting these values was an important aspect of the analysis as it allowed for verifying findings from self-reported lessons in ROWDs with empirical

data from lesson observations. This was especially useful as the size of the self-reported sample from ROWDs (448 cases = lessons) allowed for meaningful statistical analysis.

The data from the lesson observations did not allow gaining insights into inter-lesson frequency because the number of lesson observations of each teacher was too small. Hence, only the intra-lesson frequency was analysed. The results are summarised in Table 4.9 below.

Table 4.9: AfL - intra-lesson frequency (lesson observations)

Teacher	Number of observed lessons	Minimum	Maximum	Median	Mean	Std. Deviation
T1	4	2	5	4.00	3.75	1.500
T2	4	1	2	1.50	1.50	.577
T3	4	2	4	2.00	2.50	1.000
T4	2	1	1	1.00	1.00	.000
T5	4	2	6	4.50	4.25	1.708
T6	4	1	4	2.00	2.25	1.500
T7	2	2	3	2.50	2.50	.707
T8	4	0	1	1.00	0.75	.500
All lessons in which AfL was used	28	0	6	2.00	2.39	1.571

The use of AfL was recorded in twenty-seven out of the twenty-eight observed lessons. Overall, the values calculated from the data gathered through lesson observations were slightly higher than those obtained through the ROWDs analysis. This is unsurprising given the smaller sample of twenty-eight observed lessons compared to 448 lessons in ROWDs. This might also have been due to the Hawthorne effect (Cohen el al., 2007) as lesson observations were overt and the teachers knew that the study was about AfL but not what the focus of the observations was. Importantly, the mean and median values obtained from the lesson observation data confirmed between-teacher variance in intra-lesson frequency.

The findings from the lesson observations also confirmed which teachers used AfL most frequently (T1, T5), moderately (T3, T6, T7) and least frequently (T2, T4, T8). Figure 4.3 illustrates this finding graphically.
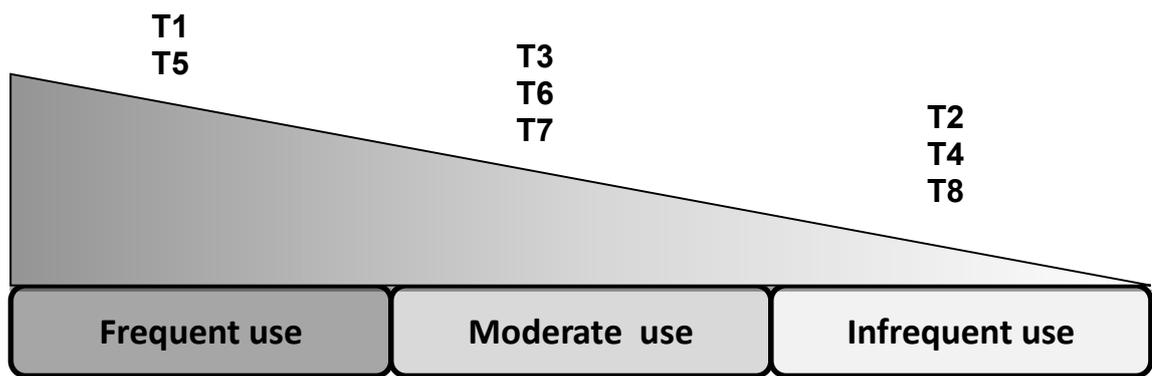
Figure 4.3: A sliding scale of the frequency of using AfL in TEYL classrooms

The findings presented so far strongly suggested that individual teachers differed in the frequency of using AfL. However, they did not provide information about the diversity in AfL technique type which the teachers implemented. Without that insight, it would be possible to claim, for example, that the teachers used just one AfL technique type. Such use would be rather limited in scope regardless of its frequency. Hence, to better understand how AfL was used in TEYL classrooms, it was of interest for this study to investigate whether teachers employed diverse AfL techniques. This focus is explored in the following section.

4.3.3.1.3 Diversity in the use of AfL in TEYL classrooms

Two measures, DivRC and DivLOC[14], were used to gain insights into the diversity of technique type. The values reported in Table 4.10 below are the total numbers of different technique types used within each data set (ROWDs and lesson observations). For example, if Success Criteria were recorded twenty times by one teacher in ROWDs, this was counted as one type of AfL technique.

---

[14] These measures are defined in Table 4.5 on page 163
DivRC – Diversity in the use of AfL reported in Records of Work Done, cross-sectional perspectives
DivLOC – Diversity in the use of AfL observed in lessons, cross-sectional perspectives

Table 4.10: The number of diverse technique types used by teachers

Measure: Data source	Measure description	Number of different AfL techniques							
Teachers		T1	T5	T6	T3	T7	T2	T4	T8
DivRC: ROWDs	The number of different AfL techniques self-reported in 448 lessons	14	12	10	9	5	3	3	2
DivLOC: Lesson Observations	The number of different AfL techniques observed in 28 lessons	8	10	5	5	3	3	3	2

The analysis suggests that the teachers who rarely used AfL tended to employ a limited number of technique types, while those who used AfL more frequently also varied the types of techniques they implemented. Specifically, the findings confirm that the use of AfL by T2, T4 and T8 was indeed rather limited with T7 implementing AfL with slightly more diversity. The remaining teachers tended to implement a large number of different AfL technique types (9-14). Similar pictures emerge from the self-reported data in ROWDs (Measure: DivRC) and the observational data from the lessons (Measure: DivLOC), thus making the findings more credible by verifying one another.

4.3.3.1.4 Summary of the findings about the frequency and diversity of using AfL

The data collected from the school documents and the lesson observations suggested that teachers implemented AfL with different frequency and diversity. The extremes of both frequency and diversity spectrums, shown in Figure 4.3 and Table 4.10 above, relate to the same teachers. Specifically, the teachers who implemented AfL with high frequency also used the greatest diversity of technique types. Those who implemented AfL rarely used a low diversity of technique type. The data concerning the remaining teachers present a somewhat more complex picture. They show that T7 implemented AfL with medium inter-lesson frequency but limited intra-lesson frequency and diversity; and that T6 and T3 implemented AfL with limited inter-lesson frequency, high intra-lesson frequency and diversity. In this cohort of teachers, it seems that greater diversity seemed to occur in the practice of those teachers who deployed AfL with high intra-lesson frequency. However, given the small sample of teachers (n=8) in this study, data could not be collected to enable testing whether this could also be true beyond this sample. Gaining insights from larger samples of teachers would be a worthwhile future research focus. This could have significant implications for teacher development programmes and the implementation of AfL as it would then indicate whether training programmes should

aim to develop teachers' knowledge of various technique types in order to inform implementation.

Additionally, the whole data set was split into two subsets based on learner age. The same quantitative analyses used for the entire data set were employed with each subset separately. No differences were observed between the frequency and diversity of using AfL with classes of 7-9 and 10-11 year olds. This meant that teachers who used AfL frequently with younger children also did so with older age groups. This was a useful finding as it suggested that the teacher and not the age of the students could be an important factor in impacting on how AfL is implemented in a TEYL context. This has important pedagogical implications as it highlights the significance of teachers as mediators of assessment practices.

The current section has discussed findings that strongly suggest that the teachers implemented AfL in TEYL classrooms with different frequency and diversity and did so similarly in the two age groups in the study. However, the data reported in this section provided only a snapshot of the use of AfL in a TEYL context. It did not offer any insights into changes over time. The following section takes a longitudinal perspective on how AfL was implemented in a TEYL context.

4.3.3.2 Frequency and diversity in the use of AfL – a longitudinal perspective

Incorporating the longitudinal perspective was important for this study as it facilitated arriving at a more comprehensive picture of how AfL was implemented in TEYL. It involved gaining longitudinal data about the frequency and diversity of using AfL and the way in which AfL techniques were implemented: purpose, type of language skills etc. The data were extracted from the responses to the delayed teacher questionnaire (Appendix 13) that was administered to teachers in October 2013: i.e. sixteen months after the main data collection stage finished in May 2012.

The main aim of the questionnaire was to gain insights into how teachers continued to use AfL sixteen months after the study had ended. In Question 2 (Q2) of the delayed questionnaire, teachers were asked to indicate how often they used different AfL techniques on a scale from 1 (*never*) or (*almost never*) to 5 (every lesson) or (*almost every lesson*). Employing such a scale was justified by the findings from the cross-sectional data. Namely, teachers on the left of the frequency continuum, as shown in Figure 4.4

below, reported using AfL techniques in almost every lesson, while those at the extreme right of the frequency continuum, reported using them almost never. The list of AfL techniques was sourced from the findings presented in Part One of the current chapter. However, space for the respondents to add other techniques, if they had used them, was provided in the questionnaire. Figure 4.4 below presents the results of the analysis of the answers to Q2. It indicates that three teachers (T1, T5, T7) reported frequent use of AfL; two (T2, T4), moderate use; and three (T3, T6, T8), infrequent use. The distribution of the teachers on the frequent-infrequent use continuum resembled the one obtained in May 2012. However, importantly the positions of five of the teachers on the continuum were different.
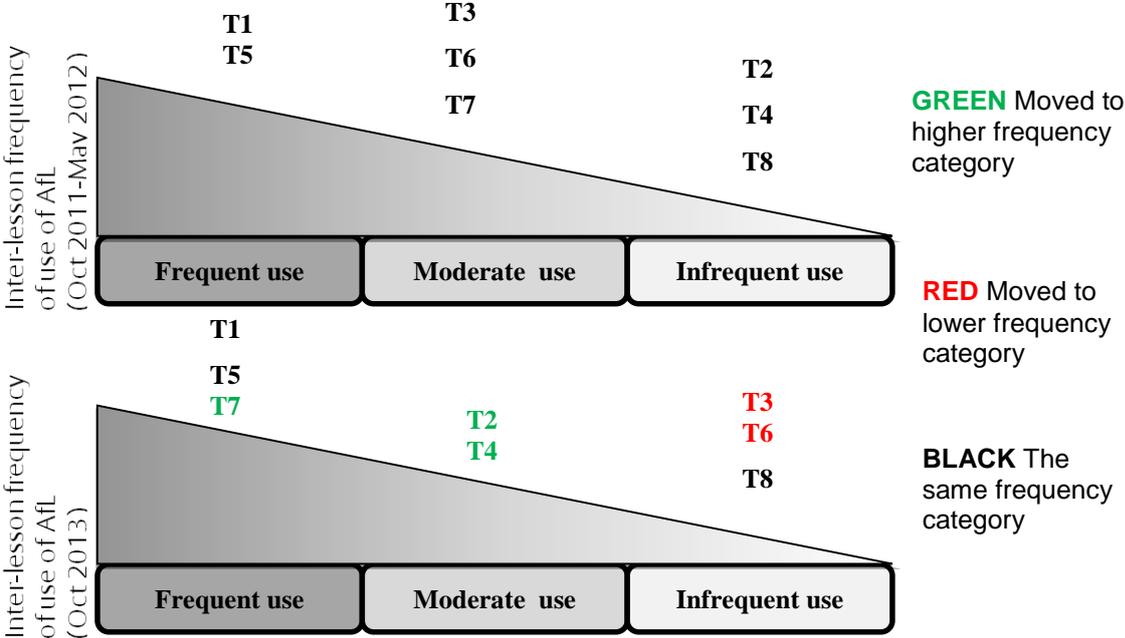
Figure 4.4: A sliding scale of the frequency of using AfL showing changes over time

The teachers' positions on the frequency continuum from May 2012 (top half of Fig. 4.4) and October 2013 (bottom half) were compared in order to provide insights into changes in the frequency of implementing AfL techniques as reported by the teachers. The findings indicate that all the teachers who were previously in the mid-range of frequency changed their position on the scale; T7 moved to the 'frequent use' category, while T3 and T6 both moved to the opposite side of the continuum. The teachers who used AfL frequently in 2012, i.e. T1 and T5, remained in that category, while T2 and T4 moved from the 'infrequent use' to the 'moderate use' category. T8 remained in the 'infrequent use' category. Overall, this finding confirmed between-teacher variance in the use of AfL

but also indicated that many teachers reported that they changed frequency in the use of AfL.

The measure DivTDQL[15] was applied to the data collected from Q2 on the delayed questionnaire to gain insights into the diversity of techniques used by the teachers sixteen months after the end of the study. The techniques used by the teachers in terms of frequency within a lesson were considered representative of the diversity of use. In other words, if a teacher used the given technique type sometimes, often or almost every lesson, it was possible to say that that particular technique added to the diversity of the technique type in the practice of that particular teacher. Whereas the technique types that the teachers reported as never, almost never or rarely being used were considered as not used often enough to provide diversity. For this reason the techniques in the 'never' and 'rarely' categories were excluded from this analysis. The total number of different technique types was recorded for each teacher. This was compared with the numbers obtained from the cross-sectional phase. The results indicated that there was a greater diversity of technique type in October 2013 compared to May 2012 (Table 4.11). The majority of teachers used quite a large number of different techniques (8-15) and only one teacher (T8) used markedly fewer (4). The findings suggested that the diversity of technique type increased over time.

Table 4.11: Diversity of technique type used by teachers one academic year later

	T5	T1	T4	T7	T2	T3	T6	T8
Number of technique types recorded in ROWDs in the cross-sectional phase (May 2012)	12	14	3	5	3	9	10	2
Number of technique types recorded in the longitudinal phase (October 2013)	15	14	14	13	10	8	8	4

The results from the cross-sectional and longitudinal phases pertaining to frequency and diversity were compared. They show that one teacher, T8, reported using AfL with limited frequency and diversity. It could in fact be argued that this particular teacher did not use AfL enough to offer useful insights into the implementation of AfL. Nevertheless, identifying T8 as an outlier was useful as it indicated that there could be factors that might inhibit implementing AfL in TEYL classes. This issue warrants future research.

---

[15] This measure is defined in Table 4.5 on page 163
DivTDQL – Diversity in the use of AfL reported in the Delayed Teacher Questionnaire: longitudinal perspective.

While comparing the values of diversity in technique type (Table 4.10, p.171) and places on the frequency continuum from May 2012 and October 2013 (Fig. 4.4), it was observed that the majority of teachers reported using AfL with medium to high frequency and all but one used it with medium to high diversity. This finding is summarised in Figure 4.5 below.

	May 2012			October 2013		
	High Frequency	**Medium Frequency**	**Low Frequency**	**High Frequency**	**Medium Frequency**	**Low Frequency**
High Diversity	T1 T5	T3 T6	-	T1 T5 T7	T2 T4	-
Medium Diversity	-	-	-	-	-	T3 T6
Low Diversity	-	T7	T2 T4 T8	-	-	T8

Figure 4.5: Frequency and diversity in the use of AfL: longitudinal perspective

The above finding could be interpreted in two ways. First, it seems possible that, as teachers became more experienced users of AfL, they became more confident in implementing a wider range of technique types. A second interpretation could be that as they became more experienced in using AfL, some teachers observed that a high diversity in technique type was more appropriate for TEYL classes than a low diversity. The latter interpretation corroborates with the finding that the teachers reported (e.g. Appendix 19, Turns 8-10) that diversity in AfL technique type was needed in order to sustain interest and engagement (RQ1).

Fully investigating the reasons for the occurrence of the changes in frequency and diversity of use is beyond the scope of this study. However, some useful insights into this area have been gained through the analysis of the answers to Questions 1, 3 and 4 of the delayed teacher questionnaire.

Q1 of the questionnaire gathered information about the continuous professional development (CPD) that the teachers participated in between May 2012 and October 2013. The aim was to find out if there was a relationship between the use of AfL and CPD

activities. The outcomes of this analysis are presented in Appendix 22. The findings indicate that:

1. The teachers who reported starting to use AfL more frequently or remained in the 'frequent use' category had all observed AfL being used in lessons by colleagues. The remaining teachers had not.

2. The teachers who moved from the 'infrequent' to the 'moderate' use category had participated in AfL training specifically.

3. The teachers who had participated in a different form of assessment training used AfL less frequently over time.

Further insights into the reasons for change in the frequency of using AfL between May 2012 and October 2013 were gained from the qualitative analysis of the teachers' answers to the open-ended questions: viz. 3 and 4. The teachers were asked to provide their accounts of why they had selected those techniques that they had implemented most frequently. The results indicate that familiarity with a technique type and their perceived ease of implementation were important considerations. Table 4.12 summarises the outcomes of the analysis of the open-ended questions.

Table 4.12: Reasons for changes in diversity and frequency over time

	Teachers' reasons for choosing to use AfL	Teachers' reasons for choosing NOT to use AfL
High ↑ **Frequency and diversity in AfL use** ↓ **Limited**	• **Familiarity with AfL techniques gained through training, observations and experience of using AfL** *'I chose the techniques I used most often because I had learned about them from input sessions and from further background reading. I tried them and they worked best for me and my students.'* (T1/Q3) • **AfL techniques were an integral part of the teaching style** *'Some of the techniques I use are an essential part of my teaching style and I do not necessarily make a conscious decision to use them.'* (T7/Q3) • **AfL was considered an effective tool to focus students on expectations of outcomes and on students' own achievement** *'They* [AfL techniques] *give students an appreciation for what was achieved in the lesson as well as a sense of accomplishment. It is extremely simple and effective, and they can understand how to improve.'* (T2/Q3) • **Satisfaction with the amount of AfL that teachers used** *'I thought they* [the AfL techniques that this teacher used] *did the job and there was no need for other techniques.'* (T8/Q3) • **Ease of use** *'I used these techniques because these are the ones that I am most happy with, and I find them easy to set up and use.'* (T3/Q3)	• **Lack of familiarity with a given technique** *'The main reason that I've not used these techniques is that I'm (or was) unfamiliar with them.'* (T3/Q4) *'I may not have seen them* [the AfL techniques that I do not use] *in action. I usually have to witness the effectiveness of a technique first hand in a language lesson in order to gain a full appreciation for it.'* (T4/Q4) • **Time constraints (on preparation and within lessons)** *'This could be time consuming to prepare.'* (T6/Q4) *'Colour coding is sometimes too time consuming with small kids.'* (T5/Q4) *'The idea of learning partners was not particularly practical due to limited time in class.'* (T7/Q4)

The findings indicate that there were a number of factors that contributed to the teachers choosing whether to use AfL or not. Notably, all the reasons seem teacher-focused. That is, they took into account the teachers' preferences, experiences or knowledge. None of the reasons quoted by the teachers were learner-centred. This might imply that the teachers and their beliefs were important factors in how AfL was implemented.

The findings about the reasons for choosing to use AfL indicate that familiarity with AfL techniques, compatibility with their teaching practice and the perceived effectiveness were among the factors that impacted on whether the teachers used AfL or not. It is worth noticing that the teachers who used AfL with medium to high diversity and frequency tended to share their reasons for choosing to use AfL. These included the beliefs that AfL facilitated raising learner awareness of achievement and was compatible with the teaching methods used. It is also important to note that the whole cohort of teachers indicated there were two reasons for not including some of the AfL techniques into their practice: lack of familiarity and time constraints. These reasons confirmed that familiarity with AfL techniques was an important factor in implementation and that practical constraints, like lack of time, can inhibit implementation.

Another interesting longitudinal insight referring to the use of AfL in TEYL classes was drawn from the responses to Q5 of the delayed questionnaire. The teachers were asked to choose three of the AfL techniques that they had identified in their responses to Q2 as most commonly used and to report on which language skills or types of tasks they had used them with. The data confirm that AfL was used with the productive skills: speaking and writing. This finding is similar to those obtained from the cross—sectional phase. Hence, it provides evidence that the types of activities that AfL was used for in TEYL classes did not change over time. The teachers' answers are summarised in Table 4.13 below.

Table 4.13: Use of AfL with task types and language skills reported by the teachers in the delayed questionnaire

AFL[16]	Skills/ types of activities
WALT	Introduction to the lesson (T1, T2, T5, T7) With any activity to keep learners focused on what they are learning (T1, T7)
SC	Writing (T2, T4, T8) Arts and crafts (T2, T6) Speaking (T4, T5) Classroom management (T4, T6)
ICS	With any activity, to introduce or reflect on the learning objective (T1, T2, T3, T6)
LP	Speaking (T3, T4, T8)
SF	After any task (T4, T8)
TSAW	Writing (T5, T6)
CMWE	Writing (T7)
ITT	Speaking (T7)

The final insights provided by the answers to Q6 relate to the impact of AfL that the teachers had observed in their classes. The quotes provided in this section to exemplify the teachers' comments were copied verbatim from the written answers to the delayed questionnaire. Hence transcribing convention does not apply to them. Most teachers indicated that:

1. The learners were more aware of what they were learning and of how to improve, thus the formative function of assessment could be realised:

    'As for teaching, giving instructions became more efficient and assessing students' work too. I was amazed how honest the kids were about their own results.' (T5/QUESTIONAIRE)

    'I think they have a better idea of their own capabilities after they complete a can-do, whereas, after a multiple choice test they only know that they've passed or failed.' (T6/ QUESTIONAIRE)

    'They know what to do and miraculously somehow are able to do it. Which is not to say that they do not make mistakes. Mistakes happen but at least there is an opportunity for them to make mistakes and get those corrected and not as I

---

[16] Detailed descriptions of all AfL techniques identified in the current study are provided in Appendix 18.

remember my pre-AfL teaching, you just didn't know what your student knew because it was hard to get anything out of them.' (T3/ QUESTIONAIRE).

2. The learners were able to take responsibility for their own learning.

'When I use these techniques, my students seem to notice the purpose of each lesson and they feel more responsible for their own learning. They take pride in their achievement. As their involvement increases, they enjoy the lessons more and, obviously, learn more.' (T7/ QUESTIONAIRE)

'Before using success criteria with writing, I only used to set writing tasks as homework, which almost invariably meant that not all students completed the task. I had the impression that students in that age range (9-11) would view writing as "boring". However, they have been motivated by doing the writing tasks in this way. It has also taught the students a sense of responsibility as they do peer-correction.' (T3/ QUESTIONAIRE)

The current section has provided a longitudinal perspective to the findings about the use of AfL in TEYL. The findings indicate that over time the majority of the teachers implemented a greater diversity of AfL techniques. Finally, the data confirm that AfL was used predominantly when teaching productive skills.

4.3.4 Summary of findings for Research Question Two

This part of Chapter 4 discussed the findings that address RQ2 (2.1 and 2.2). The analysis of the data collected through lesson observations, school documents (ROWDs) and the responses to the delayed questionnaire resulted in the findings reported below.

Findings to RQ 2.1:

1. Eighteen AfL techniques were observed or self-reported by teachers to be used in a TEYL context.
2. AfL techniques served three main purposes in 7-11 year olds classes: sharing learning aims, providing feedback from teachers, learners and peers as well as measuring learners' confidence levels.
3. AfL techniques served the purpose of giving and clarifying instructions in the younger age group 7-9.

4. The majority of eachers tended to use AfL with the productive skills (writing and speaking) as opposed to the receptive skills (listening and reading).

5. There was between-teacher variance in inter- and intra-lesson frequency in the use of AfL.

6. There was between-teacher variance in the diversity of technique type implemented in TEYL lessons.

7. Four types of implementing AfL with TEYL classes were identified: high frequency and diversity; low frequency and diversity; low inter-lesson frequency but medium level of intra-lesson frequency and diversity; and medium level of inter-lesson frequency and low intra-lesson frequency and diversity.

Findings to RQ2.2:

1. Over time, the teachers who used AfL with moderate frequency tended to move towards either of the extremes of the continuum of frequency in the use of AfL.

2. Over time, most of the teachers tended to use a greater diversity of technique type.

3. The familiarity and experience of using AfL were indicated as factors that affected the implementation of AfL.

4.4 Part Three: The impact of AfL on classroom interactions

4.4.1 Introduction to Part Three

Research Question Three sought to understand the observable impact of AfL on interactions in TEYL classes. The data came from the lesson observation field notes and the transcripts of extracts from video-recorded lessons. This part of Chapter 4 commences with a report of the findings about the relationships between the use of AfL and the types of interactions that took place during lessons (4.4.2). Then the outcomes of applying the Storch (2002) model and the Variable Approach to the analysis of FL classroom interactions are presented (4.4.3). The final section (4.4.4) summarises the findings and indicates how AfL can be empirically linked to facilitating learning in TEYL classes.

4.4.2 Relationships between AfL and classroom interactions

This section reports on the findings from the quantitative analysis of the relationships between the use of AfL and the types of interactions that occurred in TEYL classes. Bivariate correlations were calculated in SPSS v19. The Pearson-product moment correlation coefficient was used to establish whether any relationship existed between the use of AfL and each of the seven types of interactions observed in the lessons: L-L, T-1L, LL, T-C, T-xLL, L-C, IND (all interval scales). The results suggest that the frequent use of AfL was positively correlated with a large number of L-L and T-1L interactions. They also suggest that a moderate negative correlation existed between T-C interactions and the use of AfL. No correlations were found between the use of AfL and the following types of interactions: IND, L-C, TxLL and LL. Hence, it seems reasonable to conclude that in the classrooms where AfL was used frequently, more one-to-one interactions took place.

In order to investigate if there were differences between the age groups, correlations were also calculated for the 7-9 and 10-11 year olds, separately. The results confirm a strong positive correlation between the frequency of using AfL techniques and a large number of L-L as well as T-1L interactions, both significant to the 0.01 level. However, a moderate negative correlation between using AfL and the number of T-C was only found in the younger group (7-9 year olds) and was significant to the 0.05 level. The Pearson correlation values are reported in Table 4.14.

Table 4.14: Relationships between classroom interactions and the use of AfL: Pearson-product moment correlation coefficients

Measure[17]	L-L	T-1L	T-C
Use of AfL in both groups	.719**	.703**	-.405*
Use of AfL in 7-9 year olds	.707**	.684**	-.561*
Use of AfL in 10-11 year olds	.753**	.731**	-.443
*. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).			

It is important to note that the Pearson-product correlation does not indicate causation but solely co-existence. Hence, the findings could not be interpreted as showing that the use of AfL caused an increased number of one-to-one interactions. What is claimed is that in classes where AfL was used, there were more one-to-one interactions compared with classes where AfL was less frequent. This might mean that there existed conditions in some lessons that facilitated both a large number of L-L and T-1L interactions and the frequent use of AfL. However, the correlation does not exclude the possibility of causation. In order to investigate if the frequent use of AfL did cause a large number of L-L and T-1L interactions, a different study design needs to be developed. This would constitute an interesting focus of future research.

The existence of the correlations presented in Table 4.14 pose an interesting question, viz.: Is there evidence that L-L and T-1L interactions during AfL can support learning? This was addressed by analysing examples of conversations transcribed from video recorded lessons. The transcribing convention was based on Walsh (2006) and is detailed in Appendix 15.

4.4.3 Classroom discourse during use of AfL

This section focuses on the qualitative evaluation of the *richness* in the use of AfL in TEYL classes. *Richness* is understood in the present study as the quality of the interactions that occurred while AfL was being used. This was evaluated through two analyses. First, holistic interaction patterns were analysed by applying the Storch's (2002) model to dyadic interactions. This analysis is reported in Section 4.4.3.1. Secondly, the Variable Approach to L2 classroom discourse analysis (Walsh, 2006) was used (see

---

[17] Interaction codes: L-L (Pairs), T-1L (teacher – individual student), T-C (teacher – whole class), IND (individual work), L-C (learner- whole class), T-xLL (teacher – group of students, e.g. T-4LL), (LL) Groups

Section 4.4.3.2). The aim was to gain insights into whether the use of AfL could support teachers aligning their use of language with the pedagogical aims for each conversation. Additionally, the analysis included making judgements about whether the conversations that occurred while implementing AfL contributed to meeting the lesson aims, and which modes they occurred in. Transcripts of classroom interactions were sourced from video recordings of the observed lessons. All extracts reported in this part exemplify interactions that took place while AfL was being implemented.

The choice to analyse the richness of dyadic interactions was motivated by the findings reported in Section 4.4.2 which indicated that there was a significant positive correlation between the use of AfL and a large number of dyadic interactions of L-L and T-1L type. As the aim of this study is to investigate the impact that AfL could be observed to make on interactions, it seems important to focus the analysis on the richness of the types of interactions that were significantly positively correlated with the use of AfL.

4.4.3.1 Applying Storch's Model to classroom interactions

Storch's (2002) model was used to analyse the holistic patterns of L-L and T-1L interactions that occurred while AfL was being used. The purpose of that analysis was to investigate if the interactions were of the types that had been shown by research to support learning: collaborative and expert/novice (e.g. Butler & Zeng, 2014).

As discussed in Chapter 3, the Storch (2002) model is concerned with the levels of mutuality and equality between interlocutors during a conversation. In the process of the analysis, all conversations were categorised into one of the four possible categories. Details of this analysis can be reviewed in Appendix 23. The results of that analysis indicate that in TEYL lessons all T-1L interactions were observed to have had low to medium equality (see Quadrants 3 and 4 in Figure 4.6 below). This is not surprising and it illustrates that the teachers tended to control the interactions. A less predictable finding is that the majority of the transcribed conversations demonstrated medium to high levels of mutuality (an expert/novice pattern) as opposed to medium to low mutuality (a dominant/passive pattern). This is an interesting finding as it suggests that, while using AfL, YLs assumed active roles during classroom conversations with their teachers. The analysis of L-L interactions indicates that the equality levels were medium to high. This could reasonably be expected due to the context of the study: more specifically, the

makeup of the groups in terms of age and language level. The mutuality levels in L-L interactions were mostly medium to high.

Applying Storch's (2002) model to the dyad interactions that occurred during the use of AfL, revealed that T-1L and L-L displayed different interaction patterns. It was also found that T-1L interactions were mostly of the expert/novice type, while L-L interactions were mostly collaborative. No dominant/dominant interactions were identified. Hence, the findings proposed that T-1L and some of the L-L interactions during the use of AfL displayed holistic patterns that had been shown to facilitate learning (Swain, 2000, Butler & Zeng, 2014). Figure 4.6 below summarises the results of this analysis. The numbers of extracts sourced from the younger age group are underlined.

| | ← Low   EQUALITY → | |
	High	
High ↑ **MUTUALITY** ↓ **Low**	Quadrant 4: Expert/novice **T-1L:** 4.3, 4.5, 4.6, 4.10	Quadrant 1: Collaborative **L-L:** 4.7, 4.8, 4.11
	Quadrant 3: Dominant/passive **T-1L:** 4.4	Quadrant 2: Passive parallel or dominant/dominant **L-L:** 4.9

Figure 4.6: Applying Storch's (2002) model to L-L and T-1L interaction

Applying the Storch's (2002) model reveals which holistic interaction patterns occurred during the use of AfL. However, it did not offer insights into whether the interactions contributed to advancing learning in relation to the learning aims stated for each lesson. To gain that insight, conversation analysis was used.

4.4.3.2 Variable Approach to investigating FL interactions

The second stage of analysing the classroom discourse employed a variable approach to investigating FL interactions. It involved using conversation analysis (CA) methodology and analysing turns and sequences of transcribed speech. The analysis focused on investigating the relationship between classroom discourse and learning. In the variable approach, every conversation is understood as situated in a number of micro-contexts. The context for each turn is created by the preceding one. Research in this area indicates

that the use of language during classroom interactions can facilitate or inhibit learning depending on whether it is congruent with the pedagogical aims or not (Walsh, 2006). For that reason, it was important for the focus of the current study to analyse whether the use of language while implementing AfL offered opportunities for aligning language with the pedagogical aims. Hence the richness of the interactions was evaluated in terms of the congruence of the teacher language use with the perceived pedagogical aims of the conversations. Furthermore, the analysis aimed to investigate whether the perceived pedagogical aims of conversations offered opportunities for contributing towards the aims of the lessons, formally recorded in ROWDs.

This section reports on a number of conversations transcribed from the video recorded lessons. It is important to note that, similar to Walsh (2006), the conversations were recorded 'under normal classroom conditions with no specialist equipment' (p. 165). This meant that there were limitations on how much conversational data could be recorded. However, employing this data collection procedure resulted in obtaining data that were particularly useful to the current study as they offered an opportunity to gain useful insights into conversations naturally occurring in TEYL classrooms and not in an experimental setting. Hence, the transcripts used in the current chapter are examples of classroom interactions but do not contain all the conversations that occurred in the observed lessons.

As discussed in Chapter 3, Walsh's (2006) modes of conversation were used as a framework for the analysis of classroom discourse. The findings are reported in four sections below, one for each of the modes of conversation: managerial, materials, skills and systems, and classroom context.

4.4.3.2.1 The managerial mode of interactions during the use of AfL

This section reports on examples of conversation that occurred in the managerial mode. This mode comprises interactions that happen when teachers set up tasks and provide instructions. It is characterised by long teacher turns, comprehension checks (e.g. Is it clear?) and transitional markers (e.g. right, ok). As could be predicted, in TEYL classrooms there were many instances of the managerial mode. These often happened when teachers were explaining something to the whole class (i.e. during T-C interactions) and included a number of T-1L interactions (Extract 4.3). There were also a number of

T-1L conversations in managerial mode that occurred when teachers were monitoring the learners' independent or group work (Extract 4.4).

EXTRACT 4.3

(7-9 year olds, using Success Criteria and Sharing Good and Bad Model)

[1] **T1:** *so we can see the two texts here yes? one is better and the other is not so*

[2] *good right? Anna* said it has a title what else makes it good?* (1)

[3] **L61:** *the text A is bigger=*

[4] **T1:** *= yes it's bigger* (1) *what do you mean by bigger* (1) *there are lots of* (1)

[5] **L61:** *words=*

[6] **T1:** *=there are longer* (1)

[7] **L61:** *sentences*

[8] **T1:** *sentences* (1) **T underlines a sentence on the board** *can you see?*

[9] *L61: yes longer sentences is better*

[10] **T1:** *good what else makes it* [the text] *good?*

In Extract 4.3, T1 was discussing the expectations for a writing task with one student (L61). The teacher demonstrated two examples of a completed writing task to the class and was eliciting the Success Criteria (SC) from L61. Initially, the teacher accepted the incorrect use of the adjective 'bigger' to describe the text (Turns 3-4). Subsequently, the teacher clarified the meaning in Turns 4-5 and provided a model of a correct adjective, i.e. 'longer', in Turn 6. T1's use of language seems be congruent with the metalanguage i.e. 'the language that children need to talk about and understand talk about grammar and discourse' (McKay, 2006: 6) that L61 needs in order to be able to describe the requirements of the task. T1's aim seems to be to elicit one criterion for successful completion of the writing task from L61. Hence, the use of language seems well aligned with that aim. The clear references to the two model texts and the list of SC provides evidence that using these AfL techniques helped to facilitate this interaction. Additionally, as the pedagogical aim of this lesson was to write an interesting newspaper article about Notting Hill Carnival, this conversation seems to have provided a direct contribution to meeting that aim.

The following extract (4.4) is an example of T-1L interaction that occurred during an independent task. T5 noticed a difficulty that the student, L53, was having in assigning him/herself a Traffic Light.

EXTRACT 4.4

(7-9 year olds, using Traffic Lights)

[1]   **T5:** *ok* (1) *which one are you going to give yourself?*

[2]   **L53:** *what?*

[3]   **T5:** *you think you're green* (1) *yellow or red?*

[4]   **L53: looking at the teacher for 2 seconds**

[5]   **T5:** *do you know eight words?* (1) *do you know eights words* Tom[18]?

[6]   **L53:** *what?*

[7]   **T5:** *do you know eight of these expressions? do you know EIGHT? yes?*

[8]   **L53:** *yes*

[9]   **T5:** *ok then* (1) *do green light* (1) *that's good*

In Extract 4.4, the teacher asked a series of questions in Turns 1, 3, 5 and 7. Each consecutive question seemed to be more specifically indicating what was required from L53. First, T5 asked a general question about how L53 is going to assess their own achievement (turn 10), then (s)he indicated that there were three possible ways of doing it (Turn 3). Turn 5 reminded L53 of how to measure his success and finally, Turn 6 provided a tangible reference list against which to measure the success. The student responded to T5's question in Turn 7 by repeating the final 'yes' uttered by the teacher. It is uncertain from this exchange if L53 actually understood what he was being asked to do or why. This conversation was clearly located in the managerial mode as the teacher was providing guidance on how to conduct self-assessment. The perceived pedagogical aim of this conversation was to support the student in completing the self-assessment. T5's use of language appears to be congruent with that aim as the assessment technique (TL) and criteria (8 words) are consistently and clearly referred to. Furthermore, this conversation seemed to offer an opportunity to contribute to the lesson aim which was to use eight or more phrases about hobbies by allowing L53 to reflect to what degree (s)he met that aim. However, from L53's limited contribution to this conversation, it could not be inferred that this opportunity was used effectively.

The analysis of the examples of classroom discourse in managerial mode indicates that teachers used AfL to raise learners' metacognitive awareness (see Section 2.2.1.3.2 for the discussion on metacognition). The teachers' language use seems to agree with the

---

[18] Pseudonym

perceived aims of the conversations. There was some evidence in Extract 4.3 that the young learners were able to engage in those conversations meaningfully. However, in Extract 4.4 it was also evident that the 7-9 year old students experienced difficulties in understanding metacognitively oriented questions. In both extracts, the learners responded to the teachers' questions when tangible references to an example of a text (4.3) or to a list of words (4.4) were made. The next section discusses examples of conversations in the skills and systems mode.

4.4.3.2.2 The skills and systems mode of interactions during AfL

This section reports on the conversations in the skills and systems mode. This mode is centred around practising and clarifying language rules and meaning. It often follows the IRF (initiation, response, feedback/follow up) pattern. In the current study, the dyadic interactions in skills and systems mode were mostly of the T-1L type and they happened while teachers were facilitating practising the new language: e.g. during the monitoring of group or individual work. The data collected in this study indicated that when AfL was used, these conversations were very frequently initiated by the learners.

The following extract (4.5) illustrates that the micro-contexts and pedagogical aims of conversations can quickly change during one, relatively short conversation. In Turn 1, L81 initiates (I) the conversation, providing the context for T7's initiation (I) of another exchange by addressing the grammatically incorrect sentence structure (Turn 2). This provides a new micro-context to L81, who responds (R) to the teacher's request (Turn 3). The following turn includes the teacher's feedback (F) on the new correct phrase. This is followed by what could be considered the teacher's response to the initial question asked by L81. The answer is not spoken but scribed on the board. This seems to include turns typical for IRF interactions, which often occur in the skills and systems mode. However, in this particular case, there seemed to be an additional IRF sequence embedded in an overarching IR sequence initiated by the learner. This conversation between an expert teacher and a novice learner is an example of a LRE (Section 2.2.2) that focuses on the form of the language that is being used.

EXTRACT 4.5

(10-11 year olds, using Success Criteria)

[1]   **L81:** *how is throwing?*

[2]   **T7:** *give me the correct question* (2)

[3]   **L81:** *how* (1) *do we spell* (1) *the word throwing*

[4]   **T7:** *excellent stuff* **T writes 'throwing' on the whiteboard**

In Extract 4.5, T7's use of language indicates that there was a problematic language form and requested a correction. This was an example of providing explicit negative feedback. The mental work on the part of the student seems significant. First, (s)he had to identify the mistake and subsequently provide a correctly formed question, i.e. modify her/his output. T7's intervention was successful and L81 was able to provide a grammatically acceptable alternative. The difficulty of the process and the amount of mental activity on the part of the student could be inferred from the pauses that occurred in the improved question. L81's aim in this exchange was to find out how to spell 'throwing' and T7's aims were to elicit a correct question structure and provide some support with the spelling. The teacher's use of language does not seem to fit with the aim of eliciting the correct question form. The teacher, T7, did not ask any questions. If T7 had said: 'How do we ask this question correctly?' that would have better aligned with the pedagogical aim. However, it is worth noting that this exchange seemed to have at least partly supported the lesson aim which was *to use was/were + ing to talk about the past*.

In the following extract (4.6), T6 was monitoring pair work while the learners were trying to decide what safety hazards they could spot in a picture of a sport stadium during a game. L102 initiated an interaction with T6 (Turn 1). T6 initially used recast to provide implicit negative feedback (Turn 2) and in doing so clearly aligned his/her language with the pedagogical aim of that conversation, which seemed to be to clarify the rule of how to use 'much' and 'many' with nouns. In Turn 3 (N.B. 'of' in the original recording), L102 noticed the correction, which T6 then successfully reinforced with a series of short questions to clarify the grammatical rule for using 'many' with countable nouns (Turns 4-8). The congruence of T6's language with the pedagogical aim was especially evident in Turn 8, when the teacher modelled a number of different possible phrases.

EXTRACT 4.6

(10-11 year olds, using Success Criteria)

[1] **L102:** *too MUCH fans?*

[2] **T6:** *yes (1) too many fans? (1) yes (1)*

[3] **L102:** *too many (1) of much?*

[4] **T6:** *because 1 fan (1) 2?*

[5] **L102:** *fans*

[6] **T6:** *so that means that much or many?*

[7] **L102:** *many*

[8] **T6:** *many fans (1) so 1 fan (1) 2 fans (1) many fans*

It seems that in the interaction reported in Extract 4.6, following T6's negative feedback, L102's output is modified. This is an example of a LRE between T and L, which focused on grammar. The role of the AfL technique in facilitating the conversation seems evident. One criterion for success was to choose one of the two words ('more' or 'much') and use these in a sentence correctly. As the SC appeared to draw the learner's attention to the language form, thus it might have encouraged the learner to ask the teacher about the accuracy of their sentence. This conversation also directly contributed to supporting the learning aim for the lesson, which was to use '(too) much/many', 'not enough' with countable and uncountable nouns. This extract provides another example that suggests that AfL could provide a background to discussing key issues central to teaching and learning by facilitating the occurrence of LREs.

The dyadic interactions in skills and systems mode that were recorded in TEYL classrooms while AfL was being used, indicate that the teachers use of language was only sometimes congruent with the pedagogical aims of the conversations. However, the analysis indicates that the aims of conversations seem to directly support the learning aims set for the lessons. This is an interesting finding as it provides empirical evidence indicating that, when AfL is used, teachers and students focus their efforts on working directly towards meeting the learning aims. This corroborates with the teachers' beliefs about the use of AfL in TEYL classrooms.

The next section explores conversations that happened in the materials mode.

4.4.3.2.3 The materials mode of interaction during AfL

This section reports on examples of interactions in the materials mode. Notably, this set of interactions often included very short teacher turns or the teacher was silently listening to the conversations between learners or interacting non-verbally with them. Walsh (2006) attributes such teacher participation to the fact that the conversations are guided by the materials which the learners are using.

In Extract 4.7 below, the 7-9 year olds were working with their Learning Partner (LP) to decide what they should do during the following weekend. Their task was to agree on one activity. This short extract demonstrated how quickly the aim of conversation, and indeed the roles adopted by learners, changed.

EXTRACT 4.7
(7-9 year olds, using Learning Partners)

[1] **L33:** cinema

[2] **L40:** let's go to see Harry Potter at half past eight am

[3] **L33:** am? (1) or pm?(1)

[4] **L40:** evening?

[5] **L33:** yeah pm

[6] **L40:** and we will meet at half past eight pm=

[7] **L33:** =pm

In Extract 4.7 multiple aims could be identified: in Turn 2 – L40 aims to develop L33's suggestion by using it in a whole sentence; in Turns 3-7, L33 aims to clarify the suggestion; in Turn 4, L40 checked their own understanding of 'pm' by paraphrasing it to 'evening'. It should be noted that these students did not speak the same mother tongue (L33 spoke Polish; L40, French), hence they could only negotiate meaning entirely through English. This activity was completed while the teacher was monitoring the pair work and listening to the learners' conversations. Although the teacher did not participate in this exchange, one peer, L33, was clearly more capable than the other in knowing the meaning of 'am' and 'pm'. This conversation provides an example of the ZPD in operation as L33, through negotiating the meaning, helped to clarify L40's understanding. An important move, which began the negotiation, was the request for clarification in Turn 3. In that turn, L40 received negative feedback on her/his performance. This led to L40

modifying the output, in Turn 6. By employing the LP technique, the teacher facilitated an opportunity for this conversation to occur. Additionally, it could be observed that the aim of this conversation clearly supported meeting the learning aim of the whole lesson, which was 'to make suggestions using Why don't we..?, Shall we...? and Let's ...'. It is also important to note that the negotiation of meaning in Turns 3-7 was another example of a LRE which occurred during the use of AfL. Unlike the previous examples, this time the focus was on the meaning of the lexical items 'pm' and 'am'.

The following extract (4.8) is another example of one learner acting as the more capable peer during a conversation. The two learners were completing a task, which was to work with each other as Learning Partners (LP) to decide what colours to use in a Christmas themed picture. The children had been told that they could use a particular colour if it had been agreed jointly with their LP. The aim of the conversation in Extract 4.8 was to decide which colour to use for the Christmas tree.

EXTRACT 4.8

(7-9 year olds, using Learning Partners)

[1] **L7:** *can we put on it a black?*

[2] **L8:** *it's green for Christmas tree* (1)

[3] **L7:** *hmm* (2)

[4] **L8:** *I think you know it green* **Both LL reach for green colouring pencils**

In Extract 4.8, L8 seems to be more able to name colours accurately and in Turn 2 provides a model and an explanation of which colour should be selected. This conversation contributed to students working towards the learning aim for the lesson, which was to practice using colours and numbers. Employing the AfL technique called Learning Partners, which seemed intertwined with the task design, seemed to encourage conversation. Importantly, this interaction also offered an opportunity to clarify the meaning of the words 'black' and 'green' for L7, and hence could be considered an example of a lexical LRE. However, it is acknowledged here that there could be a different interpretation of L7's intention; in suggesting the use of black (Turn 1), L7 might have known the two colour names but wanted to convince her partner to use a somewhat less standard colour for the tree. But as both students coloured their trees green and L7 did not attempt to convince his/her partner, the latter interpretation seems less likely.

As is evident from the extracts discussed in this section, the materials mode prompted L-L type conversations, during which the teachers listened and often did not participate. Two points seem worth highlighting with regard to these conversations. First, both examples seemed rather fluent, with very few or no pauses. This was a notable observation in a TEYL classroom as it demonstrated that younger learners at early stages of language development could communicate with relative ease in a FL, given the right level of scaffolding. Secondly, that observation shows that young children are able to act as more able peers for one another. The analysis of both extracts (4.7 and 4.8) also demonstrates that using AfL techniques could contribute to scaffolding and facilitating L-L interactions.

4.4.3.2.4 The classroom context mode of interaction during AfL

The fourth mode in the Walsh' (2006) framework is the *classroom context mode* which allows learners to discuss their interests in a less constrained manner than otherwise. A small number of conversations in classroom context mode were recorded. But no such interactions while AfL was being used could be identified in the video recordings of the lessons. Hence, they were not included in this analysis. This could indicate that AfL did not facilitate conversations in the context mode in TEYL lessons. It is worth noting that this is the least constricted type of conversation and the observations made in this section could indicate that, while AfL was being used, only more tightly structured conversations occurred. This is a useful observation as it indicates that when AfL is used, the conversations that occur focus on the learners' understanding of instructions for tasks (managerial mode), completing tasks (materials mode) or clarifying the language rules and meaning (skills and systems mode).

The next section reports on extracts of classroom discourse that could not be classified within a single mode.

4.4.3.2.5 Mode side sequences during the use of AfL

Walsh (2006) identified a number of mode side sequence patterns (see Section 3.3.3.2.2 B). The analysis of the extracts from the classroom discourse indicates that, in the TEYL classroom, some conversations could not be classified as belonging to a single mode but they did display side sequences in a number of conversations.

The following extract (4.9) was recorded when the students were using WALT to assess each other. The task was to take turns in asking grammatically correct questions (using the pattern Has/Have…yet?) about pictures provided by the teacher. The two learners were completing the task, while the teacher was listening to their conversation.

EXTACT 4.9

(10-11 year olds, using WALT)

[1]   **L143:** *has* (1) *he* (1) *tidied the kitchen yet?*

[2]   **L145:** *yes he has* **L2 draws a tick on the corresponding picture** (2) *have he*

[3]   *taken the rubbish out yet?* (1)

[4]   **L143:** *has he* (1)

[5]   **L145:** *ok* (1) *he* (2) *has* (2) **L2 draws a tick on the corresponding picture**

[6]   **L143:** *no* (1) *but has he taken the rubbish yet* (1)

[7]   **L145:** *no* (1) *this now* (1) **L2 points to a different picture**

[8]   **L143:** *but this one is has he taken the rubbish yet*

[9]   **L145:** *yes he has* (1) *now you*

[10]   **L143:** *ok has he tidied his room yet?*

Initially in Extract 4.9, the conversation was centred on the task (Turns 1-3). However, in Turns 4, 6 and 8, L143 deviated from the materials mode and attempted to correct L145's grammatical mistake implicitly by recasting the incorrect phrase. This provides another example of a grammar LRE. No modification of L145's output occurred, presumably because L145 did not engage in the secondary mode, but seemed to remain in the materials mode throughout the conversation. This observation suggests that conversation modes may play a role in facilitating the modification of the output. The aim of that conversation from L143's perspective, seemed to be to correct the grammatical mistake made by L145. However, each attempt seems to be misunderstood by L145; hence the aim was not achieved. Evidently, L143 was the more capable peer in this conversation as (s)he knew the correct grammatical form. But L145 aimed to complete the task, i.e. to move on to the next picture, perhaps because L145 did not notice the implicit correction that was offered by L143. This implies that L145 was operating in the materials mode throughout the whole conversation. The fact that L145 did not seem to appreciate L143's message could explain the break in communication. In Turn 10, L143 seemed to return to the materials mode by moving on to the following picture. Hence, this conversation was classified as a mode side sequence pattern: materials-skills and systems-materials;

the materials mode is the main mode and the skills and system mode is secondary. The conversation in Extract 4.9 offered an opportunity to support the pedagogical objective of the lesson. However, due to the break in communication L145 did not seem to effectively use this learning opportunity.

Extract 4.10 below was sourced from a lesson in which T3 used Smiley Faces (SF) and the 'What Are we Learning Today' (WALT) type question with a class of 7-9 year olds.

EXTRACT 4.10

(7-9 year olds, using Smiley Faces and What Are we Learning Today – type question)

[1]   **T3:** *This is Stas*[19] (1) *can he do it* (1) *can you talk about beach activities using*

[2]   *like/love and ing words* **T points to the board which has WALT written on it** (2)

[3]   *the same statement* (1) **T points to a speech bubble in the activity**

[4]   **L15 (reads from the speech bubble):** *I like to sleeping*

[5]   **T3:** *What do you think* (1) *is that a happy face* (1) *medium face* (1) *or a frown* (1)

[6]   **L15:** *medium=*

[7]   **T3:** *= medium=*

[8]   **L15:** *=medium*

[9]   **T3:** *yeah* (1) *this is not so good* (1) *to sleeping* (1) *do people sleep on the beach*

[10]   **L15: shakes head**

[11]   **T3:** *sometimes* (1) *but what's better* (1)

[12]   **L15:** *sleeping*

[13]   **T3:** *what can he do to get a happy face* (3) *can he cross something here* (1)

[14]   **L15:** *I like sleeping*

[15]   **T3:** *yes* (1) *this is better*

In Extract 4.10, Turns 1-2, T3 reminded the learner what the WALT question was. Once L15 had read out the sentence in the speech bubble in the picture, T3 asked if the imaginary character in the picture could answer 'yes' (i.e. draw a happy Smiley Face) to the WALT question (Turn 5). T3 and L15 agreed that a straight face was the most appropriate. This choice meant that the sentence on the board was considered to be almost correct but a small change was needed. In Turns 9-14, T3 elicited the correct sentence structure from L15. This conversation started in the main skills and systems mode (Turns 1-8), then temporally deviated to classroom context mode (Turns 9-12) and returned to

---

[19] The name of a fictional character in the picture to which T3 was referring.

skills and systems mode (Turns 13-15) in which T3 and L15 continued discussing how to complete the task. Hence the conversation exemplifies a skills and systems – classroom context – skills and systems pattern. It could also be noted that the learners did not verbally partake in the context mode, i.e. they contributed no comments about their experience of seeing people sleep on the beach. The language used by the teacher in this conversation did not contain any phrases that were aligned with the target phrase 'I like sleeping'. Hence, T3's use of language was considered not to be matching the aim of the conversation, which was to correct the grammatical form. The pedagogical aim for the lesson was to practice the 'like/love plus progressive participle' sentence structure (Turns 1-2). This conversation clearly offered an opportunity for the teacher and the learners to work towards meeting that aim. By employing the smiley faces technique, the teacher engineered conditions for focusing the learners' attention on the form of the language, hence providing another example of how using AfL can facilitate an occurrence of grammar LREs.

In the following extract (4.11), learners were using Two Stars and a Wish (TSAW) AfL technique while completing a jumbled sentences task. They had been asked to put words in order to make grammatically correct sentences and use a TSAW template to mark two things they could do well in the process (two stars) and one that they needed to improve in the future (a wish).

EXTRACT 4.11
(10-11 year olds, using Two Stars and a Wish)
[1] **L123:** *what time*
[2] **L125:** *when*
[3] **L123:** *did you*
[4] **L125:** *what time when did you*
[5] **L123:** *nie dobrze* [Eng. *not good*]
[6] **L125:** *nie no co ty?* może być [Eng. *no what are you saying? it's ok*]
[7] **L123:** *kiedy ty rano wstałeś?* (1) [Eng. *when did you get up in the morning?*] *when did*
[8] *you get up in the morning*
[9] **L125:** *chyba o której rano wstałeś?* [Eng. m*aybe what time did you get up in the*
[10] *morning?*]
[11] **T3:** *English please*

[12]   L123: *what time* (1) *what time did you get up in the morning=*

[13]   L125: *=what time* (1) *did you co?* [Pol. co = Eng. *what*] **L4 writes on TSAW template.**

[14]   L123: *get up in the morning*

[15]   L125: *ok next one*

Extract 4.11 was almost an entirely dyadic interaction, with the teacher joining in once as the third interlocutor to request that the learners speak English (Turn 11). This conversation seemed to begin in materials mode as the students were deciding on the order of words in Turns 1-4. This was followed by a discussion about whether to begin the question with 'what time' or 'when' (Turns 5-14) which could be classified as the secondary skills and systems mode. In Turn 15, L125 places the conversation back in the materials mode by suggesting moving on to the next set of words. Although the students used a significant amount of Polish when discussing suggestions, notably their contributions seemed to aim to collaboratively arrive at a decision about which phrase the question should start with. In Turn 5, L123 provided negative feedback to L125. In Turns 7-14, L143 continued to provide the correct form, but this did not seem to lead to a modification of output by L125. It resulted in L125 making a note on the TSAW template. Although the video recording did not allow for seeing what L125 wrote and in which category it was written (star or wish), the fact that a note was made indicated that L125 reflected on the conversation or recorded a part of that reflection.

The aim of the lesson was 'to ask grammatically correct questions about the past'. Through the conversation the learners effectively worked towards meeting that pedagogical aim. More importantly, perhaps, the conversation exemplified how two learners negotiated the correct grammatical form to be used in the question. Extract 4.11 provided an example of how a grammar LRE occurred during the use of the Two Stars and a Wish technique. However, it is not evident that any modification of output occurred.

4.4.3.2.6 Summary of findings from employing the Variable Approach in analysing FL interactions

The results of the variable analysis of classroom discourse were synthesised and are presented below in Table 4.15. Column One reports the number of each extract; Column Two, the perceived pedagogical aim of conversation; Column Three, the learning aim for the lesson recorded in ROWDs; Column Four indicates if the language used by the teacher was congruent with the pedagogical aim of the conversation; Column Five, if the

conversations offered opportunities to contribute to the pedagogical aims for the lessons; Column Six, whether the opportunities that occurred were effectively used to facilitate learners to achieve the pedagogical aims; and Column Seven indicates the nature of the LREs that were evident from the interactions. The cases of L-L type conversations, where one of the learners acted as the more capable peer and the alignment of the language with the perceived aims of the conversations were recorded as not applicable (N/A). This choice is consistent with Walsh (2006) who researched the congruence of teacher, not learner, language with the pedagogical aims of conversations.

Table 4.15: Outcomes of applying the Variable Approach to analysing classroom interactions Variable analysis of classroom conversations which occurred during the use of AfL

1. Extract (age group)	2. Perceived aim(s) of the conversation	3. Learning aims for the lesson	4. Align-ment	5. Opportunity to support the lesson aims	6. Effective use of opportunity	7. Language Related Episodes
Managerial mode						
4.3 T-1L	To draw the learner's attention to a success in the criterion for writing.	To write an interesting newspaper article about Notting Hill Carnival	Yes	Yes	Yes	No
4.4 T-1L	To support the learners in assigning themselves a Traffic Light	To use eight or more phrases about hobbies	No	Yes	No	No
Skills and systems mode						
4.5 T-1L	L: to find out how to spell 'throwing' T: to elicit a correct question structure and to support with spelling.	To use was/were + ing to talk about the past	No	Yes	Yes	Grammar
4.6T-1L	L: To check if 'too much fans' was correct; T: To correct a mistake: 'too much fans'	To use (too) much/many, not enough with countable and uncountable nouns	Yes	Yes	Yes	Grammar
Materials mode						
4.7 L-L	To negotiate the meaning of 'pm'	To make suggestions using Why don't we..?, Shall we...? and Let's	N/A	Yes	Yes	Lexical
4.8 L-L	To select a colour for a Christmas tree	To practice speaking through games	N/A	Yes	Yes	Lexical
Mode side sequences						
4.9 L-L	To correct a mistake of using 'have' with a singular 3rd person pronoun	To ask grammatically correct questions using Has/have...yet?	N/A	Yes	Yes	Grammar
4.10 T-1L	To correct a mistake in the following sentence: I like to sleeping	To practice 'like/love plus ing' structure	No	Yes	Yes	Grammar
4.11 L-L	To decide whether to start a question with 'what time' or 'when'	To ask 8 questions correctly, about activities that a person did yesterday.	N/A	Yes	Yes	Grammar

The data in the current study, summarised in Table 4.15, indicate that the teachers did not align their use of language with the aims of the conversations in three out of five extracts. This suggested that the use of AfL did not seem to be related to ensuring alignment of teacher language use with the perceived aims of the conversations. This finding is similar to Walsh (2006), who reported little congruence between teacher language use and the perceived aims of classroom T-1L interactions.

However, the data summarised in Table 4.15 indicate a number of interesting results which may offer an explanation of how AfL could facilitate learning, currently missing from the published research on AfL (e.g. Bennett, 2011). Firstly, it could be observed that all T-1L and L-L interactions offered opportunities for contributing to the lesson aim(s). As might be expected (or perhaps hoped for), almost all T-1L interactions (except in Extract 4.4) were effectively used to contribute to the pedagogical aim(s). Moreover, it was especially useful to note that all L-L interactions that created learning opportunities were used effectively by the learners. It was evident from the analysis that when AfL was used, the conversations focused on achieving the learning aims for the lesson. This was evidenced particularly by the absence of the context mode in the data. This is a useful finding as it indicates that AfL could facilitate conversations that support the pedagogical aims of the lesson effectively. It also corroborates with the findings about teacher beliefs that AfL helped to focus efforts in the classroom on ensuring achievement.

Secondly, the findings indicate that LREs occurred in extracts recorded during the use of AfL (Column 7). LREs were observed in all interactions, except for those in managerial mode (seven out of the nine extracts). This indicates that when AfL is used in managerial mode, the conditions may not be conducive to the occurrence of LREs. It seems especially interesting to note that there appears to be a relationship between conversation modes and the types of LREs that occurred. Grammar LREs were observed in skills and systems mode (4.5 and 4.6) and in mode side sequences[20], which included the skills and systems mode as the main (4.10) or the secondary (4.9 and 4.11) mode. This could indicate that when interlocutors operate in the skills and systems mode, the LREs are more likely to have a grammar focus. The majority of interactions that displayed grammar LREs were of the T-1L type. The occurrence of the skills and systems mode, even as the secondary mode in the mode side sequence, could be related to the types of LREs that interlocutors

---

[20] Side sequences included 4.9: materials – skills and systems – materials; 4.10: skills and systems – classroom context (T only) – skills and systems; 4.11: materials – skills and systems – materials.

engage in. The materials mode in L-L interactions was observed to co-occur with lexical LREs. This observation provides further support for the interpretation that conversation modes might be related to the types of LREs that occur during the use of AfL.

The analysis of Extracts 4.3-4.11 indicates that some conversations provided opportunities for providing interlocutors with negative feedback. These results are summarised in Table 4.16 below.

Table 4.16: Negative feedback and modifications of output

Ex-tract	Mode	Type	Negative feedback	Modification of output	Output but no modification	Lack of output
4.3	MAN	T-1L	No	N/A	N/A	N/A
4.4		T-1L	No	N/A	N/A	N/A
4.5	SS	T-1L	Yes	Yes	No	No
4.6		T-1L	Yes	Yes	No	No
4.7	MAT	L-L	Yes	Yes	No	No
4.8		L-L	Yes	No	No	Yes
4.9	MAT-SS-MAT	L-L	Yes	No	Yes	No
4.10	SS-CC-SS	T-1L	Yes	Yes	No	No
4.11	MAT-SS-MAT	L-L	Yes	No	Yes	No
Key to codes: T-1L: teacher-learner interaction; **L-L**: learner-learners interaction; **MAT**: materials mode; **SS**: skills and systems mode; **MAN**: management mode; **CC**: Classroom context mode.						

As indicated by the analysis presented in Table 4.16 above, all cases when LREs occurred were initiated by negative feedback from one of the interlocutors. In some cases it led to a modification of output. Modification occurred in both extracts that displayed the skills and systems mode (Extracts 4.5 and 4.6), in one in materials mode (Extract 4.7), and in the one mode side sequence in which the main mode was skills and systems (Extract 4.10). In one instance (Extract 4.8), the negative feedback was not followed by any verbal output but led to a change in the learner's action thus to a modification of behaviour. In the remaining two extracts (4.9 and 4.11), negative feedback did not result in any modification of the output that followed it. Both of those extracts (4.9 and 4.11) were in mode side sequences, where the main mode was MAT. This finding indicates that feedback seemed to be more often used to modify output in the skills and systems mode: a finding that is similar to Oliver and Mackey (2003, see discussion in 5.5.2). This suggests that conversation modes should be taken into account when discussing negative

feedback in TEYL classrooms. However, it should be noted that the current study did not aim to analyse the amount of negative feedback in interactions that occurred during AfL. Instead, it set out to describe the largely under-researched area of the use of AfL in TEYL. Hence, the findings can only point to a useful path for future enquiry, but do not provide a quantitative insight into the amount of negative feedback and modification of output.

4.4.4 Summary of findings for Research Question Three

Research Question Three aimed to investigate the observable impact of AfL on classroom interactions. The findings are summarised below.

1. There was a strong positive correlation between the frequency of using AfL techniques and the number of L-L interactions: statistically significant at the 0.01 level.

2. There was a strong positive correlation between the frequency of using AfL techniques and the number of T-1L interactions: statistically significant at the 0.01 level.

3. A moderate negative correlation was found between the frequency of using AfL techniques and the number of T-C interactions with learners aged 7-9: statistically significant at the 0.05 level.

4. The majority of T-1L interactions and L-L interactions displayed holistic patterns that had been shown in previous research to facilitate L2 learning (Swain, 2000, Butler & Zeng, 2014).

5. The students assumed active roles during T-1L and L-L interactions, as evidenced by medium to high levels of mutuality.

6. There was no evidence that using AfL techniques supports aligning teacher's language with the pedagogical aims of conversations.

7. T-1L and L-L conversations during the use of AfL offered opportunities for supporting the learning aims of the lesson.

8. Most of those opportunities were effectively used.

9. LREs occurred during the use of AfL in materials and skills and systems conversation modes as well as in mode side sequences that included those two modes.

10. Grammar LREs occurred in skills and system mode and in mode side sequences that included the skills and systems mode as the main or the secondary mode.

11. Lexical LREs occurred in materials mode.

12. Negative feedback on output was provided in skills and systems, materials and mode side sequences.

13. Feedback was more often used in skills and systems mode than in any other mode.

4.5 Concluding remarks for Chapter Four

This chapter has reported the findings for the three research questions. The summary of findings for each RQ has been presented in the respective parts of the chapter. These findings are synthesised in the following chapter and discussed in the context of the literature review presented in Chapter 2.

Chapter Five: Discussion

5.1 Introduction to Chapter Five

The aim of the current chapter is to interpret the findings reported in Chapter 4, discuss their implications and consider how these findings relate to the current knowledge base presented in the literature reviewed in Chapter 2. The discussion also evaluates the extent to which the findings address each research question.

This chapter is organised in five sections. First, Section 5.2 provides a synthesis of the findings for each research question. Subsequently, in Section 5.3, teachers' understanding of AfL is related to the theoretical frameworks of AfL. In Section 5.4, the discussion considers how teachers implement AfL in TEYL classrooms by discussing them in the context of theories of cognitive development and FLL in childhood. Following that, Section 5.5 discusses the role of interaction in second language learning. The discussion in each section is extended by explicitly identifying the contribution of the findings in the current study to research on FLL in TEYL contexts. Sections 5.3-5.5 also discuss the pedagogical and research implications of the findings of the current study.

5.2 Synthesis of the findings

The aim of the present study has been to examine how AfL is understood and used by YL teachers and what observable impact it has on interactions in the TEYL classroom. It is important in that it responds to calls for research into AfL (e.g. Bennett, 2011; Rea-Dickins, 2001), assessment of YLLs (e.g. McKay, 2006) and intends to extend understanding of how assessment can facilitate learning in TEYL contexts by addressing the issue of the lack of empirical evidence for how AfL could facilitate learning (Dunn & Mulvenon, 2009). The current section synthesises the findings relating to each research question in turn.

5.2.1 The Findings for RQ1

Research Question One was: How do teachers understand AfL after receiving a limited amount of training and using AfL techniques for at least one academic year when teaching English to young learners aged 7-11?

The findings reported in Chapter 4 indicate that YL teachers understand AfL as a type of classroom-based assessment that is integrated into the teaching and learning processes.

Most teachers reported that such integration could be enacted in TEYL classrooms through: raising learners' awareness of the learning objectives, focusing their efforts on achieving those objectives, building in continuous assessment opportunities, providing feedback that aims to move learning forward and using the teaching task as a context for assessment opportunities. The findings also propose that some teachers considered these ways of incorporating assessment to be motivating and enjoyable for learners.

Additionally, the findings suggest that the majority of the teachers considered AfL techniques to be compatible with the teaching methodology that they were using in TEYL classes. The compatibility was interpreted by the teachers in two ways. The first was that practices similar to AfL techniques had already been used by some of the teachers in the study. However, the terminology they might use to describe them differed from AfL terminology. The second explanation of the compatibility of AfL with the teaching methodology indicated that AfL was easy to incorporate as activities that were typically used with YLLs. However, a lack of compatibility between AfL and the school's policy to report summative grades to parents also emerged.

Furthermore, the findings addressing RQ1 provided insights into the teachers' beliefs about how they implemented AfL. They indicate that AfL was not implemented in an ad hoc manner; rather, it was planned by the teachers who used it and time was set aside in the lesson for implementing AfL techniques. The majority of teachers reported that they employed various AfL techniques in TEYL classrooms and did so at three different times during some lessons: while introducing the learning objectives for the lesson and the instructions for each task and/or while the learners were in the process of completing tasks and/or in the final stages of a task and/or a lesson. The findings also suggest that most teachers considered diversity of technique type to be important in TEYL classes as it might encourage learners' engagement. The findings also indicate that that those teachers who implemented AfL were selecting different technique types to fit different purposes. Most teachers reported that they employed AfL mainly for the following purposes: to share learning aims; to give instructions for the tasks and the criteria for success; to provide feedback from the teacher and/or peers; to facilitate self-reflection and to measure learners' confidence about their own achievement.

The teachers who used AfL also reported what impact they observed it to have in the lessons with YLLs aged 7-11. Most significantly, they reported that the use of AfL creates

conditions that enable young learners to sustain independent and pair work and enable teachers to provide more individually tailored support for learners. The findings suggest that the purpose of such ongoing, classroom-based monitoring and support is to move learning forward. Additionally, the teachers' reports indicated that the quality of interactions between the learners changed when AfL was being used and as a result YLLs were more capable and willing to collaborate as opposed to compete with peers.

This chapter discusses the YL teachers' understanding of AfL within the theoretical frameworks proposed by Black and William (2009) and Swaffield (2011). In doing so, the study contributes new insights into AfL, which are specific to TEYL contexts and reported from teachers' perspective.

5.2.2 The Findings for RQ2

Research Question Two consisted of the following two subquestions: 2.1 How do teachers' translate their understanding of AfL into classroom practice in a TEYL context with students aged 7-11 in a private language school in Poland?; and 2.2 Do teachers report any changes in their practice of using AfL over time?

The findings for RQ2.1 provided an inventory of various AfL techniques and insights into their implementation. First, the findings indicated between-teacher variance in frequency and diversity of technique type. This resulted in the identification of four possible ways of implementation: *frequent/diverse*, *infrequent/diverse, infrequent/not-diverse* and *frequent/not-diverse*. However, the findings from deploying more detailed frequency measures, inter– and intra-lesson frequency suggest that diversity in implementing technique type occurred together with high intra-lesson frequency while low intra-lesson frequency was observed to occur together with relatively low diversity. This suggests that the teachers did not tend to use the same technique types within one lesson. Hence, the four types of implementation identified through the empirical data were:

1. *intra- and inter-lesson frequent/diverse*
2. *intra-lesson frequent but inter-lesson infrequent/diverse*
3. *intra- and inter-lesson infrequent/not-diverse*
4. *intra-lesson infrequent but inter-lesson frequent/not-diverse*

All four types of implementation were evident from the data of both age groups (7-9 and 10-11). Findings to RQ2.2 indicated that over time, the majority of the teachers

implemented AfL with a growing level of diversity in technique type. This finding suggests that developing diversity in technique type is a common aspect of implementing AfL in TEYL classes. That might point to the importance of employing a diverse number of techniques in such a context. This interpretation corroborates the findings for RQ1, which suggested that teachers believed it was important to use various technique types to ensure the learners' engagement. The frequency and diversity in the use of AfL changed over time at the inter-teacher level but remained largely unchanged across the whole sample. The longitudinal findings suggest that familiarity with technique types, the perceived ease of implementation, as well as the experience of using or observing AfL in lessons might affect the range of technique types used by individual teachers.

Another important empirical finding to RQ 2.1 suggests that fitness for purpose might be an important consideration in implementing different AfL technique types. Overall, three main purposes for implementing were identified:

> **Purpose 1:** To share learning objectives and expectations
>
> **Purpose 2:** To monitor performance
>
> **Purpose 3:** To check achievement

For each purpose, there was some age-related variance in implementation. With reference to Purpose 1, the implementation of AfL for giving and clarifying instructions was observed only in the younger age group. Within Purpose 2, the implementation of AfL for measuring learners' confidence was more evident in the lessons with the 7-9 year olds than in the lessons with the older group. Within Purposes 2 and 3, more reliance on peer and teacher feedback was observed in the younger age group, while with the older age group, the teachers employed more techniques that encouraged self-assessment and self-monitoring of own progress. This suggests that there might be a relationship between the implementation of AfL and the learners' growing levels of metacognitive control. In the case of all three purposes, a greater reliance on literacy skills was observed in the older age group. Additionally, the findings indicate that, depending on the purposes it served, AfL was utilised in various parts of the lessons and the tasks.

The findings to RQ2.1 also revealed that AfL techniques were implemented predominantly when productive skills were being taught. This was also confirmed in the longitudinal phase (RQ2.2). This suggests that AfL techniques might be useful in providing additional scaffolding in speaking and writing tasks, especially if these are

open-ended. This interpretation points to a possible relationship between the implementation of AfL and the design of the tasks within which it is implemented. Furthermore, the findings from analysing Records of Work Done confirmed the teachers' reports that the use of AfL was not easily compatible with the reporting of the summative assessment results that teachers were required to carry out by the school.

It is believed that the above insights into the use of AfL demonstrate that the current study has explored the what, when, how, who and why of the implementation of AfL in TEYL classrooms, which is not currently available in the literature.

5.2.3 The Findings for RQ3

Research Question Three was: What is the observable impact of AfL on classroom interactions in a TEYL context?

Useful new insights emerge from investigating classroom interactions during the use of AfL. Overall, the findings indicate that when AfL was used, there were more opportunities for the learners and teachers to engage in one-to-one interactions during the lessons. Significant positive correlations were found between the use of AfL and the number of T-1L and L-L interactions. Also, the moderate negative correlations between T-C and use of AfL indicated that the lessons with 7-9 year old classes in which AfL was used were less teacher fronted than those in which AfL was not implemented.

The analysis of extracts of transcribed interactions during the use of AfL provided empirical evidence pertaining to how AfL could facilitate learning in TEYL classes. First, the holistic interaction patterns identified during the use of AfL were largely of the types that had been found to support learning, i.e. the expert/novice and the collaborative pattern (e.g. Butler & Zeng, 2014; Swain, 2000). Furthermore, the Variable Approach to analysing classroom interactions indicated that the one-to-one conversations occurring during the use of AfL were either of the following modes: materials, managerial, skills and systems, or displayed mode side sequences (Walsh, 2006). There were no extracts demonstrating that YLLs engaged in the classroom context mode when using AfL. That finding provides useful new insights into the nature of interactions between teachers and learners and among learners. The analysis also points to little alignment between the teachers' use of language and the pedagogical objectives of the conversations, hence indicating that teachers did not use language in a manner that may facilitate FLL (Walsh,

2006). This finding is similar to what Walsh (2006) found. Interestingly, however, the majority of the interactions that occurred during the use of AfL offered opportunities for the teachers and students to move learning forward in relation to the pedagogical objectives of the lessons and almost all of those opportunities were efficiently used. This finding corroborates the reports of most of the teachers that using AfL helps to focus learners' efforts on achieving the learning objectives set for each specific lesson.

Moreover, the findings suggest that the potential pedagogical merits of the interactions that occur during the use of AfL might be related to the mode of conversation. During all the conversations in the materials mode and in the skills and systems modes as well as in the mode side sequences accompanying those modes, LREs were observed. The skills and systems mode and the mode side sequences accompanying that mode seemed to facilitate the occurrence of grammar LREs while in the materials mode lexical LREs occurred. In the skills and systems mode and its accompanying mode side sequences, the LREs resulted in modification of output. These findings are discussed with reference to the body of evidence on the contribution of interactions to the development of FL (e.g. Oliver, 2002; Oliver & Mackey, 2003) and the criteria for effective assessment conversations proposed by Ruiz-Primo (2011).

5.3 Teachers' understanding of AfL in a TEYL classroom

Despite the attention that AfL has received from governments and researchers since the 1990's (Dann, 2002; Klenowski, 2009; Swaffield, 2011) it lacks a firm theoretical model (Davison & Leung, 2009). Bennett (2011) refers to this lack of an established theoretical framework and its associated inconsistent use of terminology as a definitional issue. In the attempt to answer the first research question, this research drew on the understanding of AfL discussed by Wiliam and Black (2009) and Swaffield (2011). Both of their frameworks focused on generic, i.e. non-domain specific, characteristics of AfL. This section examines teachers' understanding of AfL by comparing it to both these frameworks. This method was adopted in order to identify similarities, which may indicate generic characteristics of AfL, and differences, which offer the opportunity to gain insights into aspects of AfL that are specific to a TEYL context. In doing so, the discussion extends the current body of knowledge about AfL.

5.3.2 AfL terminology

Before the discussion turns to the theoretical frameworks, I wish to revisit the terminology adopted here. The key terms are AfL and FA. These are not understood as synonyms in the current thesis. The distinction between them is discussed fully in Section 2.3.2.3. Since a consistent use of the terminology of AfL is lacking in the field, the studies discussed in this chapter were selected because their authors adopted similar definitions of AfL (regardless of which of the two key terms they used). The theoretical framework adopted in the current study is consistent with the notion of formative for-learning assessment (Leung, 2004) and informal formative assessment (Ruiz-Primo, 2011) as discussed in Section 2.3.2.3.

The lack of a consistent use of terminology is not only problematic from the research perspective. Perhaps more importantly, it may inhibit understanding and effective implementation of AfL by policy makers and teachers (e.g. Harlen, 2005; Swaffield, 2011). In effect, the possible benefits for learning might not be capitalised on. Although some attempts at clarifying the theoretical framework and the terms associated with it have been made (e.g. Black & Wiliam, 2009; Klenowski, 2009; Ruiz-Primo, 2011), the definitional issue (Bennett, 2011) still persists. While the for-learning purpose of AfL seems to be generally agreed upon by researchers (e.g. Black & Wiliam, 2009; Ruiz-Primo, 2011; Swaffield, 2011) and is evident in the findings related to the teachers' understanding of AfL in the present study, the use of the term 'assessment' seems more problematic. Notably, Black and Wiliam (ibid.) use the term 'practice' instead of 'assessment' in their definition, as discussed in Section 2.3.2. However, this term is not helpful as it seems too generic. Swaffield (ibid.) also recognises the problematic use of the term 'assessment' and attempts to deal with it by defining the term based on its etymology.

> 'The word 'assessment' has its roots in the Latin verb assidere meaning 'to sit beside' (...). The picture of someone sitting beside a learner, perhaps in dialogue over a piece of work, represents much more accurately assessment as a support for learning rather than assessment as a test of performance.' (p. 434)

However, she also recognises that the current (mis)understanding of the term 'assessment', i.e. signifying predominantly summative practice, might have been caused by the dominance of examinations and testing in education. Although, somewhat more

encouraging than Black and Wiliam (2009), this attempt at dealing with the terminology issue does not seem to be entirely helpful because adopting Swaffield's understanding of the word 'assessment' would inevitably cause tensions in the terminology of summative assessment, which, in contrast to AfL, seems rather well established.

It seems logical to propose that a term should to some extent describe the practical implementation of the concept that it refers to. Findings of the current study suggest that in TEYL classrooms AfL is enacted by teachers, learners and peers who share learning objectives, determine where learners are in relation to those objectives, and use that information in order to move learning forward (5.3.3). These findings are consistent with the body of research discussed in Section 2.3.2. It seems useful to consider how the term AfL signifies such practical implementation. First, the most obvious relationship seems to be between the term 'for learning' and the use of AfL to move learning forward. As indicated above, this aspect is not problematic. Secondly, making judgements about where learners are in relation to a learning goal seems to entail a summative component of AfL. This interpretation is consistent with Tarras (2005), who argues that 'all assessment begins with summative assessment (which is a judgement) and that formative assessment is in fact summative assessment plus feedback which is used by the learner' (p. 466). Although this conceptualisation acknowledges that AfL has a summative component, it does not seem very helpful because it could be argued that it encompasses any summative assessment activity that is followed by feedback, e.g. a test with a numerical grade. Moreover, as Ruiz-Primo (2011) rightly notices and as is evident in the findings of the present study, AfL does not begin with making judgments. Instead, Ruiz-Primo (ibid.) proposes 'clarifying the learning expectations (goals) as a prerequisite for collecting, interpreting, and acting on the information with the intention of improving student learning.' (p. 18). Bennett (2011) offers a useful way of conceptualising the summative component of AfL. He argues that each assessment activity has a primary and a secondary purpose. While AfL's primary purpose is for learning, it also has an of learning purpose which enables accomplishing the primary purpose. Although Bennett (ibid.) does not define the term *primary*, his discussion suggests that it refers to the main reason why the assessment is being conducted. Hence, it could be argued that there is a summative component to AfL that occurs in the middle stages of implementing AfL and is secondary in nature.

Finally, the findings of the current study confirm that AfL happens through interactions between teachers, learners and peers. This corroborates with other studies about AfL (e.g. Black & Wiliam, 2009; Ruiz-Primo, 2011). The social aspect of AfL is not reflected in any of the terms used in the literature despite the fact that it seems to be the vehicle for conducting the primary, i.e. for learning, purpose of AfL. It should be noted that, for AfL to be effective, social interactions should be 'dialogic and interactive in nature' (Ruiz-Priomo, 2011, p. 18) and should display high levels of mutuality (Butler & Zeng, 2014). Hence, it seems reasonable to suggest that a term referring to the idea of a socially situated interaction (verbal and non-verbal) that aims to move learning forward would describe the concept of AfL as defined in Section 2.3.2 more accurately. The current discussion is not intended to further proliferate terminology but is considered necessary as it contributes to the development of a terminology that would have a well-defined theoretical framework based on its pedagogical implementation and would describe the concept that it refers to more accurately. As discussed in this section, the terminology of AfL and FA currently used in the field does not seem to satisfy that criterion. The above proposition is examined in the context of the findings of the current study and suggestions are made about how it can contribute to moving our understanding of AfL forward[21].

As no better alternative is currently available, the term AfL is used throughout the discussion to refer to the concept defined in Section 2.3.2.

5.3.3 Teachers' understanding of AfL compared to the Black and Wiliam (2009) framework

The discussion now focuses on comparing the findings for RQ1 with the Black and Wiliam (2009) framework. Because their theoretical framework is not domain-specific, it is especially valuable to compare it to the teachers' understanding of AfL as the similarities and differences may reveal domain characteristics specific to a TEYL context. Such a context has at least two specific features: curriculum area (FL English) and learners (young children). It is believed that by evaluating whether the teachers' understanding of AfL is consistent with that of Black and Wiliam (ibid.), the discussion

---

[21] Following completion of writing up the current chapter in February 2015, I emailed the teachers who participated in the study with the following question: 'If AfL were not called AfL, what would you call it for the name to show what it is?' The answers I received included: 'negotiated learning', 'empowered learning', 'learning to learn', 'learning and teaching techniques that support learning', 'learning development techniques'. Notably, none of the five teachers who responded used the word 'assessment'.

does not only contextualise the findings for RQ1 but also provides interesting insights into the teaching and learning processes that occur in TEYL classrooms.

Black and Wiliam (2009) propose a theoretical framework that incorporates five aspects of AfL. The framework incorporates Ramaprasad's (1983) three stages of teaching. Within each stage, Black and Wiliam (ibid.) consider the roles and responsibilities of the participants in assessment procedures (teachers, peers and learners). The authors argue that the overall purpose of implementing AfL, which is to move learning forward, is also shared with cognitive acceleration programmes and dynamic assessment. However, the authors point to the prescriptive nature of these programmes that makes them different from AfL. Importantly, they comment that the formative function of assessment, which is central to AfL practices, is also inherent in cognitive acceleration programmes and dynamic assessment. However, discussing that area is beyond the scope of the current chapter. Instead, the five aspects of AfL cited by Black and William (ibid.) are explored in more detail.

Aspect One has teacher-led and learner-led components. In enacting the first aspect of AfL, the teachers clarify the learning objectives and criteria for success while their students can enact this aspect by understanding them. These students can also share their understanding with their peers. Hence, it seems that Aspect 1 encompasses different actions that are intended to ensure that all participants of the teaching and learning process understand what is being taught. This aspect seems to be an integral part of the teaching and learning that occurs in classes. That is to say, learning objectives can be shared and, hopefully, understood within the context of a task, a lesson or perhaps a longer unit of work. Hence, the implementation of this aspect might be sensitive to the educational context in which it is enacted. This is an important feature as it indicates that there is value in considering the practical implementation of the framework in different subject domains, a view shared by Black and Wiliam (ibid.), who point out that 'what counts as a good explanation in the mathematics classroom would be different from what counts as a good explanation in the history classroom, although they would also share certain commonalities' (p.27). Here it is the teachers' understanding of AfL specifically in the domain of teaching English to Young Learners that is explored. Thus, the discussion illustrates teachers' understanding specific to TEYL classrooms and allows for comparisons with the generic aspects proposed by Black and Wiliam (ibid.).

Aspects 2 and 3 are related to teacher-led actions. In order to fulfil the overall aim of moving learning forward, teachers implement various strategies that provide evidence of where learners are in reference to the learning intentions and criteria for success clarified through Aspect 1. Black and Wiliam (2009) cite questioning as an example of a strategy that may be employed within Aspect 2. Aspect 3 seems to naturally follow Aspect 2. It includes strategies that teachers can implement to guide learning and that are informed by the information collected through Aspect 2. Black and Wiliam (ibid.) propose comment-only marking as a strategy for implementing Aspect 3.

Aspects 4 and 5 are learner-led. Aspect 4 considers peers as important agents in the assessment process while Aspect 5 points to learners' agency in assessing their own learning. This agency can be enacted, for example, by conducting peer- or self-assessment. However, Black and Wiliam (ibid.) do not explore that area in detail. Therefore, how learners could establish where they or their peers are in relation to the learning objectives and how that might differ from guiding themselves or their peers towards achieving the objectives remains open to interpretation. Despite that, Aspects 4 and 5 do provide the useful proposition that learners can have a role in monitoring and guiding their own and/or their peers' learning. A summary of the framework discussed by Black and Wiliam (2009) is provided in Table 5.1 below.

Table 5.1: Aspects of AfL, adapted from Black and Wiliam (2009, p. 8)

	Column A: Where is the learner going?	Column B: Where is the learner right now?	Column C: How does the learner get there?
Row 1: Teacher	**Aspect 1.** Clarifying learning intentions and criteria for success	**Aspect 2.** Engineering effective classroom discussions and other learning tasks that elicit evidence of students' learning	**Aspect 3.** Providing feedback that moves learners forward
Row 2: Peer	Understanding and sharing learning intentions and criteria for success	**Aspect 4.** Activating students as instructional resources for one another	
Row 3: Learner	Understanding learning intentions and criteria for success	**Aspect 5.** Activating students as the owners of their own learning.	

It seems especially valuable to examine whether the findings of the current study confirm that the aspects of AfL proposed by Black and Wiliam (ibid.) can be identified in a TEYL context and whether any differences exist. Such a comparison may offer insights into the stages of learning and the roles of the participants. The discussion is organised in two sections. First, it focuses on the aspects of AfL that are teacher-led. Secondly, learner-led aspects are discussed. To facilitate the discussion, the findings addressing RQ1 are summarised in Appendix 24.

The findings about the teachers' understanding of AfL in a TEYL context have been mapped out against the five aspects of AfL discussed above. The outcome is reported in Table 5.2 below. The numbers referring to each theme correspond with the numbers reported in Chapter 4 and summarised in Appendix 24.

Table 5.2: Teachers' understanding of AfL mapped out against the Black and Wiliam (2009) framework

	Column A: Where is the learner going?	**Column B:** Where is the learner right now?	**Column C:** How does the learner get there?
Row 1: Teacher	**Aspect 1:** **Theme 4 (part 1):** AfL techniques could be used to serve the purposes of giving and clarifying instructions, sharing aims and criteria for success **Theme 1f:** Teaching and learning can be integrated with assessment by providing structure and focus to lesson planning	**Aspect 2:** **Theme 1c:** Teaching and learning could be integrated with assessment by teachers continuously building AfL opportunities into lessons. **Theme 6a:** When the students were able to work independently, the teacher could spend time on monitoring work more effectively and providing individual support.	**Aspect 3:** **Theme 4 (part 2):** AfL techniques could be used to serve the purposes of feedback and measuring learners' confidence about their learning. **Theme 1d:** Teaching and learning could be integrated with assessment within an AfL framework giving meaningful explicit feedback that helps to move learning forward (from teacher, peers and own reflection).
Row 2: Peer	_	**Aspect 4:** **Theme 1d:** see above **Theme 6b:** Using AfL enabled the teachers to introduce more pair work in the lessons. **Theme 6c:** The students collaborated rather than competed when working together.	
Row 3: Learner	**Theme 5b:** Using AfL helped the learners become more responsible for their own learning.	**Aspect 5:** **Theme 1:** Teaching and learning can be integrated with assessment by **a:** making learners more aware of what they are learning, **b:** focusing learners on achieving the learning objectives, **d:** giving meaningful explicit feedback that helps to move learning forward (from teacher, peers and own reflection). **Theme 5b:** Using AfL helped the learners become more responsible for their own learning.	

Table 5.2 suggests that there are many similarities between the Black and Wiliam (2009) framework and the understanding of AfL reported by the teachers in the current study. These, with references to the TEYL context, are explored in the following sections. First, the teacher-led aspects are discussed in Section 5.3.3.1. This is followed by the learner-led aspects of AfL in Section 5.3.3.2. Throughout, references are made to the aspects and themes presented in Table 5.2.

5.3.3.1 Teacher-led aspects of AfL

As is evident from Table 5.2, all the aspects that are teacher-led (See Row 1 in Tables 5.1 and 5.2) were reported in the current study. This suggests that the role of the teacher in deploying AfL in TEYL classes is similar to the role that teachers assume in other contexts. This interpretation seems plausible as there was significant evidence in the data which indicated the teachers reported 'clarifying goals and criteria for success' (Aspect 1 of the Black and William (ibid.) model) by using AfL techniques for the purposes of giving and clarifying instructions, sharing aims, criteria for success (Theme 4 in Table 5.2, Column A, Row 1). Some additional evidence to support the above interpretation is provided by the teachers reporting that AfL provided 'structure and focus to teachers' lesson planning' (Theme 1f).

The second aspect of the role of the teachers in enacting AfL as discussed by Black and Wiliam, (ibid.) indicates that teachers would engineer 'effective classroom discussions and other learning tasks that elicit evidence of students' learning'. This is also strongly reflected in the understanding of AfL shown by the teachers in the study. They reported that they 'continuously build AfL opportunities into lessons' (Theme 1c in Table 5.2). Presumably these enabled them to elicit evidence of the students' learning. Additionally, the teachers indicated that when they used AfL, the dynamics of the classroom changed so that 'when students are able to work without a teacher's help, the teacher could spend time on monitoring work more effectively and providing individual support' (Theme 6a in Table 5.2).

Finally, the third aspect identified by Black and Wiliam (ibid.) was also reflected in the themes obtained from the teacher interviews and the focus group. Aspect 3 indicated that a teacher's role in AfL is to 'provide feedback that moves learners forward'. The data in the current study suggests that most of the YL teachers used AfL techniques for 'giving feedback and evaluating learners' confidence about their learning' (Theme 4 in Table 5.2,

Column 3, Row 1). Importantly, the teachers reported that they understood AfL as a way of 'providing a framework for giving meaningful feedback that helps to move learning forward (from teacher, peers and own reflection)' (Theme 1d in Table 5.2). Evidently, this understanding also indicates that the data provides evidence for Aspects 4 and 5 of the Black and Wiliam (ibid.) framework. The discussion of these learner centred aspects of AfL as proposed by Black and Wiliam (ibid.) is provided in Section 5.3.3.2.

The similarities between the teacher-led aspects of AfL in the findings of the current study and the Black and Wiliam (ibid.) framework, discussed above, provide interesting insights when considering other research. The theme corresponding to Aspect 1 in Table 5.1 (i.e. 'Clarifying learning intentions and criteria for success') was very frequently expressed in the interview data. This might suggest that it is an important feature of AfL in a TEYL context. It seems plausible to infer that the use of AfL helped the teachers achieve Aspect 1 when setting up lessons and tasks by providing a familiar structure to those stages of the lessons. This interpretation is reflected in the data. For example, T7 commented that 'within the lesson it [AfL] provides a REALLY good structure of how things are organised and I think the students have benefitted from that'. This is an interesting insight into the use of AfL, especially when compared with the findings of the research that suggests that FLL in childhood may benefit from using familiar task types as this may facilitate collaboration in completing the task (Pinter, 2007). Research about adult contexts suggests that tasks with more structure can facilitate different aspects of linguistic output, including complexity (Bygate, 1996) or fluency (Skehan & Foster, 1999). As the current study suggests that AfL might contribute to providing a structure for lesson activities, it is plausible to infer that AfL could help learners collaborate to complete a task and perhaps produce better quality output (see Section 5.5.2 for further discussion of output).

The above interpretation could have important pedagogical implications as it suggests that, by systematically employing AfL in TEYL classrooms to clarify the learning objectives and criteria for success, teachers may be able to use the time in class more effectively. Hence, the systematic use of AfL techniques could have pedagogical merits. However, it should be noted that the present study did not evaluate the design of tasks or lessons that were observed. Hence, there remains the possibility that a sufficient level of scaffolding that supports learners in producing a higher level of complexity or fluency

could be obtained without AfL. Further research could provide more detailed insights into the possible positive impact of using AfL for clarifying learning objectives and criteria for success on learning in TEYL classes and on the quality of output produced by learners (see also 5.4.2.1 for the discussion on timing in the use of AfL).

Teacher-led Aspect 2 (Table 5.2, Column B, Row 1) was reflected in Themes 1c and 6a in the findings of the current study. These highlighted two ways in which AfL can be operationalised in TEYL contexts. First, Theme 1c indicates that planning for the use of AfL is important. This implies that implementing AfL requires a teacher's conscious efforts and does not rely solely on reacting to classroom circumstances. Presumably, the planned use of AfL provides opportunities for gathering information on how learners progress with reference to learning objectives. Such information is a pre-requisite to enacting Aspect 2. Considered from the socio-cultural perspective, which maintains that teachers can gain insights into the potential development of their learners through enabling interaction with more capable peers (Ohta, 1995; Swain, 2000), it seems that, for the effective practical implementation of Aspect 2, interactions that occur during the deployment of AfL techniques should be carefully considered. This is further explored in Section 5.5.

Secondly, some indication of the manner in which Aspect 2 can be enacted in TEYL classrooms is provided by Theme 6a[22]. The teachers participating in this study believed that AfL enabled the learners to sustain individual and pair work, during which time the teachers could participate in an increased number of T-1L conversations thus providing opportunities for interventions that could move learning forward. This observation is interesting from two perspectives. First, it corroborates with the findings of research suggesting that teacher-learner interactions are important for learning to occur in TEYL classrooms (Ellis & Heimbach, 1997). Furthermore, Theme 6a suggests that YL teachers may not consider a whole class teaching situation (T-C interactions) as conducive to moving learning forwards effectively, despite the relatively small class sizes (8-12 children). This observation suggests that, in TEYL contexts, not the class size but the

---

[22] Theme 6: Using AfL created opportunities for a larger number of one-to-one interactions between teachers, learners and peers.

a) When students were able to work without a teacher's help, the teacher could spend time on monitoring work more effectively and providing individual support (T-1L interactions)(see Appendix 24).

teaching and/or assessment methods used in the lessons may play an important role in creating conditions favourable for learning. In an extensive review of the literature concerned with YLLs, Kubanek-German (1998) reported that 'class size was not found to be influential' (p.199) in a European context, where classes are relatively small compared with educational contexts in other parts of the world, for example, in some Asian countries. A small scale study conducted by Jagatić (1993) compared the FL teaching and learning of students aged 6-7 between two groups of different sizes: an experimental group of 15 with a control group of 30 learners. The author reports a number of differences between the two classes that she interpreted as beneficial to FLL in the smaller group. Most significantly, these included more opportunities for interactions between peers and the teacher, less teacher-centred lessons and more opportunities for continuous assessment. Jagatić's (ibid.) interpretation of the classroom conditions that could benefit learning corroborates in a number of studies that demonstrate the positive impact of interactions on FLL by children. The findings of the current study do not contradict Jagatić's (ibid.) findings but they suggest that factors other than class size alone may lead to creating conditions that are beneficial for learning and that these factors seem to be related to the adopted teaching and assessment methods. The practical implications of this interpretation are also very important as they confirm that one of the most important factors that can help to engineer effective opportunities for moving learning forward is what teachers do in the classrooms. It suggests that, in TEYL classrooms, learning can be advanced more effectively through teacher-learner as opposed to teacher-class types of interactions. This has implications for teacher initial and in-service development programmes.

The third, and final, teacher-led aspect of AfL proposed by Black and Wiliam (2009) is the provision of feedback that moves learning forward. At this point, the concept of *good feedback* needs defining. Working within the Black and Wiliam (ibid.) framework, Wiliam (2011) argues that good feedback is crucial to moving learning forward and that 'the use of assessment information to improve learning cannot be separated from the instructional system within which it is provided' (p. 4). He defines feedback as 'information generated within a particular system, for a particular purpose (...) but [feedback] requires an additional condition, that it actually improves student learning, for it to be counted as good.' (p. 4).

Theme 1d ('Teaching and learning could be integrated with assessment within an AfL framework giving meaningful explicit feedback that helps to move learning forward (from teacher, peers and own reflection') indicates that most teachers participating in this study were of the opinion that AfL enabled them to orchestrate situations in the classroom during which feedback could be provided. The teachers indicated that the sources of feedback could be teachers themselves, and the reflections of peers and learners. However, the findings for RQ2.1 indicate that feedback giving practices in TEYL classes are complex and may depend on the age related characteristics of young learners. These insights are discussed in Section 5.4.2. Of importance to the discussion of the teachers' understanding of AfL is that teachers consider feedback provision an important aspect of AfL in the TEYL context. This suggests that the teachers thought that learners aged 7-11 would benefit from receiving feedback and/or that it was appropriate to enable conditions for providing feedback in TEYL classrooms (see Section 5.4.3. for a discussion of feedback provided through AfL).

This section has argued that the teachers' understanding of the teacher-led aspects of AfL in this study is largely consistent with the Black and Wiliam (2009) framework. The themes relate to the role of the teaching in all three stages of learning: establishing where learners are in their learning, where they need to get to and how to best get there. The frequency of occurrence of the themes related to teacher-led aspects of AfL in the data suggests that teachers play an important role in the implementation of AfL.

Having explored the teacher-led aspects of AfL, the discussion continues by considering the learner-centred aspects of the Black and Wiliam (ibid.) framework.

5.3.3.2 Learner-centred aspects of AfL

The learner-centred aspects of AfL proposed by Black and Wiliam (2009) focus on the role of learners and peers. The discussion will now consider Aspects 1 and 4 (See Table 5.2, Row 2) with reference to the role of peers in AfL.

The teachers' reports about peer learning in TEYL lessons and its relationship with AfL are indicated in Themes 6b and 6c[23] (Table 5.2). There is one apparent discrepancy

---

[23] Theme 6: Using AfL created opportunities for a larger number of one-to-one interactions between teachers, learners and peers.
b) Using AfL enabled teachers to introduce more pair work in the lessons (L-L interactions).
c) Students cooperated rather than competed when working together (L-L interactions).

between the Black and Wiliam's (ibid.) framework and the understanding of AfL as reported in the findings of this study; the teachers in this study did not recognise the young learners' roles in sharing the learning objectives and criteria for success. This finding does not seem surprising given the low levels of language proficiency of the learners in the study (86 children within A1, 56 within A2 and 6 within B1). It suggests that perhaps if the learning objectives and criteria for success were connected with new language which was to be taught in the given lesson, then perhaps YLLs could not offer peer support, simply because they did not know the new language. This suggests that the use of AfL might be linked to the type of programme being taught; and in more content oriented programmes the role of peers in introducing the learning objectives and success criteria might be greater.

The findings of the current study indicate that teachers are the predominant sources of learning objectives and success criteria. This provides an interesting insight into the nature of a TEYL class by suggesting that teaching is organised according to the objectives decided upon by the teacher, presumably based on the curriculum and may not incorporate child-initiated objectives. A number of reasons for this finding can be inferred from the context in which the study was based. These include the limited number of teaching hours combined with the necessity to teach a prescribed target language within one academic year; or alternatively the low language level of the learners that inhibited a more child-centred approach. Either one of those or a combination of both could indicate that contextual factors play an important role in what happens in TEYL classrooms. Especially when we consider the curricular context of the present study, i.e. English as a school subject, the finding does not seem surprising. It seems to suggest that in contexts where language is the target content and the means of teaching, it may not be appropriate or possible to engage children actively into completing Aspect 1 ('Sharing learning intentions and criteria for success').

The findings relating to the role of peers, suggest that learners can, and do, play an active role once the learning objectives and criteria for success have been set. Findings 6b and 6c collectively indicate that YLLs enact Aspect 4 (Table 5.1, Columns B and C, Row 2, i.e. 'Activating students as instructional resources for one another'). However, Theme 6b ('Using AfL enabled teachers to introduce more pair work in the lessons') seems to indicate that in TEYL classrooms peer learning may be limited to situations when children are working in pairs. Importantly, Theme 6c ('Students collaborated rather than competed

when working together') suggests that using AfL techniques seems to encourage collaboration instead of competitiveness in situations when children work in pairs or small groups. It is useful to note that Butler and Lee (2010) reported that in a context where learners were competitive, their teachers 'worried about providing students any feedback because doing so might increase competitiveness' (p. 26). Perhaps by encouraging '*working together and not being competitive'* (T1/INT) though the use of AfL, teachers could contribute to facilitating conditions conducive to giving feedback. The relationship between feedback provision and the use of AfL is discussed further in Section 5.4.3. Additionally, Theme 6c is interesting because it indicates that the use of AfL was seen as a way of improving the quality of interactions between learners when they were completing tasks. This suggests a possible positive impact of using AfL on interactions in TEYL classrooms. This area is explored further in Section 5.5 where empirical findings about interactions are discussed

The final aspects of AfL proposed by Black and Wiliam (2009) are concerned with the relationship that the learners have with their own learning. Wiliam (2011) argues that Aspect 5 ('Activating students as owners of their own learning') 'clearly draws together a number of related fields of research, such as metacognition, motivation, attribution theory, interest and, most importantly, self-regulated learning' (p. 12). The findings reported in the current study seem to correspond to three of the areas indicated by Wiliam (ibid.), namely Theme 1a is related to the notion of metacognition; 1b, to motivation; and 5b, to self-regulated learning.

The concept of self-regulated learning is discussed by Boekaerts and Cascallar (2006) in terms of *dual processing*.

> [T]he dual processing self-regulation model distinguishes between two main pathways; the growth and the well-being pathway. Students who want to reach a goal (e.g., increasing their competence in a domain, making new friends, or helping others) initiate activity in the growth pathway because they value that goal and are prepared to put energy in its pursuit (i.e. self-regulation is energized from the top down). By contrast, students who are primarily concerned with their well-being, initiate activity in the well-being pathway; they focus on cues in the learning environment that signal unfavourable learning conditions, obstacles and drawbacks.

At such a point, they use energy to prevent (further) negative events from occurring (cue-driven or bottom up self-regulation). (Boekaerts & Cascallar, 2006, p. 202)

Boekaerts and Cascallar (ibid.) propose that, when confronted with a learning activity, learners can draw on a number of metacognitive or motivational strategies in order to regulate their learning. They define metacognitive strategies as 'orienting oneself before starting on an assignment, collecting relevant resource material, integrating different theoretical viewpoints, monitoring for comprehension, and assessing one's progress' (p. 200). The review of literature in Chapter 2 suggests that these may be beyond the ability of younger learners. It would be a valuable focus for future research to investigate whether and how the use of AfL techniques can support the development of metacognitive awareness and the control of strategies in YLLs. The other group of strategies indicated by Boekaerts and Cascallar (ibid.) includes motivation, which is usually considered by TEYL research together with other affective factors, most commonly attitude (Dörnyei, 2001) and more recently with self-concept (Mihaljević Djigunović & Lopriore, 2011). Discussion of affective factors is provided in Section 5.4.2.2.

The possible relationship between AfL and self-regulation indicated by the above interpretation would be an interesting focus for future research. It seems especially valuable as studies with children in L1 contexts have indicated that successful 5th grade readers in the U.S deployed self-regulatory strategies in reading (Owings, Petersen, Bransford, Morris & Stein, 1980). Another study (Harris & Graham, 1992), from within a similar educational context, which focused on writing, suggested that children who were able to deploy self-regulation of the type that Boekaerts and Cascallar (ibid.) consider top-down were able to work independently and monitor their own work and that their on-task engagement increased. These studies indicate that children aged 9/10 are cognitively able to self-regulate their reading and writing. It would be interesting to investigate if similar pedagogical gains can be expected in FL classrooms and whether using AfL can help to develop metacognitive and motivational skills needed for effective self-regulation, e.g. through planning, monitoring or evaluating own learning.

It is also interesting to note that a relationship between self-regulation and AfL has been investigated in tertiary education. The findings indicate the important role of feedback in facilitating self-regulation (Nicol & Macfarlane-Dick, 2006). Nicol and Macfarlane-Dick (ibid.) argued that:

'there is strong evidence that feedback messages are invariably complex and difficult to decipher and that students require opportunities to construct actively an understanding of them (e.g. through discussion) before they can be used to regulate performance' (p. 3)

They proposed that self-generated feedback may be more effective as it is easier for students to translate into action and may not evoke negative emotional reactions, which external feedback could. They proposed a model for self-regulation supported by seven feedback principles (Appendix 25). Similar studies based in TEYL contexts are not found in the literature but the concern about learners enacting feedback received from teachers and striving for positive emotional reactions to feedback certainly seem valid in TEYL contexts.

5.3.3.3 TEYL specific aspects of AfL

A number of themes identified in the current study are not represented in the Black and William (2009) framework. These are: 1f, 2, 3, 4 (partly[24]), 5a. The focus of the discussion now shifts to consider if these beliefs may offer TEYL specific interpretations of AfL and if so, what those interpretations are.

The findings of the current study indicate another teacher related aspect of AfL that seems to be specific to the TEYL context: viz. 'providing structure and focus to teachers' lesson planning'. The Black and Wiliam (ibid.) framework focuses on the processes that occur during the lessons, while the teachers' understanding reported in the current study indicates what seems to be a very important characteristic of AfL: that it is planned and hence, presumably, used purposefully. This is an important observation, especially in the light of the findings for RQ2.1, which suggest that the purpose(s) of using AfL techniques are an important factor that impacts on how AfL is implemented in TEYL classes. The findings relating to RQ2 (2.1 and 2.2) are fully discussed in Section 5.4.

Themes 2, 3 and 5a could not be mapped out against Black and Wiliam's (2009) theoretical framework because they are more pedagogically focused themes. Not surprisingly, perhaps, the teaching practitioners repeatedly discussed themes related to implementing AfL in the lessons as this is extremely valid to their everyday classroom

---

[24] There was one purpose for using AfL that was identified in the current study but was not evident in the framework discussed by Black and Wiliam (2009). This was the purpose of using AfL to measure students' confidence about their learning.

experience. However, given the focus of the discussion in this section, which is on comparing teachers' understanding of AfL with the adopted theoretical framework of this study, it is important to acknowledge that the teachers' understanding encompasses considerations related to practical implementation. Such considerations cannot be compared to a theoretical framework; therefore it seems justified to postpone the discussion of practical implementation until Section 5.4 where empirical findings about the use of AfL are discussed.

Finally, Theme 4c indicates that some teachers participating in the study considered AfL to be a way of evaluating learners' confidence about their achievement in relation to the learning objective. This is an interesting finding for two reasons. Firstly, this purpose for using AfL was more evident in the younger age group, thus highlighting differences in the purposes for using AfL with the two age groups. Secondly, and more importantly, the nature of such differences is interesting. Unfortunately, the findings do not indicate the reasons why the teachers wanted to know how confident their learners were. However, it could be speculated that, in the cases of learners who indicated low confidence in own achievement, the teachers could implement pedagogical interventions. These would aim to help the learner move their learning forwards in relation to the learning objective, if the learner's confidence was justifiably low. Alternatively, teachers could highlight the success of a certain learner, thus helping those learners who did achieve their objective despite not feeling confident, to develop confidence. This is an interesting finding as research suggests that feelings of success in language learning can be motivating to young learners (e.g. Cable et al., 2010) and thus could contribute to learners developing positive self-concept (see Section 5.4.2.2).

5.3.4 The teachers' understanding of AfL compared to Swaffield's (2011) conceptualization

The previous section discussed the teachers' understanding of AfL in the context of an influential theoretical framework proposed by Black and Wiliam (2009). However, policy makers' interpretations and resulting practical implementations of that framework have been criticised (Harlen, 2005; Swaffield, 2011), predominantly because they seem to encompass all classroom based assessment practices into AfL and not to capitalise on its for learning purpose. The criticism is connected predominantly with the way in which government policies in some educational systems have interpreted the formative nature

of AfL (see Section 2.3 for discussion of this issue). Swaffield (ibid.) provides a useful critique of the language used in the guidance documentation produced by the Department of Education in England and Wales. She argues that the way AfL had been described in those very influential documents has caused AfL to be misinterpreted such that any classroom based assessment is considered to be AfL (see also 2.3). She argues that AfL can be considered genuine only when its results are used immediately to improve the learning within which they are embedded. She considers 'now' in terms of timing and 'pupils' as beneficiaries of genuine AfL as described in Table 5.3. below.

It should be noted that the role of teachers in Swaffield's (ibid.) interpretation is not as significant as it is in the Black and Wiliam (2009) framework. Rather, the focus is on the learners. She discusses the roles and responsibilities of the various participants of the learning process but concludes that priority in defining AfL should be given to the beneficiaries and the timing of assessment. Importantly, both frameworks agree that the formative impact of assessment on learning and not any other area of educational work (e.g. programme design, teacher professional development) is what defines AfL.

Table 5.3: YL teachers' beliefs about AfL in the context of Swaffield (2011) conceptualisation

Swaffield (2011) conditions for genuine AFL	YL teachers' beliefs about AfL
Timing ('now'): happens simultaneously with teaching and learning	Theme 1
Beneficiaries ('pupils'): information which is gathered moves children's learning forwards	Themes 4 and 5 Theme 6 (see discussion in Section 5.5)

As demonstrated in Table 5.3, the teachers' understanding resembles Swaffield's (2011) conceptualisation of genuine AfL. It seems plausible to infer that if assessment is integrated with teaching and learning processes, it satisfies the immediacy requirement set by Swaffield (ibid.). Furthermore, the teachers' understanding satisfies Swaffield's (ibid.) learner-beneficiaries conditions through provision of feedback on the ongoing learning (Theme 4) and facilitating classroom conditions that may support learning (Themes 5 and 6).

5.3.6   Assessment conversations in AfL

As discussed in the final paragraphs of Section 2.3.2.3, Ruiz-Primo (2011) proposes the notion of informal formative assessment (IFA), which consists of 'small-scale, frequent opportunities teachers have for collecting information about their students' progress towards the learning goals they have in mind' (p. 16) and this is consistent with the definition of AfL adopted in the current study (see Section 2.3.2.3). The themes identified through the analysis of the data drawn from teacher interviews and the focus group indicated that 'AfL could be implemented in a TEYL context using a number of different AfL techniques' (Theme 3) which 'could be used to serve the purposes of giving and clarifying instructions, sharing aims, criteria for success, feedback and evaluating learners' confidence about their learning' (Theme 4). Theme 3 seems to correspond with the initial part of Ruiz-Primo's (ibid.) definition by suggesting that AfL may provide 'small-scale, frequent opportunities' (p. 16) for collecting assessment information. Theme 4 corresponds to the concept of 'collecting information about their students' progress towards the learning goals they have in mind' (p.16) by specifying that AfL techniques can be used to give feedback on where learners are in their learning. It also seems to go beyond what Ruiz-Primo (ibid.) suggests as it includes informing learners about the aims which the teacher has in mind and providing support in the form of success criteria for meeting those aims. This suggests that teachers' understanding of AfL is similar to, albeit somewhat broader than, Ruiz-Primo's (ibid.) definition of informal formative assessment.

The above is an important interpretation as Ruiz-Primo (2011) argues convincingly that such assessment occurs predominantly through assessment conversations, i.e. 'dialogues that embed assessment into an activity already occurring in the classroom' (p. 17). The author proposes that informal formative assessment can be effective in facilitating learning if 'assessment conversations are learning goal-guided' (p. 17) and 'dialogic and interactive in nature' (p. 18). Theme 1 (AfL implies a degree of integrating assessment with teaching and learning) suggests that it could facilitate the occurrence of assessment conversations. Themes 1a, 1b and 1c[25] suggest that conversations that occur during the

---

[25] Theme 1: AfL implies a degree of integration of assessment with teaching and learning by:
      a. making learners more aware of what they are learning
      b. focusing learners on achieving the pedagogical aims
      c. teachers continuously building AfL opportunities into lessons

use of AfL are goal oriented, while the findings presented in Section 4.4 provide evidence that they are dialogic and interactional in nature. Hence, it seems that interactions are an important feature of how AfL is enacted in classrooms and analysing their nature could offer insights into how AfL impacts learning in TEYL classrooms.

This section has discussed teachers' understanding of AfL in the context of an influential theoretical framework (Black & Wiliam, 2009) and a critique of implementing that framework (Swaffield, 2011). The discussion has highlighted that the teachers' understanding of AfL, specific to a TEYL context, is largely consistent with the current conceptualization of AfL found in the literature. A number of TEYL-specific characteristics were also identified and discussed. These include: conscious teacher effort in planning for using AfL; a relationship between AfL and affect; and a lack of peer role in sharing learning objectives and success criteria. The following section discusses the findings that emerged from investigating how the teachers implemented their understanding of AfL into TEYL.

5.4 Fitness for purpose in implementing AfL in a TEYL context

One of the aims of the current study is to report on how AfL could be implemented in TEYL classes. The literature review suggests that an overarching principle in implementing assessment is its fitness for purpose (James, 2013). The findings of the present study indicate that the purpose of using AfL guides its implementation in TEYL classes. Three main purposes of using AfL were identified: to set objectives and expectations, to monitor learning and to check achievement (see Table 4.3, p. 159). This was reported by the teachers through the interviews and the focus group. It was also evident from the observational data which indicated links between the purpose of using AfL: language skills (AfL was used predominantly when teaching productive skills (see Table 4.2, p.152), the learners' ages (the way in which each of the purposes was enacted differed between the two age groups, 7-9 and 10-11 year olds, in the current study (see Table 4.3, p.158 and Table 4.4, p. 161) and the timing (different purposes were operationalised at different stages of the tasks or lessons, see Table 4.3, p.159 and Fig. 4.1, p.144). The discussion that follows will focus on each of these areas separately. Sections 5.4.1 discusses the insights offered by the findings about the use of AfL with teaching productive skills. Section 5.4.2 focuses on the insights gained through the

analysis of difference in implementing AfL in the two age groups. Finally, Section 5.4.3 explores the insights provided by the timing of using AfL.

5.4.1 Scaffolding learning with AfL techniques

The findings indicate that AfL was used primarily during the teaching of the productive skills (writing and speaking) or in tasks that required the learners to produce output in the FL in response to prompts provided by the teacher, either orally (mainly involving vocabulary) or in writing (mainly involving grammar). The discussion attempts to interpret why such tasks might lend themselves well to employing AfL. It is important to note the term *skill* is used in this thesis to refer to what some researchers call *macroskills* (Nunan, 1989): viz. reading, writing, speaking and listening. It is also acknowledged that in any activity, different skills may be intertwined (Nunan, ibid.). Whenever, the discussion refers to a 'speaking task', it means that developing speaking was the pedagogical aim for the task and speaking was likely to be the skill which the learners were expected to use most in the process of accomplishing the task.

The review in Chapter 2 points to the importance of task type in obtaining useful assessment data and supporting language learning in L2 contexts. Two issues were highlighted: familiarity with the task type (Pinter, 2007) and the amount of structure provided by the task (Skehan & Foster, 1999). The discussion considers both areas.

First, Skehan and Foster's (ibid.) distinction between inherently structured tasks and ones that require more on-line processing is adopted here to facilitate interpretation of the finding that AfL is less frequently used with listening and reading tasks than with speaking and writing tasks. Although their research was based in adult contexts, adopting this criterion seems justified here as most teachers in the current study also reported that the amount of structure provided in a task may be important in task completion: 'it [AfL] helps to teach, you could say, because it gives you structure (...) I think it's also very helpful for learners in completing tasks' (T5, INT). In order to interpret this finding, it seems useful to first consider whether the types of tasks that tend to be used in TEYL classes for teaching the four language skills differ in terms of the degree of internal structure and the amount of on-line processing required. As the current study did not evaluate the teaching and task types implemented in TEYL classes, the discussion considers the task types discussed by Szpotowicz and Szulc-Kurpacka (2009). This book was selected for the following reasons: it provides a recent perspective on teaching of

young learners in TEYL contexts, was published in Poland for the educational context of that country and was the most frequently borrowed book from the study school's resource cupboard during the data collection period (as was evident from the sign-out sheet on which teachers recorded the books which they borrowed from that resources area).

Table 5.4 below illustrates the different task types that may be used to teach each of the four language skills. Each task was judged by the researcher as to whether it was likely to be structured or whether it might require on-line processing. As a result, the task types drawn from Szpotowicz and Szulc-Kurpacka (ibid.) were allocated to one of the two categories in Columns A and B. The allocation of the task types was based on the descriptions of these tasks provided by Szpotowicz and Szulc-Kurpacka (ibid.), an analysis of examples of tasks drawn from the course books used in the study school and the researcher's ten years' of experience in TEYL. Columns C, D and E report which AfL technique types were found to be used with each of the language skills in the present study. As fitness for purpose was discovered to be an important factor in implementing AfL, the techniques are reported in three columns, each one corresponding to one purpose for using AfL: setting objectives and expectations (Column C), monitoring performance (Column D) and checking achievement (Column E).

Table 5.4: Task types used in TEYL classrooms to teach the four language skills and AfL technique types used for each skill

	Skill	Szpotowicz and Szulc-Kurpacka, 2009		The current study: AfL techniques used for each purpose		
		Column A: Task types which are likely to require on-line processing	**Column B: Task types which are likely to have an inherent structure**	**Column C: Setting objectives and expectations**	**Column D: Monitoring performance**	**Column E: Checking achievement**
Productive skills	Speaking	1. Role plays 2. Information gap activities 3. Describing people and objects (narrative)	1. Singing 2. Class surveys 3. Asking and answering questions	Success Criteria Increased Thinking Time Learning Partners	Increased Thinking Time Learning Partners	Smiley Faces Find the Fib Traffic Lights Learning Partners
	Writing	1. Creative writing: descriptions and stories 2. Functional writing: letters, postcards, film/book reviews 3. Dialogue journals	1. Writing labels on pictures/objects 2. Copy/write words in different categories 3. Copy/write a sentence and add a missing word 4. Running dictations	Success Criteria Sharing Good and Bad Models	Indicating Mistakes Without Explanation Perfect Purple and Red to Remember	Two Stars and a Wish Perfect Purple and Red to Remember Next Steps Colour Coding Learning Partners
Receptive skills	Listening	-	1. 'Listen and do' activities 2. Bingo 3. Simon Says 4. Chinese whispers 5. Picture dictation 6. Mime what you can hear 7. Follow the route 8. Listening grids 9. Listen and point	-	-	-
	Reading	-	1. Matching words and phrases 2. Labelling pictures/objects with words 3. Predicting words from initial sounds 4. Rearranging jumbled words/letters 5. Guess the missing word 6. Arranging sentences in a text 7. Memory games 8. Clapping syllables	-	Increased Thinking Time Indicating Mistakes Without Explanation	-

Table 5.4 indicates that most task types deployed in TEYL classrooms tend to have an inherent structure. However, some of those used when teaching productive skills are likely to require more on-line processing. In this study, AfL techniques were found to be used predominantly with tasks that focused on speaking and writing for each of the three main purposes of using AfL. Hence, it seems plausible to infer that AfL techniques might provide scaffolding that facilitates on-line processing by young language learners. This interpretation could provide some explanation of how AfL can support YLLs in completing production tasks. However, understanding the nature of this support would require further research.

The review of the literature in Chapter 2 also suggested that learners may be able to perform better and collaborate more effectively when they are familiar with the task type (Pinter, 2007). The findings of the current study indicate that a relatively small number of AfL techniques (n=18) could be identified through the data. The findings also indicate that individual teachers implemented AfL in idiosyncratic ways, using a different number of technique types: between 2-12 (in the cross-sectional phase) and between 4-15 (in the longitudinal phase). It is conceivable that, over a certain period of time, children could become familiar with the technique types. This interpretation is supported in teachers' reports about how AfL could provide scaffolding[26]. It would be of value for future research to investigate whether familiarity with technique types could support learners in completing a task.

The above discussion seems to suggest that YLLs may need additional support to complete tasks requiring on-line processing and that AfL techniques might provide such scaffolding. It would be a valuable focus for future research to investigate in more detail how AfL could provide scaffolding. In adult contexts familiarity with tasks structure has been reported to have a positive impact on measures of fluency, complexity and/or

---

[26] As reported in Section 4.2.4:

Theme 5: Using AfL helped the learners sustain individual, pair and group work for longer, without the need for support from the teacher by:

    A. providing scaffolding

        'they get used to this that they're given steps in what to do and they need this and they want this and they really (1) and it makes them actually complete the whole task rather than shout out finished (1) because they know exactly what they need to do' (T5/INT)

accuracy within different language skills (e.g. Bygate, 1996; Skehan & Foster, 1999). The current study provides some evidence that by using AfL some gains in supporting better production might be possible in TEYL classes with learners as young as 7 in them. Arguably, more research is needed in this area to investigate how using AfL can impact on various measures of language production in TEYL contexts. The literature review presented in Chapter 2 suggests that children learn languages implicitly and may not have the cognitive resources to analyse language (Section 2.2.1). Hence one of the challenges of such research would be to establish whether all measures used in adult contexts, especially accuracy, are valid for TEYL contexts.

Considered from a socio-cultural perspective, the interpretation that AfL techniques may provide scaffolding which facilitates on-line processing by YLLs seems to raise a further question, namely: whether or not, if given appropriate scaffolding through AfL, learners would be able to complete tasks within their ZPD more effectively. Understanding this would have important implications for explaining how AfL can contribute to facilitating learning. If the affirmative were true, it could be concluded that AfL could be used to facilitate learners' work within their ZPDs. As discussed in Chapter 2, by creating an intermental development zone (IDZ) through interaction, learners with their teachers or their peers can work together to meet the learning objectives. The current study has provided preliminary insights into this area by analysing those interactions between teachers, learners and peers that occurred during the use of AfL. This is reported in Section 4.4 and discussed in Section 5.5.

5.4.2 Age-related factors in implementing AfL

One of the aims of the current study is to investigate how AfL can be implemented in classes of learners aged 7-11. The findings indicate that implementation may differ depending on learner age. This section discusses the insights obtained from the findings about differences in when and for what purposes, AfL techniques were deployed in the two age groups. Section 2.2 has discussed the issues relevant to how children develop cognitively and how they learn FL in childhood while Section 2.3 explored the conditions for assessment practices that would be appropriate for YLLs. These considerations are helpful because they inform the interpretation of the findings about implementation of AfL. The aim of the current section is to consider how the findings of the current study

related to the fitness for purpose in using AfL could extend the understanding of assessment and learning in TEYL reviewed in Sections 2.2 and 2.3.

5.4.2.1. AfL and cognitive demands of following instructions

As is evident from Table 4.3 (p.159), within each of the main purposes for using AfL, there were some age-related differences. Within Purpose 1 (Setting objectives and expectations), the teachers in the study tended to use AfL to discuss the criteria for success and expected outcomes in both age groups, but in the younger age group, they also used AfL to give and clarify instructions. This latter use of AfL was not observed in the older age group. This finding may be linked to the cognitive development of children in the younger age group. As discussed in Section 2.2.1.3.1, children have short attention spans. This may simply mean that teachers may have found it more helpful to use pictorial representations of instructions: recorded visually and displayed, for instance, as step-by-step Success Criteria. This seems an interesting interpretation when considered together with the development of memory in childhood.

As discussed in Section 2.2.1.3.3, working memory (WM) has been argued to be important in processing and comprehending language (Gathercole & Baddeley, 1993; Wen & Skehan, 2011). However, WM has a limited capacity that is shared between storage and processing, and when more storage is required less processing can occur (Daneman & Merikle, 1996). It seems possible that by providing tangible scaffolds for remembering and understanding instructions for tasks, teachers might have facilitated the process of freeing up some of the WM space for learning. This would mean that when teachers used AfL techniques to give and clarify instructions in the lessons with the younger participants, they lessened the requirements posed by the instructions on the WM. Hence, children were able to use more of their WM on completing the tasks. This interpretation is supported by research suggesting that, as processing becomes more effective with age (Siegler, 1994), it takes up less of the WM. Hence, as they mature, learners rely on the exemplar-based (Skehan, 1996) system to a lesser degree and do not have to remember instructions verbatim as their younger counterparts would. Therefore, older children might not need additional support to understand and follow instructions. This interpretation is consistent with Case (1972), who reported that physiological resources needed for effective processing develop faster from the age of 8, and with a study reported by Siegler (1994) indicating that the ability to use WM effectively

develops between the ages of 6-15. The findings of the current study could be interpreted to mean that, as learners were developing in using their WM effectively, the older age group (10-11) had reached a stage when less support was needed from the teacher to understand instructions. However, in the younger age group (7-9), the learners benefitted from support in understanding and following the instructions that were provided by the majority of the teachers in the form of AfL techniques.

The discussion above has direct implications for assessment. It indicates that, by employing techniques that supported the learners' understanding of instructions, the teachers might have been able to create conditions that allowed them to collect valid assessment data. This could be possible because the learners' performance in tasks, which teachers could observe, was less likely to be inhibited by difficulties in understanding or following instructions. This discussion also resonates with the critique of Piaget's study, which considered lack of age-appropriate instructions to be a methodological shortcoming in Piaget's work (Winer et al., 1992).

5.4.2.2 AfL and the affective perspective on learning

Within Purpose 2 (Monitoring learning), there is more evidence of employing AfL techniques to measure the learners' confidence about their learning in the younger age group. The research reviewed in Section 2.2.3 has indicated that children's motivation, self-concepts and attitudes towards FLL may change over time as they accumulate their own experiences of FLL (e.g. Mihaljević Djigunović & Lopriore, 2011; Nikolov, 1999). The findings of the current study indicate that AfL could be used to measure learners' perceptions of their own achievement. In cases where children are having difficulty recognising their own achievement, presumably teachers could intervene to make them more aware of what they have learnt. An important feature of AfL could be to provide teachers with this type of insight. It would be especially useful as research has suggested that children's motivation can decrease over time (Matsuzaki Carreira, 2006) but that perceptions of success could contribute to learners developing positive motivation towards language learning (Cable et al. 2010; Lopriore & Mihaljević Djigunović, 2009; Masgoret et al., 2001). This interpretation has useful pedagogical implications when considered in the light of the suggestion that positive attitudes towards FLL could be maintained given favourable conditions (Cenoz, 2003) as it suggests that using AfL could help teachers to create such conditions by identifying how confident learners are and by

nurturing the development of confidence in those needing it. It also seems useful to note that motivation has been interpreted (Boekaerts & Cascallar, 2006) to be an internal part of self-regulation in the learning process, which may support learning.

However, the above discussion does not explain why AfL techniques might have been implemented to measure the learners' confidence about their own learning more frequently in the younger age group. One interpretation could be that some teachers observed that it was somewhat beneficial for learning in that age group. The research reviewed in Chapter 2, provides insights into the relationship between positive affect and achievement in the productive skills (e.g. Mihaljević- Djigunović, 2006) and in listening and oral production (Mihaljević Djigunović & Lopriore, 2011). However, that body of research does not seem to indicate that 7-9 year olds' achievement would benefit more from positive affect than that of 10-11 year olds. On the contrary, it was found in the ELLiE study that affective factors had a impact on achievement that was stronger at the age of 10-11 than at the age of 7-8. Hence, it cannot be claimed that the feeling of success seems to be more important for the younger children. Therefore, there might have been a different reason for employing more AfL techniques to measure the learners' perception of their own success in the younger age group (7-9 year olds) than in the older one (10-11 year olds).

Perhaps, it would be more informative to consider the finding from the perspective of cognitive individual differences (IDs), taking into account especially the development of attention in childhood. It seems that AfL techniques were a useful tool in providing tangible prompts (e.g. a green Traffic Light, a happy Smiley Face[27]) that would draw the learners' attention to their achievement. Thus AfL might be helpful in drawing learners' attention to the issues that their teachers would like them to notice. This may be useful in classes of younger learners because their ability to direct attention and ignore unnecessary information might not yet be well developed (Ridderinkhof & van der Molen, 1997). If this were the case, then it may be interesting to note that the teachers wanted their learners to recognise that they were being successful in their language learning. Perhaps, the teachers wanted to achieve the same effect in the older age group but simply used AfL for that purpose less frequently and were able to bring about feeling on success it in a different way. The above interpretation seems more informative about the nature of AfL

---

[27] Please see Appendix 18 for descriptions and examples of AfL techniques.

than about its effect in TEYL contexts. This points to the conclusion that the AfL techniques reported by some teachers were appropriate for drawing the younger learners' attention to their own achievement.

It is useful to note that AfL techniques could be used for supporting learners in building up positive self-concepts by making them notice their own achievements. This has important pedagogical implications, especially when considered in the context of the body of research suggesting that developing motivation and a positive attitude to FLL are often explicitly stated as expected outcomes of TEYL programmes (Edelenbos et al., 2006). AfL could contribute to meeting these outcomes. Further research in the future is needed to investigate the relationship between the use of AfL techniques and affective factors.

5.4.2.3 Types of AfL technique and levels of L1 literacy

The findings of the current study also suggest that the techniques used with the younger age group (7-9 year olds) relied on pictorial representations whereas the ones used with the older age group (10-11) depended more on the learners' literacy skills. This finding can be explained by the children's development in both L1 and FL. In Poland, at the time of the study, children started primary education and began developing their literacy skills at the age of 6/7 (Leowiecki, 1999). Hence, 7-9 year olds were at the early stages of literacy development, whereas their older counterparts tended to be more confident readers and writers. The findings that AfL techniques used with 10-11 year olds are more reliant on literacy skills could indicate that the implementation of AfL in TEYL classes depends on contextual factors, such as learner characteristics.

5.4.2.4 AfL and metacognitive development

Another area where age related differences are highlighted is feedback. The findings of the current study indicate that the feedback provided to the younger learners (7-9 year olds) was predominantly from the teacher or peer(s) and self-assessment was rare. Self-assessment was found in the older age group (10-11 year olds). This finding can be explained in terms of the development of metacognitive control (Flavell et al., 1993), i.e. the effective use of skills such as monitoring and assessing one's own progress (Boekaerts & Cascallar, 2006). As discussed in Section 2.2.1.3.2, research indicates that, although learners develop metacognitive awareness around the age of 3-5, they only develop the

control aspect of metacognition around the ages of 8-10 (Flavell et al. 1993) or 9-10 (Owings et al. 1980). Hence, it seems possible that monitoring and self-evaluating achievement was beyond the developmental stages of learners aged 7-9 in the current study. This interpretation corroborates with the findings reported by Gu, et al. (2005), who found that, in a FL context, children aged 9 were able to use a wider range of metacognitive strategies than 7 year olds. This would explain why self-assessment was less evident in the data collected about the younger age group. This finding has implications for understanding assessment in TEYL contexts. It highlights the importance of accounting for the metacognitive development of children when providing feedback on assessment. Specifically, it warns against assuming that YLLs are able to reflect on their own learning and suggests that older children may be more capable of doing so than younger learners.

It seems important to note that successful young language learners have been shown to be more effective at using strategies to monitor their own learning (Chamot & El-Dinary, 1999). Additionally, children aged 9-12 have been shown to be able to develop metacognitive awareness and control in FLL (Goh & Taib, 2006; Vandergrift, 2002). Another strand of research concerning the development of metacognition in FL contexts has shown that feedback received from teachers or peers could contribute to learners developing the ability to monitor and evaluate their own learning (Butler & Lee, 2010; Hawe & Dixon, 2014). Hence, an interesting focus for future research would be to investigate if using AfL techniques could facilitate the development of metacognition in YLLs. This could have pedagogical implications, especially when considered together with the research on self-regulation suggesting that metacognitive strategies, alongside motivation, are an important component of the growth pathway of self-regulation that can benefit learning.

5.4.3 Feedback through AfL

The findings of the current study that indicate a relationship between the timing and the purpose of using AfL suggests that in the initial stages of a lesson and/or a task AfL was predominantly used to set the objectives and expectations; during a task it was implemented to monitor learning, which included providing feedback on elements of the task or on the process or completing it; and in the final stages of a lesson or a task it was used to focus on providing feedback on achievement. This suggests that feedback

provision was an important element of AfL and that it served two purposes as specified above.

The findings of the current study indicate that teachers use a greater diversity of technique type for providing feedback than for any other purpose (Fig. 4.1, p. 154). It seems useful to consider this finding in relation to contextual (the timing of providing feedback) and developmental (the characteristics young learners) factors. Feedback provision occurs in the final parts of tasks. This implies that children are expected to focus their attention on the task at hand for a certain amount of time before they receive feedback. As discussed in Chapter 2, young learners' control of attention develops as they mature. By considering these two factors, it seems plausible to infer that introducing diversity in technique type may be appropriate when working with young children as it helps to address the issue of their short attention span. This has a number of important pedagogical implications. First, Pinter (2011) suggests that learners need to focus their attention on what they are doing in order to commit new information to memory. It would be interesting for future research to investigate whether the use of AfL techniques and, more specifically, diversity in technique type, can contribute to learners' on-task engagement.

It also seems relevant to the current discussion to recall that between-teacher variance was evident in the frequency and diversity of implementing AfL techniques. This finding is relevant to the current discussion because it was found that most teachers tended to implement greater diversity in technique type and increased the diversity over time. Similar observations were not made for frequency of use. This corroborates with the discussion in the previous paragraph and confirms that diversity in technique type might be important in TEYL classes.

It is vital for the present discussion to consider the types of feedback AfL techniques were used for. Reflecting on the nature of each AfL techniques (Appendix 18), it could be inferred that the techniques which teachers reported as 'used for feedback' provided explicit positive information about what children were able to demonstrate in their performance: e.g. 'two stars' in TSAW or 'perfect purple' in PPRR. These techniques also included an element which allowed setting a developmental target for the learner: 'the wish' in TSAW and 'red to remember' in PPRR. Another group of techniques allowed communicating about the degree to which the learning objective was met. For example, green in Traffic Light meant 'fully met', amber: 'partially met', and red: 'not

met'. The findings indicated that this group of techniques was often used in conjunction with others that were explicitly used to set developmental targets: e.g. as with Next Steps. Hence, the notion of 'feedback' in the findings from TEYL classes indicate that AfL techniques were used to provide explicit feedback about success in learning in relation to the learning aims and developmental targets for the future.

The above understanding of 'feedback' is different to the one often reported in FL research. A large body of research focusing on the impact of feedback on developing speaking (e.g. Mackey & Silver, 2005; Oliver, 1995) in FL, considers implicit feedback as recast 'where a teacher reformulates a learner's non-target-like form' (Oliver & Mackey, 2003, p. 519) or as negotiation strategies that include repetition, clarification requests and comprehension checks. The focus of that research has been on understanding how different types of teacher feedback could impact on learners' performance, usually according to fluency, complexity and accuracy measures. Oliver and Mackey (ibid.) also note that in this tradition '(w)hen feedback is provided in response to learners' non-target-like production, it is generally termed negative feedback.' (p. 519). Feedback understood in this way is relevant to the discussion of classroom interactions in Section 5.5.

The current discussion will focus on feedback on writing. Research investigating feedback on writing in L2 classes comes predominantly from tertiary education in EAP/ESL contexts (e.g. Conrad & Goldstain, 1999; Ferris, Pezone, Tade & Tinti, 1997; Paulus, 1999) with a very limited number of studies of TEFL settings with younger learners that only examine the secondary school context (Furneaux, Paran & Fairfax, 2007; Lee, 2004). One of the reasons for this lack of research on practice and the impact of feedback on writing in TEYL contexts at primary level may be due to the fact that children in primary schools are rarely expected to write long texts in a FL. Nevertheless, the development of writing skills is normally included in TEYL curricula. Szpotowicz and Szulc-Kurpacka (2009) suggest that it can focus on letter, word, sentence and/or text level. The research on feedback, however, has focused predominantly on text level writing. This implies that little is known about feedback practices and their impact on the development of FL writing in childhood.

The research available from secondary contexts suggests that teachers use direct (words) or indirect (symbols) strategies when marking writing and that they focus their feedback predominantly on error correction, which is something that the students rely on (Lee,

2004). Broader insights are offered by Furneaux et al. (2007) who report finding in their study that the secondary school L2 teachers adopted six different roles through their feedback provisions, with those of Provider and Initiator being the most common and the other roles being: Supporter, Adviser, Suggester, Mutator (for definitions of these roles, see Appendix 26). The Provider role entailed annotating a learner's writing with 'the correct form by substitution, addition, deletion or reordering of an item of language or punctuation' (p. 78). The Initiator '(a)lerts by providing a specified (lexical, grammatical, stylistic, semantic, discoursal, mechanical) or unclassifiable (dotted lines, circle, question mark etc.) alert (...) provided there is no actual correction' (p. 76). Although the Furneaux et al. (ibid.) study did not explore the impact of these different types of feedback on learning, some evidence from a different educational context suggests that adult students believe that by exploring problematic areas themselves, as opposed to being given the correct form by the teacher, they learn more (Chandler, 2003). Similar insights into learners' perceptions of feedback and its effectiveness or empirical, quasi-experimental investigations about the impact that teacher feedback may have on writing are not available from within TEYL, or even secondary school, contexts. This highlights a gap in TEYL research.

The present study extends the body of research reported above by offering insights into feedback on writing practices in classes of 7-11 year olds. The findings suggest that teachers, peers and learners can provide written feedback on YLLs writing. Furthermore, the findings suggest that AfL techniques can be used as strategies for such feedback. The AfL techniques identified in the current study as used for the provision on feedback on writing in TEYL classes were TSAW, PPRR, CC, NS, IMWE[28]. The first three of these techniques include components that allow the feedback provider to highlight the positive aspects of the written work. This might facilitate the role of a Supporter, i.e. somebody who '(r)esponds positively to the text with either symbols (++) or comments' (Furneaux et al., 2007, p. 77). This finding could be partially explained by the fact that, unlike Furneaux et al.'s (ibid.) study, the teachers in the present study knew the learners that they provided feedback to and had presumably developed positive relationships with them. It is possible to infer that teachers could purposefully choose the techniques that

---

[28] TSAW – Two Stars and a Wish, PPRR – Perfect Purple and Red to Remember, NS – Next Steps, CC – Colour Coding, IMWE – indicating mistakes without explanation. For descriptions and examples of each technique, see Appendix 18.

may facilitate the Supporter role in order to foster that positive relationship and/or provide encouragement to young learners. Regardless of such contextual considerations, the findings of the current study indicate that AfL has the potential for facilitating the Supporter role in providing feedback on writing to YLLs. In doing so, it could raise children's awareness of their success in FLL and, thus could be motivating. This interpretation is supported by research indicating that learners' feelings of success may have a positive impact on their motivation (Cable, et al. 2010; Vilke & Vrhovac, 1995, cited in Mihaljević Djigunović, Nikolov and Ottó, 2008). Motivation is discussed together with other affective factors in Section 5.4.2.2.

Another function of feedback provided with the help of AfL techniques is to focus learners' attention on areas that need improvement. This can be done either directly (TSAW, NST) or indirectly (CC, PPRR, IMWE). These techniques could lend themselves to the roles of Initiator (CC, PPRR, IMWE), Suggester or Adviser (TSAW, NST) as defined by Furneaux et al. (ibid.). These interpretations are consistent with Lee (2004), whose study focused solely on grammar correction, and who reported that the teachers provided both direct and indirect feedback. Furneaux et al. (ibid.) extend Lee's (ibid.) focus in their investigation to include feedback on lexis, style, discourse, semantics and the mechanics of writing. The current study suggests that three out of the six roles identified by Furneaux et al. (ibid.) that teachers adopt while providing feedback on writing (Supporter, Advisor, Initiator) could be enacted with AfL techniques in classrooms with 7-11 year olds. The Supporter role could be enacted by providing positive comments (for example, Point 6, Appendix 18); the Advisor role by providing explicit points for development (for example, Point 12, Appendix 18); and the Initiator role by indicating mistakes without explanations (for example, Point 16, Appendix 18). Interestingly, the most common role in Furneaux et al. (ibid.), i.e. Provider, does not seem to be facilitated by AfL techniques. This might be explained in terms of a potentially greater focus on meaning and not on language form in TEYL classes. This interpretation is consistent with the body of research discussed in Chapter 2, which suggests that children learn implicitly and benefit from implicit instruction. It would be useful for future research to explore this area further by investigating the roles that teachers and peers adopt when providing feedback on writing through AfL techniques. It would also be useful to investigate whether teachers supplement feedback given through AfL by,

perhaps, adopting the Provider role and, more importantly, gain insights into the impact that feedback on writing has on learning and on learner motivation in TEYL contexts.

To supplement this discussion, Appendix 18, contains examples of feedback on writing provided through AfL techniques. These should be treated solely as examples because the study was not designed to collect systematic data about feedback on writing.

5.4.4 Factors which could facilitate or inhibit the implementation of AfL

A number of findings discussed in the current chapter indicate that the implementation of AfL is closely linked to the context in which it happens. The findings about the timing of using AfL provide insights into this within the context of a lesson by indicating the need for implementation to account for the individual characteristics of the learners, many of which are related to their age and stages of cognitive development (5.4.2). Furthermore, as shown in the studies of AfL in secondary school TEFL (e.g. Lee & Coniam, 2013) contexts, the implementation of AfL is related to the curriculum and the educational context in which it happens as. In fact, according to Lee and Coniam (2013), implementing AfL may be hindered if it is not compatible with the broader context of a school and educational system, especially in contexts where students are expected to take high-stake summative tests. A similar lack in the compatibility of AfL with the practice of summative reporting was found by some teachers in the present study (Theme 2c, Appendix 24).

The finding that teachers understood summative and formative practices as incompatible is worthy of further consideration. This issue has been discussed by Lee and Coniam (2013), who researched the implementation of AfL in writing in the secondary school context in Hong Kong. They argued that the exam driven educational culture, which valued accuracy, led to feedback giving practices which inhibited the formative function of AfL. First, the teachers tended to correct every error in writing and, secondly, they provided numerical grades for each piece. Lee and Coniam (ibid.) refer to a study by Butler (1988), who demonstrated that when presented with a numerical grade, learners may not pay attention to the formative comment that accompanies the grade. In the current study, the learners were not preparing for any high stake external exams. Therefore external examinations did not feature as a factor that might inhibit the implementation of AfL. Other factors might have contributed to the teachers' experiences as they reported them. These are discussed below.

It is important to recall that the findings of the current study also indicate that at certain times during the semester when summative reports were due, the teachers tended not to use AfL techniques, hence providing empirical evidence that summative reporting can inhibit the implementation of AfL. This raises the possibility of an external factor, other than examinations, that shaped classroom assessment practice: the school's reporting policy. As the teachers were expected to report numerical grades to parents, they tended to opt for employing testing procedures in the lessons preceding the reporting. Furthermore, it could not be ignored that there was another stakeholder in the process of reporting assessment data: namely the children's parents. The findings of the current study indicate that parental expectations impacted on the teachers' assessment practice. Presumably, such a consideration is especially valid in a context where the parents pay fees for the language course since this affects the teacher-parent dynamics which may include aspects of a service provider – customer relationship. Hence, parental opinions may possibly have a greater influence in a fee-paying context than in other non-fee-paying settings.

It is also useful to consider the tension between the formative and summative functions of assessment from a theoretical perspective. Harlen and James (1997) propose a number of distinctions between the two functions. Most notably, they argue that progress is conceptualised on an individual basis in AfL, while in summative assessment it is concerned with public criteria. They propose that 'it is not helpful, to be concerned with strict criterion-referencing in formative assessment' (p. 366). This suggests that externally mandated criteria may not be compatible with embedding AfL in the classroom as they may not allow the flexibility that is needed for setting individual steps for learners, who might progress at different rates. Harlen and James (1997) identify the requirement for reliability as another important distinction between the formative and summative functions of assessment. They argue that validity and usefulness are important in the formative function of assessment, while summative assessment 'requires methods which are as reliable as possible without endangering validity' (p. 373). This seems to suggest that if reliability is not required in AfL, while summative tests must ensure it, the two functions cannot be subsumed under the same paradigm. Indeed, in the psychometric testing paradigm, the reliability and validity of summative assessment is a requirement. But as Teasdale and Leung (2000) argue 'psychometric approaches may not provide an adequate response to pedagogic and policy developments' (p. 163). For the pedagogic

and learning functions of assessment (Rea-Dickins, 2001) to be realised in the 'pedagogical paradigm' (Torrance, 1995, p. 55), reliability may be compromised. This highlights an important issue in considering tensions between AfL and summative assessment.

What is important to note is that the findings of the current study also indicate a number of individual and contextual factors that might facilitate the implementation of AfL. The contextual factors are to do with a generally perceived compatibility of using AfL with the teaching methods employed in TEYL classes. This is similar to the findings of Gattullo's (2000) study reporting that teachers implemented techniques that were easily compatible with their teaching. This finding corresponds to Cheng et al. (2004), who discuss the importance of the nature of the course and the needs and levels of students in shaping classroom-based assessment. Cheng et al. (ibid.) also include teachers' knowledge and experience as well as their attitudes and beliefs in the list of factors that may shape assessment. The longitudinal perspective of the current study provides evidence that teachers consider their own familiarity with technique type as an important factor in shaping their practice over time. Additionally, the teachers' experience of observing AfL being used by colleagues was identified as a factor that facilitated the implementation of AfL. This finding is consistent with the findings reported by Lee and Coniam (2013), who found that teachers' previous experience of using AfL and opportunities for collaboration with other teachers facilitate the implementation of AfL.

It is interesting to note that the contextual differences between the teachers participating in this study were not obvious. Nonetheless, the teachers seemed to implement AfL with different levels of diversity and frequency. Four types of implementation were identified (intra- and inter-lesson frequent/diverse; intra-lesson frequent but inter-lesson infrequent/diverse; intra- and inter-lesson infrequent/not-diverse; intra-lesson infrequent but inter-lesson frequent/not-diverse). Additionally, the contextual considerations did not change significantly between 2012 and 2013 but changes in the frequency and diversity of using AfL in the practice of individual teachers did occur. This provides an interesting insight into the nature of the factors that may affect the implementation of AfL, because it suggests that the individual characteristics of the teachers could account for some of those changes. These might include teachers' beliefs, previous learning experiences,

training and professional practice (Borg, 2003). However, investigating this area in detail is beyond the scope of the current study.

5.4.5 The construct of assessment in TEYL contexts

As discussed in Chapter 2, more clarity is needed on what level of language proficiency could realistically be expected from young children in ELL programmes (Enever, 2011; Johnstone, 2000). This signifies that the construct (the what) of assessment of YLLs is a valid question for research to address. However, it has been argued that available scales of reference, e.g. the CEFR, are not appropriate for YLLs in terms of setting attainment targets (Enever, 2011) and that they may need to be adapted to smaller steps which would be sensitive to the relatively slow progress made by children (Nikolov & Mihaljević Djigunović, 2011). Some attempts made to date to clarify that issue have included adapting the CEFR descriptors for use with YLLs (e.g. Hasslegreen, 2005) and collecting data of learners' performance to build up an empirically driven understanding of what performance at each level could constitute (Huhta et al., 2014). However, it should be noted that the reported insights into learner performance only refer to an adolescent context (13-16 year olds). Additionally, Butler (2009) argues that the issue of clarifying what constitutes good performance in the primary school is vital as it also has implications for evaluating the quality of teaching in ELL programmes.

Although the primary aims of the current study did not include investigating what constitutes the construct of assessment in the TEYL classrooms, the findings about the implementation of AfL providing some interesting insights into that area are worthy of mention. The evidence comes from two sources. One, the AfL techniques identified in lesson observations indicate that AfL practice offers opportunities for assessing learners' ability with reference to the task in hand. Two, the analysis of conversations taking place during the implementation of AfL points out that teachers and learners focused on assessing and achieving the learning objectives set for the lessons. These two strands of evidence indicate that, in the context of the current study, what is assessed (i.e. the construct of assessment) is closely related to the teaching objectives set for the tasks or the lessons. This suggests that it may be necessary to consider more situated and context sensitive definitions of the construct of assessment in TEYL contexts.

The above argument is further supported by calls for aligning assessment practice with the foci of the teaching programmes (Inbar-Laurie & Shohamy, 2009) and is consistent

with the calls for developing attainment targets that would reflect the gradual development of YLLs abilities (Nikolov & Mihaljević Djigunović, 2011). Applying a more situated approach to clarifying the construct of assessment also promises the opportunity for integration within each of the four models of implementation of ELL as proposed by Johnstone (2009, see Section 1.2).

Another important consideration in discussing the construct of assessment in TEYL contexts is related to the development of positive affective dispositions: motivation especially being one of the expected outcomes of TEYL programmes (Mihaljević Djigunović, 2015). The findings in the current study, based on teacher interviews and lessons observations, suggest that a number of AfL techniques (e.g. Smiley Faces, Thumb Up/Down, Traffic Lights) offered opportunities for the learners to indicate how confident they felt about their own achievement: hence, presumably building up their self-concepts and/or enhancing their motivation (Masgoret et al., 2001). These findings are more evident in the younger age group (7-9 year olds), suggesting that this component of the construct of AfL may be more appropriate for younger learners. It seems useful to note that the younger learners were simultaneously at a lower level of proficiency in English (see Table 3.2, p. 101). Hence, the question remains whether it is the age or the level of proficiency that determines whether or not the development of an affective disposition should be incorporated in the construct of assessment in a TEYL context. Nonetheless, both, the age and the level of English constitute the context in which AfL was implemented. Hence, the impact of either, or both, in determining what is important to assess, seems to support adopting a situated approach to clarifying the construct of assessment.

Further support for considering a situated approach to clarifying the construct of assessment in TEYL contexts is provided by research on the development of memory in childhood. The review in Chapter 2 indicates that young learners may prefer to operate within a lexical mode of communication due to their reliance on the exemplar-based system. But as the effectiveness of the rule-based system develops with time, learners become more able to analyse language. This has implications for teaching YLLs by suggesting that children may rely on implicit learning without analysing the language. This has also been suggested by studies exploring the importance of the starting age of instruction on achievement (e.g. Muñoz, 2006). Working on the principle that '(i)t is axiomatic that the way that children learn best be reflected in the way that they are

assessed' (McKay, 2006, p. 47), it seems reasonable to argue that assessment techniques should account for the cognitive development of learners, especially bearing in mind their ability to analyse language. However, even in a class of children who are of the same age, individual learners may differ in their cognitive development. Therefore, as suggested in Section 2.3.2.2, children may be able to reach the same learning objectives by following different trajectories. Investigating this in detail would constitute a useful focus for future research. But what is important to the current discussion is the suggestion that adopting a situated approach could offer the possibility of incorporating individual characteristics, such as the learner's stage of cognitive development, into the construct of assessment.

This section has proposed that, given the variety in programme type Johnstone's (2009) four models and the affective and cognitive individual differences between young learners, it may be appropriate to consider adopting a situated approach to clarifying the construct of assessment in ELL contexts. This would address the need for aligning assessment practice with the foci of the teaching programmes, the way in which learners learn and for adapting the existing reference frameworks to small steps that would evidence the gradual progress made by YLLs.

5.4.6 Concluding remarks for the section

The above discussion has highlighted the importance of fitness for purpose in implementing AfL in TEYL contexts. It has suggested that implementing AfL might be important in providing scaffolding for speaking and writing tasks, especially those that are likely to require on-line processing and do not have an internal structure. With regards to fitness for purpose in different age groups, it seems that the cognitive development of learners is an important factor in shaping the implementation of AfL. Most importantly, the development of working memory, cognitive inhibition and metacognitive control have been highlighted as the factors that might shape feedback giving practices in AfL. The discussion has also indicated that AfL may be linked to fostering positive affective dispositions of YLLs.

The discussion has also explored an area of incompatibility in AfL with summative assessment. It suggests that contextual factors, such as policy or stakeholders in the assessment process, can inhibit implementing AfL if their expectations do not align with the principles of AfL. The discussion has also explored the plausible sources of such a lack of alignment by suggesting that the formative and summative functions of

assessment operate within different paradigms: psychometric and pedagogical, respectively. Hence, issues of criteria and reliability may render the two functions incompatible.

Finally, the discussion has reflected on the insights about the construct of assessment in TEYL contexts that were provided by the current study. It has been argued that a situated and context specific approach to clarifying the construct of assessment might be most appropriate.

The discussion in this section has indicated how AfL may be linked to facilitating learning (e.g. through fostering positive affect or providing feedback). The following section continues that theme by discussing the findings from the data concerning the interactions that occurred during the use of AfL.

5.5 AfL and interactions

This section discusses the findings relating to interactions in TEYL classes. The findings indicate that the use of AfL techniques correlate positively with the number of T-1L and L-L[29] interactions. The findings also suggest that the holistic interaction patterns of conversations occurring during the use of AfL techniques are of the types that have been shown to facilitate learning, namely the collaborative and expert/notice type. Furthermore, the findings indicate that although the teachers did not align their language use with the perceived aims for conversations, the perceived conversation aims tended to be aligned with the learning objectives for the lesson. Finally, it is evident from the analysis that LREs often occurred when AfL was used but not in the managerial mode. A relationship between the conversation modes and the types of LREs was also reported; in the materials mode interlocutors engaged in lexical LREs while in the skill and system mode or the mode side sequences which involved that mode (see 3.3.3.2.2B), the interlocutors engaged in grammar LREs. Additionally, more modification of output was observed in the skills and systems mode.

This section is divided in two subsections. The first one discusses the findings relating to the holistic interaction patterns. The second one examines the conditions for learning during L-L and T-1L interactions.

---

[29] T-1L – teacher – one learner; L-L – learner-learner.

5.5.1 Holistic patterns

The findings relating to T-1L and L-L interactions occurring while AfL was being used indicate that these interactions were characterised by a medium to high level of mutuality. They could be classified as either collaborative (L-L) or expert/novice (T-1L) type. These types of interactions were reported to have pedagogical merits (e.g. Kim & McDonough, 2008; Watanabe & Swain, 2007). This body of research, however, comes from adolescent and adult settings and a similar research focus has not been explored in TEYL classrooms, with the exception of a study reported by Butler and Zeng (2014). Butler and Zeng (ibid.) found that learners aged 8/9 rarely engaged in collaborative dialogues and tended to use a formulaic language and fixed turn taking more frequently than 10/11 year olds. The findings of the current study do not provide insights into interactional features that would enable comparisons with Butler and Zeng's (ibid.) results. However, they do contribute to that research focus by suggesting that when AfL techniques were used, learners younger than those in the Butler and Zeng (ibid.) study entered collaborative dialogues. This suggests that AfL might have provided scaffolding that could benefit collaboration by children. However, the claims of the current study are limited due to the relatively small number of interactions in the dataset. It would be valuable for future research to explore this issue with a larger sample.

The identification of collaboration as a pattern of interaction that occurred during use of AfL has important pedagogical implications when considered in the context of studies investigating Long's Input Hypothesis and Krashen's Comprehensible Input Hypothesis. As discussed in Chapter 2, L2 research indicates that input can be made comprehensible through interactional modification of input that happens chiefly through the negotiation of meaning. A number of studies have indicated that comprehensible input (Ellis et al., 1994) and interactional modification of meaning (e.g. Mackey, 1999; Pica, 1994) can contribute to language learning. The potential contribution of negotiating meaning to language learning has been extensively researched in the field of SLA. The interactional modification of meaning has been shown to facilitate conditions and processes that are considered significant in learning a second language (Pica, 1994), having positive effects on second language production and comprehension (Gass & Varonis, 1994) and on vocabulary learning (Maleki & Pazhakh, 2012). Research to date also suggests that in the process of negotiating meaning, learners' attention may be on language form as well as on meaning as they attempt to convey messages clearly (Gass, 2013). Mackey (1999)

also argues that learners need to participate actively in such interactions, while (Ellis et al., 1994) demonstrated that learners who observed interactional modification also benefited These studies suggest that, by creating conditions in which teachers, learners and peers can modify input through the negotiation of meaning, teachers could facilitate language learning in their classrooms.

However, the majority of studies on the interactional modification of meaning have been carried out in adult settings; only a few have explored the negotiation of meaning between children and children and their teachers. Oliver (1998, 2000) reported a study focussing on identifying negotiation strategies that young learners employ depending on their age, language proficiency and the availability of a native speaker. The findings indicate that children can negotiate for meaning and do so most often when two low proficiency speakers are paired together. Less negotiation for meaning among children was observed in pairs where one of the speakers had high language proficiency, or both had high language proficiency, or one or both were native speakers (Oliver, 2002). These studies suggest that YLLs are capable of meeting the condition of negotiating for meaning that is crucial for interactional modification of input.

An important finding of the present study is the empirical evidence suggesting that the use of AfL could facilitate creating conditions conducive to learning in TEYL classes. Specifically, the findings of the quantitative analysis indicated that the use of AfL positively correlated with opportunities for one-to-one (L-L and T-1L) interactions and that these correlations were statistically significant. This is an important finding when considered in the context of the research discussed above as it indicates that AfL can help create conditions that enable interactional modification of meaning. This interpretation has implications for the practice of TEYL as it suggests that, by embedding AfL in classroom practice, teachers can, not only collect information about where learners are in their learning and use it for setting the next steps, but, more importantly, teachers can enact the learning function of assessment as discussed by Read-Dickins (2001), i.e. to enable learning through assessment. The discussion continues by exploring further insights provided by the findings of the current study into the relationships between AfL and creation of classroom conditions that can contribute to moving learning forward.

The discussion so far has indicated that holistic interaction patterns occurring most frequently during the use of AfL (collaborative and expert/novice) were of the type that

can contribute to creating conditions conducive to language learning and that the use of AfL contributes to creating conditions in which such interactions can occur. Additionally, it is important to note that research focussing on how collaborative and expert/novice patterns can contribute to learning (Kim & McDonough, 2008; Ohta, 1995; van Lier, 2014; Watanabe & Swain, 2007; Williams, 2001) point to the vital role of LREs in the process of language learning. The findings of the current study indicate that LREs occurred frequently during the use of AfL in the study. Hence, the occurrence of LREs seemed to be related to the modes of conversation in which the interlocutors operated and sometimes resulted in the modification of output. That area is discussed in the following section.

5.5.2 Modifications of output in different conversation modes

First, it is important to note that each analysed episode of classroom discourse was sourced from a broader context, i.e. a lesson, but which also constituted a set of micro contexts with their own pedagogical aims. The analysis of the alignment of the teachers' language use with the perceived pedagogical aims of each conversation indicated little congruence. That finding was consistent with what Walsh (2006) found. However, a different picture emerged when the extracts were considered in the context of a whole lesson. The findings of that analysis indicate that the perceived conversation aims were congruent with the pedagogical aims of each lesson, formally reported by teachers. Additionally, they suggest that the interactions in the majority of reported extracts offered opportunities for the learners to work towards meeting the pedagogical aims of the lesson. This suggests that the use of AfL could offer opportunities for directing learners' attention towards the learning objectives. While this may be considered a rather limited vision of learning, it certainly seems to be one that can benefit learning in the short term. This empirical finding of the current research prompts questions related to the curriculum and course design for YLLs. More specifically, if AfL can contribute to learners achieving their short term, i.e. lesson learning objectives, then, if over a longer period of time learning objectives offer opportunities for gradual development of skills, it seems plausible to suggest that the implementation of AfL in TEYL lessons could contribute to medium to long term benefits for learning. This interpretation indicates that the positive impact that AfL may have on learning is dependent on the lesson and curriculum contexts within which it is implemented.

The micro-contexts of each interaction were characterised by the different conversation modes (Walsh, 2006) within which they occurred. The findings obtained through applying the Variable Approach to the analysis of classroom interactions (Walsh, 2006) indicate that while AfL was being used, conversations happened within three classroom modes: managerial, materials and skills and systems. The findings also indicated two mode side sequences: materials-skills and systems-materials and skills and systems-classroom context (T only)- skills and systems. In a study conducted with learners aged 6-12 in an ESL context in Australia, Oliver and Mackey (2003) analysed similar modes, which they referred to as 'contexts'. They identified the following contexts: management, communication, content and explicit language-focused exchanges. Table 5.5 compares the current study modes with the Walsh (ibid.) modes and the Oliver and Mackey (ibid.) contexts. Walsh's (ibid.) terminology is adopted here.

Table 5.5: Comparison of Walsh (2006) modes and Oliver and Mackey (2003) contexts of classroom interactions with the current study

Walsh (2006) modes	Managerial	Classroom context	Skills and systems	Materials
Oliver and Mackey (2003) contexts	Management	Communication	Explicit language-focused	Content
The current study modes	Management	-	Skills and systems	Materials

The modes in the second and third columns are defined very similarly in the Walsh and the Oliver and Mackey (ibid.) studies. In the fourth column, the explicit language-focused mode is defined as 'discussions and instruction about English phonology, morphology, the lexicon (including the meaning of words), and syntax' (Oliver & Mackey, 2003, p. 524), whereas Walsh (2006) in addition to all these characteristics includes 'practice of sub-skills' (p. 66) in the definition of the skills and systems mode. Despite that difference, these two modes are considered here to be largely similar. The definitions of the remaining two differ, reflecting the different educational contexts in which they were developed: TEFL (Walsh, ibid) and ESL (Oliver & Mackey, ibid.). The materials mode is concerned with practising language with relation to a piece of material, e.g. a text in a course book, while the content mode focuses on 'imparting knowledge or eliciting information from the learners about a curriculum content' (Oliver & Mackay, 2003, p. 523). Considering both classifications is useful as each of them shares a different

characteristic with the context of the current study; Oliver and Mackey (ibid.) developed their modes with children, while Walsh (ibid.) developed his in a TEFL context and subsequently also used them for analysing children's interactions. Additionally, Walsh's (2006) modes offered opportunities for more in-depth analysis as they included mode side sequences. The interactions recorded in the current study (Section 4.4.3.2) demonstrate three of the four modes proposed by Walsh (ibid.). The classroom context mode was not recorded. This was interpreted to mean that, during the use of AfL, the conversations did not tend to allow for flexibility in topics related to the learners' interests and experiences. Instead, they incorporated a mode that included management of the classroom activities and conversations related to the target language (skills and systems mode) or to the task in hand (materials mode).

The analysis of interactions recorded during the use of AfL, indicate that negative feedback initiated the negotiation of meaning, which in four out of seven LREs resulted in the modification of output by the learners (Table 4.16, p.203). It seems important to consider the opportunities for learners to modify output. The discussion in Chapter 2 indicates that, by producing output and modifying it following negative feedback, learners may 'process language more deeply – with more mental effort' (Swain, 2000, p. 99). The findings of the present study, when compared with the outcomes of Oliver and Mackey's (2003) study, provide some interesting interpretations. Oliver and Mackey (ibid.) found that teachers were most likely to provide feedback in explicit language-focused and content modes, while learners were most likely to modify their output following feedback only in the explicit language-focused mode. The findings of the current study suggest that interactions occurring in the skills and systems mode, and perhaps in mode side sequences including the SS mode, could facilitate modifications of output, similarly to the interactions in explicit language focus mode, as reported by Oliver and Mackey (ibid.). Importantly, all T-1L interactions during LREs resulted in modified output. This is especially interesting when compared with the Mackey et al. (2003) claim that 'teachers often provide information and answers, and thus fewer opportunities for modifications' (p. 58). The findings of the current study suggest that when using AfL the teachers were providing opportunities for modifications and the learners used those opportunities to modify output, and that this seemed to occur more often in the skills and systems mode. However, LREs did not occur in the managerial mode. This has important pedagogical implications as it suggests that by employing AfL techniques, teachers can

change the nature of interactions which they have with their learners to facilitate language processing through encouraging modification of output.

However, it should be noted that the current study did not aim to analyse the amount of negative feedback in interactions occurring during AfL. Instead, it set out to describe the largely under-researched area of the use of AfL in TEYL classrooms. Hence, the collected data did not allow for direct comparisons with the Oliver and Mackey (ibid.) study. This focus would certainly constitute a worthwhile path for future research.

It is important to note that the claims of the present study are also limited by the nature of the dataset as it only includes examples of interactions and not the whole set of all interactions that occurred in observed lessons. The examples were sourced from the video recorded lessons. What is perhaps more valuable is that the above discussion highlights a gap in TEYL research and possible paths of inquiry that would be highly valuable for future studies to explore.

5.5.3 Concluding remarks for the section

This section has interpreted and discussed the implication of findings about interactions occurring during the use of AfL in TEYL classrooms. It demonstrated that the use of AfL could be related to creating classroom conditions that may facilitate opportunities for the interactional modification of input through negotiation. This implies that teachers would adopt the roles of facilitators of learning and not of sources of knowledge. Moreover, it suggests that the use of AfL could be linked to teachers facilitating modification of output through LREs which are often initiated with negative feedback, instead of teachers 'providing answers'. Importantly, however, the discussion has stressed that classroom interactions ought to be considered according to the modes in which they occur. This is important because the provision of feedback and its use had been shown to be related to modes of conversation (Mackey & Oliver, 2003). These findings are important as they demonstrate how the learning function of assessment proposed by Rea-Dickins (2001) can be enacted in practice in TEYL classrooms. It is also crucial to comment that both types of modification, of input and of output, are believed to support FLL, with some emerging evidence that this is also true in TEYL contexts (Mackey & Silver, 2005; Oliver, 2002; Oliver & Mackey, 2003).

5.6 Summary of the chapter

This chapter has discussed the insights into teaching, learning and assessment in a TEYL context. These insights have been derived through the interpretations of the findings about teachers' understanding and use of AfL and its relationship with interactions in TEYL classes.

The discussion of the findings about the teachers' understanding of AfL in the TEYL context, in Section 5.3, indicates that this understanding is largely consistent with the Black and Wiliam (2009) theoretical framework and the conceptualisation of genuine AfL as proposed by Swaffield (2011). The discussion has suggested a number of characteristics of AfL, which, in regard to teachers' understanding, differed from the two theoretical frameworks and hence might suggest that this understanding is specific to TEYL classes. These include opportunities for fostering affective factors, the need for purposeful planning for the implementation of AfL and the lack of a role for peers in sharing learning objectives and success criteria.

The discussion in Section 5.4, indicates that teachers implement AfL differently depending on the age of their learners and that, in the older age group, implementation changes especially in terms of technique type and feedback provision procedures. These differences could be attributed to the cognitive, metacognitive and literacy development of learners. Importantly, the discussion has also explored how AfL could facilitate learning in TEYL classrooms. It suggests that AfL techniques could be used to scaffold production. Furthermore, the discussion has explored the relationship between the use of AfL and affective factors, concluding that the use of AfL has the potential to foster learners' feelings of success and in doing so could contribute to motivating YLLs. It has also been argued that the motivation to learn FL is important in TEYL classrooms as it is often stated as an explicit outcome in TEYL programmes and has been shown to facilitate the learning of productive skills.

The discussion of the findings about the relationship between AfL and interactions, presented in Section 5.5, has proposed that the increased number of L-L and T-1L interactions associated with the use of AfL can create conditions that are conducive to FLL. Specifically, it has been argued that the use of AfL provides opportunities for interactional modifications of input and output, which have been shown by research to facilitate FLL. This was evident in the higher number of L-L and T-1L interactions

positively and significantly correlated with the use of AfL. Additionally, the holistic interaction patterns occurring during those interactions were of the kinds that had previously been shown to facilitate learning (e.g. Butler & Zeng, 2014; Swain, 2000). The third source of empirical evidence for the relationship between the use of AfL and creating conditions conducive to learning is the empirical finding that interactions that took place in the materials and the skills and systems modes offered opportunities for interlocutors to provide negative feedback through lexical (in the materials mode) and grammar (in the skills and systems mode) LREs. As a result interlocutors could modify output and did so more often in the skills and systems mode.

Throughout the discussion, a number of implications for future research have been identified. These are synthesised in the following final chapter, where the limitations of the current study are also discussed.

Chapter 6: Conclusions, implications and limitations

Claims about the impact of a small scale study like the current one should be considered carefully. Nevertheless, it seems valuable for the final chapter to draw conclusions from the discussion of the findings in Chapter 5. In order to do so in an informed manner, it seems necessary to first revisit the aims of the current study (6.1) to examine whether they have been met and to confirm that the study contributes to the body of literature reviewed in Chapter 2. Perhaps the most important contribution of an exploratory and descriptive study, like the current one, is its potential to indicate directions for future research. Hence, Section 6.2 proposes directions for advancing the development of the theory of AfL and for researching implementation of AfL as a means of moving learning forward. It also seems important that a study set within the classroom offers insights that could underpin teachers' practice. Hence, the next section (6.3) examines the pedagogical implications of the findings in order to consider how they could inform the implementation of AfL and the understanding of its impact on classroom practice and language achievement. Finally, Section 6.4 discusses the limitations of the findings.

6.1 Revisiting the aims of the current study

The aims of the current study (1.4), inspired by the researcher's professional experience as a teacher of YLLs and shaped by the outcomes of the literature review (Chapter 2), were translated into the three research questions (2.4). Chapter 2 indicates three gaps in the research that supports the timeliness of and the need for the current study.

First, the number of young learners who are taught a FL has grown globally during the past decades resulting in a necessity to develop assessment methods appropriate for the needs of children. Importantly, such assessment should be informed by what is known about learners' cognitive development and the ways in which they learn languages. Additionally, in order to demonstrate achievement, assessment methods used with YLLs should be organised in small steps, sensitive to the varying rates of progress and the needs of individual learners. These issues signify a clear need for research into assessment in TEYL contexts.

The second area that underpins the relevance of the current study to the field of TEYL is the attention that AfL has attracted in other educational contexts. Three areas have received the significant attention of academics, governments and practitioners alike. The

first one encompasses the discussions of the theoretical framework(s) of formative assessment and/or AfL, their relationship(s) to summative assessment and/or AoL as well as the associated terminology. The overall picture that has emerged from these discussions indicates that this is a complex area and as yet there is no consistency across the research community with regards to these issues. Such inconsistency in the understanding and the language of AfL is seen as one of the reasons for many types of practical implementation that are subsumed under the umbrella terms of AfL and/or FA. It also has been argued by some researchers that a number of such types of implementation did not actually have a formative function, despite being called AfL and/or FA. In the light of these discussions of the theoretical conceptualisations of AfL and/or FA, this study addresses this lacuna in the TEYL context.

Another emergent area focussing on AfL in other educational contexts concerns models of the practical implementation of AfL. Empirical research in this area suggests that factors such as educational policy, teaching methods and teachers' experience could impact on the implementation of AfL. Some insights have revealed the ways in which teachers enact the formative function of assessment by pointing to the existence of formatively oriented behaviours observed in lessons and the limited time being devoted to such practices by teachers. Hence, as other researchers have requested, it seems important to explore what teachers do when they implement AfL as well as how and when they do so.

The final area of relevance to a TEYL context is related to the claims of efficacy of AfL in advancing learning. In recent years, a number of studies have indicated that embedding AfL could have a positive impact on achievement in Language Arts, Science or Mathematics. Additionally, Black and Wiliam (1998) conducted a meta-analysis of 250 articles and concluded that their work 'shows conclusively that formative assessment does improve learning' (p. 61). However, such assertions have been called into question mostly due to the perceived inadequacy of the criteria applied for including studies in the meta-analysis. In effect other researchers have called for more empirical evidence from a greater variety of educational contexts to validate the efficacy claims (e.g. Dunn & Mulvenon, 2009). As yet, I have been unable to identify any published study that would link the use of AfL empirically with achievement in a TEYL context. To examine whether AfL could contribute to moving learning forward, this study has explored its observable impact on interactions in the classroom.

Having revisited the aims of the current study, the chapter continues by considering how the findings of the current study could guide future research into AfL.

6.2 Implications for future research

As indicated in the introduction to this chapter, this section consolidates the suggestions for future inquiry made throughout Chapters 4 and 5.

An important point made in Chapter 5 is the indication that AfL could scaffold children's work while they are completing speaking and writing tasks that seem to require more on-line processing. The discussion suggests that one area of focus for future research would be exploring if and how AfL techniques can contribute to the effective scaffolding of learning of different language skills and in various age groups. This could provide insights into the relationship between the use of AfL and improved performance. This path of inquiry might offer a fruitful area that could provide empirical evidence for how AfL could facilitate learning. For example, it would be of value to conduct comparative studies of performance between learners who use AfL and those who do not. The results of such studies could provide insights into the relationship between AfL and improved performance, thus providing empirical evidence for the efficacy of AfL.

Furthermore, the findings in the current study suggest that diversity in technique type might be an important consideration in implementing AfL in TEYL classrooms. Some teachers in the current study indicated that such diversity is needed to ensure learners' interest. At the same time, studies into the assessment of YLLs have highlighted the importance of assessment instruments being interesting and engaging for young children. Hence, it would be useful for future research to explore this area by investigating if and how diversity in technique type would correlate with high on-task engagement. However, it should also be noted that designing such studies poses methodological challenges as it is difficult to quantify level of engagement.

Chapter 5 also proposed that a future focus of research could be to explore whether there might be a relationship between the use of self-assessment, an integral part of AfL, and self-regulation. A fruitful line for inquiry would be to examine if the development of metacognitive awareness and control could be supported through the use of AfL techniques. This would be a valuable research focus as the control of metacognitive strategies is an important component of the growth path of self-regulated learning.

Enabling self-regulated learning of this type has pedagogical gains in that there are implications for the roles that teachers and learners adopt in lessons. The findings of the present study suggest that teachers' roles might change from those of providers of knowledge to facilitators of learning. To understand the impact that the use of AfL could have on the roles of teachers and learners in classes more systematic research is needed.

The findings of the current study suggest that interactions taking place during the use of AfL could have pedagogical benefits for FLL. This points to yet another area of interest for future research. In Chapter 5 it was argued that AfL might facilitate conditions beneficial for learning through encouraging collaborative dialogues, providing opportunities for negative feedback and the occurrence of LREs. Furthermore, it was indicated that in the skills and systems mode, LREs result in modifications of output. It would be interesting for future studies to explore these areas with larger samples of interactional data. It would also be interesting to compare systematically how these gains might differ among various age groups. In order to make valid claims of the existence of such relationships, the design of a study needs to allow lengthier conversation between learners (similar to Swain, 2000), to analyse the occurrence of LREs in various interaction modes (following Oliver & Mackey, 2003) and to consider those LREs, referring to the task types that children complete when the interactions are recorded (similar to Butler & Zeng, 2014). By designing studies that attempt such analyses, future research could explore the actual *learning* function of assessment.

Another area worthy of future research in TEYL contexts is related to giving explicit feedback and the impact that such feedback may have on learning. The findings of the current study indicate that AfL was used for providing explicit feedback on performance to young learners. One aspect is that there may be a relationship between learners' ages and the most appropriate feedback provider; the younger learners received feedback from each other, while the older learners, in addition to receiving feedback from each other, acted as feedback providers for themselves. Another aspect relates to the roles that the feedback providers adopted through feedback and how these could support learning in TEYL contexts. For example, the significant advantage of positive reinforcement through feedback has been suggested. This is different from the findings of studies indicating that teachers of adolescents most often adopt the role of a Provider (Furneaux et al., 2007). Rather, it foregrounds the connection between learning and affective factors in TEYL

classes. It would be informative to replicate those studies examining written feedback given by teachers of adolescents, by applying their design to younger age groups.

Finally, a number of strands of evidence suggest a relationship between using AfL and building positive affective profiles of YLLs. The overall indication is that using AfL can be motivational for YLLs mostly through positive reinforcement in the feedback and through providing opportunities for building up a sense of success. By highlighting children's success in language learning through AfL techniques, classroom practice could enable a building up of positive self-concepts. This would be an important focus for future inquiry, especially in the light of research that has highlighted the decreasing level of motivation of children in the long term.

The current section paints a complex picture since many variables are at play when AfL in TEYL contexts is considered. These include cognitive, socio-cultural and affective factors, which can interact with the implementation of AfL. Since research into AfL in TEYL contexts is in its infancy, there are many unexplored areas and a great deal of research is still needed before valid claims of efficacy and models of implementation, grounded in a robust theoretical framework, can be made.

6.3 Pedagogical Implications

This section synthesizes the pedagogical implications that the findings of the current study may have. These implications should be considered together with the limitations of the study.

The findings of the current study provide insights in how the learning function of assessment (Rea Dickins, 2001) can be enacted in TEYL classes. Most significantly, they indicate that teachers can use AfL techniques to support ongoing learning by setting objectives and expectations, monitoring learning and checking achievement. This understanding of assessment as a continuous process which occurs alongside teaching and learning has implications for how teachers plan and deliver their lessons. It seems that to enact the learning function of assessment, teachers should align their assessment foci with the pedagogical objectives of lessons. Furthermore, they should ensure that young learners understand what is expected of their performance. Following that, teachers should provide opportunities for (self-)monitoring of learners' work. Building on studies which suggest that setting objectives, monitoring work and providing feedback

on achievement can be conducive to language learning (e.g. Edelenbos & Vinje, 2000), the present study proposes that TEYL teachers could create such conditions in their lessons by implementing AfL techniques.

Secondly, the findings about implementing AfL in TEYL classrooms highlight the importance of employing a diversity of technique types, presumably to encourage interest. This implies that teachers should be supported in implementing AfL through the provision of development opportunities enabling them to learn how to use different technique types. These findings also suggest the importance of ensuring that language assessment techniques are interesting and engaging for children, as shown in the findings of Hasselgreen (2000).

Thirdly, the findings of the current study indicate that AfL techniques can be used for giving explicit feedback. In contexts where learners are at low levels of language proficiency, providing informative feedback to learners may be inherently difficult. Hence, an indication that AfL can be used to facilitate that process promises that it can be a useful tool for teachers to use to improve their practice in this areas. However, it should also be noted that the findings highlight a number of differences in feedback provision practices between the two age groups in the current study. Most significantly, there seems to be more evidence that self-assessment was used in the groups of 10-11 year olds, whereas the younger children relied predominantly on feedback received from teachers and peers. These findings have useful pedagogical implications as they suggest that teachers should, at least, be aware that younger children may not be able to self-evaluate. Building on studies which suggest that young learners may be trained to self-assess accurately (e.g. Butler & Lee, 2010), the findings of the current study propose that teachers could use AfL techniques to provide opportunities for learners to practise how to self–assess. However, the findings also indicate that this may be challenging in classes of younger learners.

Fourthly, the discussion in Section 5.3.3.1 suggests that, by employing AfL systematically, teachers may be able to create conditions conducive to better language performance. Most importantly, the discussion indicates that, as children become more familiar with AfL techniques, these can contribute to increasing the degree of familiarity and structure of tasks used in lessons. Other research (e.g. Pinter, 2007) has suggested that children may be able to produce better quality output when they are familiar with the

task type which they are working on. Hence, it seems possible to suggest that, by implementing AfL techniques systematically, teachers could contribute to improving their learners' language performance by introducing a greater level of familiarity with structure within tasks.

Another finding of the current study that has implications for the practice of TEYL is that AfL techniques can be used to motivate learners. It was indicated in the teachers' understanding that they could adopt the role of Supporters through feedback practices, hence helping children to perceive themselves as successful learners, which, in turn, could contribute to building up a positive self-concept. Consequently, the research on the assessment of YLLs indicating a need for positive feedback to be provided to learners (Hasselgreen, 2005) is addressed in this study since AfL techniques could be perceived as a vehicle for such feedback.

The pedagogical implications of the findings of the current study could inform implementation of AfL in TEYL classrooms. While they seem to share many of the characteristics of the guidance produced for mainstream teachers in England (Appendix 1), they extend that guidance by emphasising that learners' age may be an important consideration in developing models for implementation of AfL. However, a larger empirical database is required before valid models of implementation of AfL in TEYL contexts could be developed.

6.4 Limitations

A number of limitations were discussed in Chapter 3 as they were related to the study design and the sample. The current section aims to draw them together and consider their effect in the context of the whole study.

First, it is acknowledged that, in a mixed-method design with a larger qualitative than quantitative component, the researcher was an important research tool (Cohen et al., 2007). This means that, in the process of coding, the researcher interpreted the information shared by the teachers in the interviews and the focus group. Furthermore, the researcher brought to the study her own professional experience as a primary teacher, a TEFL teacher and an academic researcher. Effectively, her understanding of AfL may have differed from that of the participant teachers. Hence, measures were employed to ensure that the findings of the current study could be substantiated by the data. The

following steps were taken to ensure the validity of interpretation: codes were developed in the pilot stage; an inter-rater was invited to participate, which meant that a reliability coefficient higher than 0.9 could be obtained; and, finally, the findings from analysing the interview data were validated through the focus group. Another area in which the researcher's interpretation could have affected the results of the study was in recording field notes of the lessons observations. To ensure reliability of the field notes, the lessons were video recorded to enable reviewing. It is believed that by implementing these measures the researcher's role as the main tool of the qualitative procedures did not affect the quality of the findings and that, in effect, the findings are a worthy representation of the phenomena studied.

The second area of consideration is the classroom as a complex research setting. As discussed in Chapter 3, there can be many variables interacting with one another at any one time during a lesson. Hence, the study did not aim to, and did not make any claims about, demonstrating causation. The findings for RQ3, which looked at the observable impact of AfL on interactions, reported a number of correlations. They indicated the co-occurrence of AfL and the high numbers of L-L and T-1L. The increased number of those types of interactions is relative to the whole sample of lessons. More specifically, this indicates that the number of interactions was higher compared to those in other lessons in the sample. Hence, this finding seems sensitive to the context of the current study. This may signify that, if replicated in a different context, the study design may return somewhat different results in terms of the correlation between the number of interactions and the use of AfL. Multiple factors could impact on this including teaching methodology, the way in which AfL is implemented, and contextual factors impacting on assessment as discussed elsewhere in the thesis. This may result in limiting the claims that can be made about the transferability of the findings of the current study. The thesis has aimed to acknowledge that limitation by fully reporting on the context of the study. To that end, the characteristics of the participants and the school were described in detail in Chapter 3 and the information about the teaching context was included in Chapter 4. It is believed that in doing this the thesis has indicated the nature of the contexts to which the findings may be applicable.

Thirdly, a study that included twenty-eight lesson observations conducted by eight teachers and their records of work is inevitably limited in scope, as indicated at the beginning of the current chapter. Hence, no claims are made about the generalisability of

the findings to other populations of YLLs. However, it is believed that the findings of the current study add an interesting voice to the discussions about assessment in TEYL classrooms. Most significantly, the present study extends the body of research reviewed in Chapter 2 by: reporting teachers' understanding of AfL, specific to a TEYL context; describing the practice of implementing AfL; and empirically indicating a number of areas where using AfL could be linked with language achievement.

Finally, it is important to recall that this study was a project undertaken by the researcher in order to learn and subsequently demonstrate the research skills required for the award of an academic degree. Hence, it was a developmental task for the researcher. Inevitably, it seems valuable at this stage to reflect on how this study could have been approached differently. One area seems especially valuable to consider. With the benefit of hindsight, I would consider recording classroom conversations by using a different technology, if possible, in order to capture the majority of conversations that occurred in all lessons. Then by comparing the amount of negative feedback, LREs and modifications of output in lessons in which AFL was used with those in which it was not, it would be possible to provide a more informed insight into the relationship between the use of AfL and the conditions for learning through interaction. However, using a greater amount of audio or video technology might impact on what actually occurs in the lessons: hence changing the nature of the resulting findings somewhat.

6.5 Final conclusions

This thesis explores teachers' understanding and classroom implementation of AfL as well as the observable impact of AfL on learning. It concludes that teachers' understanding of what AfL is in TEYL classrooms is largely consistent with the influential frameworks proposed by Black and Wiliam (2009) and Swaffield (2011) with the exception of the peer-role in establishing learning goals and criteria for success but with the addition of the evaluation of learners' confidence concerning their own progress. Secondly, it shows that AfL could be implemented with a variety of techniques for sharing learning objectives, criteria for success and feedback as well as, to a lesser degree, for evaluating learners' confidence and, in younger groups, also for giving and clarifying instructions. Finally, the thesis concludes that AfL could benefit learning by providing scaffolding and facilitating interactions between learners, peers and teachers that could allow for the occurrence of LREs, increasing comprehensibility by modifications of input

and encouraging internal processing by facilitating modifications of output. Furthermore, it is noted that when researching in TEYL, it is important to draw on the insights of cognitive psychology.

References

Alcón, E. (2007). Incidental focus on form, noticing and vocabulary learning in the EFL classroom. *International Journal of English Studies*, *7*(2), 41-60.

Anderson, P. (2005). *Cognitive psychology and its implication.* London: Macmillan.

ARG (Assessment Reform Group, 2002). *Assessment for Learning: 10 Principles. Research-based principles of assessment for learning to guide classroom practice* Retrieved June 8, 2012, from http://webarchive.nationalarchives.gov.uk/20101021152907/http:/www.ttrb.ac.uk/ViewArticle2.aspx?ContentId=15313.

Bachman, L. (2007). What is the construct? The dialectic of abilities and contexts in defining constructs in language assessment. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. Turner, and C. Doe (Eds.), *Language Testing Reconsidered* (pp.41-71). Ottawa, Canada: University of Ottawa Press.

Bailey, A. L. (2005). Cambridge Young Learners English (YLE) Tests. *Language Testing*, *22*(2), 242-252.

BERA, (2011). *Revised ethical guidelines for educational research.* Retrieved July 5, 2011, from https://www.bera.ac.uk/researchers-resources/publications/ethical-guidelines-for-educational-research-2011

Berk, L. (2000). *Child Development.* Needham Heights, MAA: Pearson Education Company.

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

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge: Cambridge University Press.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.). *Second language acquisition and the critical period hypothesis* (pp.161-181). London: Routledge.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, *44*(2), 235-249.

Black, T.R. (2005). *Doing Qualitative Research in the social sciences: an integrated approach to research design, measurement and statistics.* London: Sage.

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

Black, P. J. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice, 5*(1), 7-73.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability 21*(1), 5-31.

Bloom, B. S. (1984). The search for methods of instruction as effective as one-to-one tutoring. *Educational Leadership*, *41*(8), 4-17.

Boekaerts, M., & Cascallar, E. (2006). How far have we moved toward the integration of theory and practice in self-regulation? *Educational Psychology Review*, *18*(3), 199-210.

Bogdan, R. C., & Biklen, S. K.(1992). *Qualitative research for education (2nd ed.).* Boston: Allyn and Bacon.

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, *36*(2), 81-109.

Borkowski, J. G. (1985). Signs of intelligence: Strategy generalization and metacognition. In S.R. Yussen (Ed.), *The growth of reflection in children* (pp. 105-144). London: Adademic Press.

Braunek, (2013). *Powszechność nauczania językow w roku szkolnym 2011/12: raport. [The teaching of languages in academic year 2011/12: a report].* Ośrodek Rozwoju Edukacji [The National Centre for the Development of Education]. Warsaw: ORE.

Broadfoot, P. M., Daugherty, R., Gardner, J., Gipps, C. V., Harlen, W., James, M., & Stobart, G. (1999). *Assessment for learning: Beyond the black box.* Cambridge: University of Cambridge School of Education.

Burstall, C., Jamieson, M., Cohen, N. F. E. R., & Hargreaves, M. (1974). *Primary French in the balance*. Slough: NFER.

Butler, R. (1988). Enhancing and undermining intrinsic motivation: The effects of task-involving and ego-involving evaluation on interest and performance. *British Journal of Educational Psychology*, *58*(1), 1-14.

Butler, Y. G. (2009). Issues in the assessment and evaluation of English language education at the elementary school level: Implications for policies in South Korea, Taiwan, and Japan. *Journal of Asia TEFL*, *6*(2), 1-31.

Butler, Y. G., & Lee, J. (2006). On-task versus off-task self-assessments among Korean elementary school students studying English. *The Modern Language Journal*, *90*(4), 506-518.

Butler, Y. G., & Lee, J. (2010). The effects of self-assessment among young learners of English. *Language Testing*, *27*(1), 5-31.

Butler, Y. G., & Zeng, W. (2014). Young foreign language learners' interactions during task-based paired assessments. *Language Assessment Quarterly*, *11*(1), 45-75.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136-146). London: Macmillan Education.

Cable, C., Driscoll, P., Mitchell, R., Sing, S., Cremin, T., Earl, J., Eyres, I., Holmes , B., Martin, C. and Heins, B. (2010). *Languages learning at Key Stage 2: a longitudinal study.* United Kingdom of Great Britain and Northern Ireland, GB, Department for Children, Schools and Families (DCSF Research Reports DCSF-RR198). Retrieved March 15, 2011 from http://eprints.soton.ac.uk/143157/1/DCSF-RR198.pdf

Cameron, L. (2003). Challenges for ELT from the expansion in teaching children. *ELT Journal*, *57*(2), 105-112.

Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice*, *12*(1), 39-54.

Carless, D., Joughin, G., & Liu, N. F. (2006). *How assessment supports learning: Learning-oriented assessment in action*. Hong Kong: Hong Kong University Press.

Carr, M. (2008). Can assessment unlock and open the doors to resourcefulness and agency. In S. Swaffiedl (Ed.), *Unlocking assessment: Understanding for reflection and application* (pp.36-54). New York: Routledge.

Case, R. (1972). Validation of a neo-Piagetian mental capacity construct. *Journal of Experimental Child Psychology*, *14*(2), 287-302.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*(3), 386-404.

CBOS (Centrum Badania Opinii Spolecznej, 2011). *Wydatki rodzicow na edukacje dzieci w roku szkolnym 2011/12* [Parents spending on children's education in the

academic year 2011/12, Public Opinion Research Centre] Retrieved April 7, 2012 from http://www.cbos.pl/SPISKOM.POL/2011/K_134_11.PDF

Ceci, S. J., & Roazzi, A. (1994). The effects of context on cognition: Postcards from Brazil. *Mind in Context: Interactionist Perspectives on Human Intelligence*, *12*, 74-101.

Cenoz, J. (2003). The influence of age on the acquisition of English: General proficiency, attitudes and code mixing. In M. García Mayo & M. García Lecumberri (Eds.). *Age and the acquisition of English as a FL* (pp. 77-93). Clevedon: Multilingual Matters.

Chamot, A., & El-Dinary, P. B. (1999). Children's learning strategies in language immersion classrooms. *The Modern Language Journal*, *83*(3), 319-338.

Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing. *Journal of Second Language Writing*, *12*(3), 267-296.

Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, *21*(3), 360-389.

Clarke, S. (2005). *Formative assessment in action: Weaving the elements together*. London: Hodder Murray.

Cohen, L. Manion, L. & Morrison, K. (2007). *Research methods in Education* (6th ed). Oxon: Routledge.

Colby-Kelly, C., & Turner, C. E. (2007). AFL Research in the L2 classroom and evidence of usefulness: Taking formative assessment to the next level. *Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, *64*(1), 9-37.

Conrad, S. M. and Goldstein, L.M. (1999). ESL student revisions after teacher-written comments: Text, contexts and individuals. *Journal of Second Language Writing, 8*(2): 147–180.

Copland, F., & Garton, S. (2014). Key themes and future directions in teaching English to young learners: Introduction to the Special Issue. *ELT Journal*, *68*(3), 223-230.

Council of Europe (1996). *Common European framework of reference for the learning and teaching of languages*. Strasbourg: Council of Europe Education Committee.

Cresswell, J.W. (2007). *Qualitative inquiry and research design: choosing among five approaches.* London: Sage Publications.

Cummins, J. (1986). Language proficiency and academic achievement. In J. Cummins, M. Swain (Eds.), *Bilingualism in education: aspects of theory, research and practice* (pp. 138–161). London: Longman,

D'Andrade, R. (1991). The identification of schemas in naturalistic data. In Horowitz, Mardi Jon (Ed), *Person schemas and maladaptive interpersonal patterns. The John D. and Catherine T. MacArthur Foundation series on mental health and development* (pp. 279-301). Chicago, IL, US: University of Chicago Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 450-466.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, *3*(4), 422-433.

Dann, R. (2002). *Promoting assessment as learning: Improving the learning process*. London: Routledge.

Davison, C. & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly, 43*(3), 393-415.

Dawson, C. (2007) *A practical guide to research methods*. Oxford: Hot To Books Ltd.

Debyser, D., & Tagliante, C. (2001). *Mon premier portfolio*. Saint Amand: Didier.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499-533.

Denzin, N. K. (1996). The epistemological crisis in the human disciplines: Letting the old do the work of the new. In R. Jessor (Ed.), *Ethnography and human development: Context and meaning in social inquiry* (pp. 127-151). Chicago: University of Chicago Press.

Dimroth, C. (2008). Age effects on the process of L2 acquisition? Evidence from the acquisition of negation and finiteness in L2 German. *Language Learning*, *58*(1), 117-150.

Dixon, H. R., Hawe, E., & Parr, J. (2011). Enacting Assessment for Learning: the beliefs practice nexus. *Assessment in Education: Principles, Policy & Practice*, *18*(4), 365-379.

Donaldson, M. (1978). *Children's minds*. London: Fontana Press.

Dörnyei, Z. (2001). *Teaching and researching:Motivation*. London: Longman.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In J. Wiliams (Ed.), *Focus on form in classroom second language acquisition* (pp. 114-138). Stuttgart: Ernst Klett Sprachen.

Drummond, M. J. (2003) *Assessing children's learning*. London: David Fulton.

Dunn, K. E., & Mulvenon, S. W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation*, *14*(7), 1-11.

Dweck, C. (2000). *Self theories: Their role in motivation, personality, and development.* Philadelphia, PA: Psychology Press.

Earl, L. M. (2012). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oakes: Corwin Press.

Edelenbos, P., Johnstone, R., & Kubanek, A. (2006). The main pedagogical principles underlying the teaching of languages to very young learners. *Languages for the children of Europe: Published research, good practice and main principles. European Commission Report.* Retrieved June 2, 2011 from http. *ec. europa, eu/education/policies/lang/doc/youngsum_en. pdf*.

Edelenbos, P., & Kubanek-German, A. (2004). Teacher assessment: The concept of 'diagnostic competence'. *Language Testing*, *21*(3), 259-283.

Edelenbos, P., & Vinjé, M. P. (2000). The Assessment of a Foreign Language at the End of primary (elementary) education. *Language Testing*, *17*(2), 144-162.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research*, *4*(3), 193-220.

Ellis, G. (2014). 'Young learners': clarifying our terms. *ELT Journal*, *68*(1), 75-78.

Ellis, R., & Heimbach, R. (1997). Bugs and birds: Children's acquisition of second language vocabulary through interaction. *System*, *25*(2), 247-259.

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, *44*(3), 449-491.

Enever, J. (2011). *ELLiE. Early language learning in Europe.* London: British Council.

Eurostat (2013). *Educational attainment level and outcomes of education. Implementation of ISCED 2011.* Retrieved August 10, 2013, from http://ec.europa.eu/eurostat/data/database

EURIDICE (2014). *Organisation of school time in Europe: Primary and Secondary general education 2014/15 school year. Eurydice – Facts and Figures.* Brussels: Eurydice. Retrieved August 20, 2014 from http://eacea.ec.europa.eu/education/eurydice/documents/facts_and_figures/school_calendar_EN.pdf

Ferris, D. R., Pezone, S., Tade, C. R., & Tinti, S. (1997). Teacher commentary on student writing: Descriptions & implications. *Journal of Second Language Writing*, *6*(2), 155-182.

Flavell, J. H. (1992). Cognitive development: Past, present, and future. *Developmental Psychology*, *28*(6), 998-1005.

Flavell, J. H., Beach, D. R., & Chinsky, J. M. (1966). Spontaneous verbal rehearsal in a memory task as a function of age. *Child Development,37*(2), 283-299.

Flavell, J. H., Green, F. L., & Flavell, E. R. (2000). Development of children's awareness of their own thoughts. *Journal of Cognition and Development*, *1*(1), 97-112.

Flavell, J.H., Miller, P.H., & Miller, S. A. (1993). *Cognitive development* (3rd ed.)*.* New Jersey: Prentice Hall.

Flege, J. E., Munro, M. J., & MacKay, I. R. (1995). Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America*, *97*(5), 3125-3134.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*(1), 78-104.

Furneaux, C., Paran, A., & Fairfax, B. (2007). Teacher stance as reflected in feedback on student writing: An empirical study of secondary school teachers in five countries. *IRAL-International Review of Applied Linguistics in Language Teaching*, *45*(1), 69-94.

Garcia Mayo, M. D. P., & Garcia Lecumberri, M. L. (Eds.). (2003). *Age and the acquisition of English as a foreign language*. Clevendon: Multilingual Matters.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. London: Edward Arnold.

Gardner, R. C. (2010). *Motivation and second language acquisition: the socio-educational model*. New York: Peter Lang.

Gardner, R. C., & MacIntyre, P. D. (1993). A student's contributions to second-language learning. Part II: Affective variables. *Language Teaching*, *26*(01), 1-11.

Gass, S. M. (2013). *Input interaction and the second language learner*. New York: Routledge.

Gass, S. M., & Varonis, E. M. (1994). Input, interaction, and second language production. *Studies in Second Language Acquisition*, *16*(03), 283-302.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: a critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*(3), 259-272.

Gattullo, F. (2000). Formative assessment in ELT primary (elementary) classrooms: an Italian case study. *Language Testing*, *17*(2), 278-288.

Ghatala, E. S. (1986). Strategy-monitoring training enables young learners to select effective strategies. *Educational Psychologist*, *21*(1-2), 43-54.

Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. New York: Psychology Press.

Glaser, B. G. (1992). *Basics of grounded theory analysis*. Mill Valley, CA: Sociology Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Goh, C. (1997). Metacognitive awareness and second language listeners. *ELT Journal*, *51*(4), 361-369.

Goh, C., & Taib, Y. (2006). Metacognitive instruction in listening for young learners. *ELT Journal*, *60*(3), 222-232.

Gopher, D. (1993). The skill of attention control: Acquisition and execution of attention strategies. In D.E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 299-322). Palatino: Asco Trade.

Granena, G., & Long, M. (Eds.). (2013a). *Sensitive periods, language aptitude, and ultimate L2 attainment*. Amsterdam: John Benjamins Publishing.

Granena, G., & Long, M. H. (2013b). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, *29*(3), 311-343.

Gray, D. E. (2004). *Doing research in the real world*. London: Sage.

Griffiths, C. (2003). Patterns of language learning strategy use. *System*, *31*(3), 367-383.

Gu, P. Y., Hu, G., & Zhang, L. J. (2005). Investigating language learner strategies among lower primary school pupils in Singapore. *Language and Education*, *19*(4), 281-303.

Guba, E. G. (1990). The alternative paradigm dialog. In E.G. Guba (Ed.). *The paradigm dialog* (pp.17-30). London: Sage

GUS (Glowny Urzad Statystyczny, 2012). *Oswiata i wychowanie w roku szkolnym 2011/2012* [Education and teaching in the academic year 2011/12, Central Statistical Office of Poland] Retrieved April 7, 2012 from http://stat.gov.pl/obszary-tematyczne/edukacja/edukacja/

Hammersley, M., & Traianou, A. (2012). *Ethics in qualitative research: Controversies and contexts*. London: Sage.

Harlen, W. 2005. Teachers' summative practices and assessment for learning: Tensions and synergies. *The Curriculum Journal, 16*(2), 207 – 223.

Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education*, *4*(3), 365-379.

Harris, K. R., & Graham, S. (1992). Self-regulated strategy development: A part of the writing process. In M., Pressley, J. T. Guthrie, & K. R. Harris, (Eds.). *Promoting academic competence and literacy in school* (pp. 277-309). Academic Press.

Hasselgreen[30], A. (2000). The assessment of the English ability of young learners in Norwegian schools: an innovative approach. *Language Testing*,*17*(2), 261-77.

Hasselgreen, A. (2005). Assessing the language of young learners. *Language Testing*, *22*(3), 337-354.

---

[30] Please note that when this article was published, the author's name was misspelled. Although the author's name is Hasselgreen, her name in the published article is spelled: Hasselgren. The correct spelling of the author's name is used in this thesis but it is acknowledged that the article can best be identified using the following reference:
Hasselgren, A. (2000). The Assessment of the English Ability of Young Learners in Norwegian Schools: An Innovative Approach. *Language testing*,*17*(2), 261-77.

Hawe, E. M., & Dixon, H. R. (2014). Building students' evaluative and productive expertise in the writing classroom. *Assessing Writing*, *19*, 66-79.

Hild, G., & Nikolov, M. (2010). Teachers' views on tasks that work with primary school EFL learners. In M. Lehmann, R. Lugossy, & J. Horvath, (Eds.). *UPRT 2010: Empirical Studies in English Applied Linguistics* (pp. 47-62). Pesc: Lingua Franca Csoport.

Hill, K., & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, *29*(3), 395-420.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288.

Huhta, A., Alanen, R., Tarnanen, M., Martin, M., & Hirvelä, T. (2014). Assessing learners' writing skills in a SLA study: validating the rating process across tasks, scales and languages. *Language Testing*, *31*(3), 307-328.

Inbar-Lourie, O., & Shohamy, E. (2009). Assessing young language learners: What is the construct? In M. Nikolov (Ed.). *The age factor and early language learning* (pp. 83-96). Berlin: Mouton de Gruyter.

Jagatić, M. (1993). Are available teaching time and the number of pupils important factors? In M. Vilke & Y. Vrhovac (Eds.). C*hildren and foreign languages I (*pp. 153–161). Zagreb: Faculty of Philosophy, University of Zagreb.

James, M. E. (2013). *Educational assessment, evaluation and research: The selected works of Mary E. James*. London: Routledge.

Jarvis, J. and Robinson, M. (1997). Analysing educational discourse: an exploratory study of teacher response and support to pupils' learning, *Applied Linguistics*, *18*(2), 212–228.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech, Language, and Hearing Research*, *50*(5), 1280-1299.

Johnstone, R. (2000). Context-sensitive assessment of modern languages in primary (elementary) and early secondary education: Scotland and the European experience. *Language Testing*, *17*(2), 123-143.

Johnstone, R. (2009). An early start: What are the key conditions for generalized success? In J. Enever, J. Moon, & U. Raman (Eds.), *Young learner English*

language policy and implementation: international perspectives (pp. 31-42). Reading: Garnet Education Publishing.

Kennedy, T.J., Nelson, J.K., Odell, M.R.L. and Austin, L.K. (2000). The FLES attitudinal inventory. *Foreign Language Annals 33*(3), 278-289.

Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research*, *12*(2), 211-234.

Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective. *Assessment in Education: Principles, Policy and Practice, 16*(3), 263-268.

Klenowski, V. (2011). Assessment for learning in the accountability era: Queensland, Australia. *Studies in Educational Evaluation*, *37*(1), 78-83.

Knell, E., Haiyan, Q., Miao, P., Yanping, C., Siegel, L. S., Lin, Z., & Wei, Z. (2007). Early English immersion and literacy in Xi'an, China. *The Modern Language Journal*, *91*(3), 395-417.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Addison-Wesley Longman.

Krashen, S. D., Long, M. A., & Scarcella, R. C. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly 13*(4), 573-582.

Krippendorff, K. (2012). *Content analysis: An introduction to its methodology*. London: Sage.

Kubanek-German, A. (1998). *Primary foreign language teaching in Europe-trends and issues*. Cambridge: Cambridge University Press.

Laine, E. J. (1988). The Affective Filter in Foreign Language Learning and Teaching. Report 2: A Validation Study of Filtering Factors with a Focus on the Learner's FL Self-Concept. *Jyvaskyla Cross-Language Studies, No. 15*. Retrieved on March 20th 2015 from http://files.eric.ed.gov/fulltext/ED303992.pdf

Lam, R., & Lee, I. (2010). Balancing the dual functions of portfolio assessment.*ELT journal*, *64*(1), 54-64.

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing*, *13*(4), 285-312.

Lee, I. (2007). Assessment for learning: Integrating assessment, teaching, and learning in the ESL/EFL writing classroom. *Canadian Modern Language Review/La Revue canadienne des langues vivantes*, *64*(1), 199-213.

Lee, I., & Coniam, D. (2013). Introducing assessment for learning for EFL writing in an assessment of learning examination-driven system in Hong Kong. *Journal of Second Language Writing*, *22*(1), 34-50.

Legac, V. (2007). Foreign-language anxiety and listening skill in Croatian monolingual and bilingual students of EFL. In J. Horvath, & M. Nikolov (Eds.) *UPRT 2010: Empirical Studies in English Applied Linguistics* (pp. 217-243). Pesc: Lingua Franca Csoport.

Leow, R. P. (2000). A study of the role of awareness in foreign language behavior. *Studies in Second Language Acquisition*, *22*(04), 557-584.

Leowiecki, T. (1999). W poszukiwaniu sensu i strategii reform oświatowych, [In the search of sense and the strategy of educational reforms] in *Oświata na wirażu [Education on a crossroad]* (pp.11-19). In K. Denek, & T. M. Zimny, (Eds.) Częstochowa: Wyzsza Szkola Pedagogiczna..

Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice, and change. *Language Assessment Quarterly: An International Journal*, *1*(1), 19-41.

Leung, C., & Mohan, B. (2004). Teacher formative assessment and talk in classroom contexts: Assessment as discourse and assessment of discourse. *Language Testing*, *21*(3), 335-359.

Lightbown, P. M., Spada, N. (2008). *How languages are learned* (2nd ed). Oxford: Oxford University Press.

Lincoln, Y. S., Guba, G. E. (1985). *Naturalistic inquiry.* London: Sage.

Little, D. (2005). The Common European Framework and the European Language Portfolio: Involving learners and their judgements in the assessment process. *Language Testing*, *22*(3), 321-336.

Long, M. H. (1981). Input, interaction, and second-language acquisition. *Annals of the New York Academy of Sciences*, *379*(1), 259-278.

Long, M. H. (1990). Maturational constraints on language development. *Studies in second language acquisition*, *12*(03), 251-285.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie, & T. K. Bhatia (Eds.), *Handbook of research on language acquisition* (pp.413-468) San Diego: Academic Press.

Long, M. H. (2005). Problems with supposed counter-evidence to the Critical Period Hypothesis. *International Review of Applied Linguistics in Language Teaching*, *43*(4), 287-317.

Long, M. (2013). Maturational constraints on child and adult SLA. In G. Granena, & M. Long (Eds.), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* (pp. 3-41). Amsterdam: John Benjamins.

Lopriore, L., & Mihaljević Djigunović, J. M. (2009). Attitudinal aspects of early EFL learning. In Szabó, G, Horvath, J. and Nikolov, M. (Eds.) *UPRT 2009: Empirical Studies in English Applied Linguistics* (pp. 3-11). Pesc: Lingua Franca Csoport.

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in second language acquisition*, *19*(01), 37-66.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, *44*(2), 283-305.

MacIntyre, P. D., Baker, S. C., Clément, R., & Donovan, L. A. (2002). Sex and age effects on willingness to communicate, anxiety, perceived competence, and L2 motivation among junior high school French immersion students. *Language Learning*, *52*(3), 537-564.

Mackey, A. (1999). Input, interaction, and second language development. *Studies in Second Language Acquisition*, *21*(04), 557-587.

Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, *27*(3), 405-430.

Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS–NNS and NNS–NNS adult and child dyads. *Language Learning*, *53*(1), 35-66

Mackey, A., & Silver, R. E. (2005). Interactional tasks and English L2 learning by immigrant children in Singapore. *System*, *33*(2), 239-260.

Maleki, Z., & Pazhakh, A. (2012). The effects of pre modified input, interactionally modified input, and modified output on EFL learners' comprehension of new vocabularies. *International Journal of Higher Education*,*1*(1), p128.

Marschollek, A. (2002). Kognitive und affektive Flexibilität durch frmde Schprachen. Eine empirische untersuchung in der primarstufe. [Cognitive and affective flexibility in foreign languages. Empirical study of first grade]. Münster: Lit.

Masgoret, A.M., Bernaus, M. & Gardner, R.C. (2001). Examining the role of attitudes and motivation outside the formal classroom: A test of the mini-AMTB

for children. In Dörnyei, Z. & Schmidt (eds.) Motivation and Second Language Acquisition. Hawaii: University of Hawaii Press.

Matsuzaki Carreira, J. M. (2006). Motivation for learning English as a foreign language in Japanese elementary schools. *JALT JOURNAL*, *28*(2), 135.

Maxwell, J. A., & Mittapalli, K. (2010). Realism as a stance for mixed methods research. In A. Tashakkori, & Teddie, C. (Eds.), *Sage handbook of mixed methods in social & behavioral research* (pp. 145-167). London: Sage.

McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.

Meara, P. (2005). *LLAMA Language Aptitude Tests*. Swansea: Lognostics.

Mercer, N. (2000). *Words and minds: How we use language to think together*. Philadelphia, PA: Psychology Press.

Mihaljević Djigunović, J. (2006). Role of affective factors in the development of productive skills. In M. Nikolov, & J. Horvath (Eds.). *UPRT 2006: Empirical Studies in English Applied Linguistics*, (pp. 9-24). Pesc: Lingua Franca Csoport.

Mihaljević Djigunović, J. (2010). Starting age and L1 and L2 interaction. *International Journal of Bilingualism, 14*(3), 303-314.

Mihaljević Djigunović, J. (2012). Dynamics of learner affective development in early FLL. *Studies in Second Language Learning and Teaching, 2,* 159-178.

Mihaljević Djigunović, J. (2015). Individual differences among young EFL learners: age- or proficiency-related? A look from the affective learner factors perspective. In: Mihaljević Djigunović, J. Medved Krajnovic, M.(Eds.). *Early learning and teaching of English: New dynamics of primary English,* (pp.10-36). Bristol: Multilingual Matters.

Mihaljević Djigunović, J., & Legac, V. (2009). Foreign language anxiety and listening comprehension of monolingual and bilingual EFL learners. *Studia Romanica et Anglica Zagrabiensia*, *53*, 327-347.

Mihaljević Djigunović, J., & Lopriore, L. (2011). The learner: Do individual differences matter? In J. Enever (Ed.), *ELLiE: Early language learning in Europe* (pp. 29-45)*.* London: The British Council.

Mihaljević Djigunović, J., Nikolov, M., & Ottó, I. (2008). A comparative study of Croatian and Hungarian EFL students. *Language Teaching Research*, *12*(3), 433-452.

Miles, M.B. & Huberman, A.M. (1994). *Qualitative data analysis* (2nd ed.). London: Sage.

Mohan, B.A. (2003). *A discourse analysis of choice decision-making.* Unpublished manuscript. University of British Columbia.

Morgan, D. L. (1988). *Focus Groups as Qualitative Research.* London: Sage.

Moyer, A. (2004). *Age, accent, and experience in second language acquisition: an integrated approach to critical period inquiry.* Clevendon: Multilingual Matters.

Munnich, E., & Landau, B. (2010). Developmental decline in the acquisition of spatial language. *Language Learning and Development*, *6*(1), 32-59.

Muñoz, C. (Ed.). (2006). *Age and the rate of foreign language learning* (Vol. 19). Multilingual Matters.

Murphy, V. (2014). *Second Language Learning in the Early School Years: Trends and Contexts: An overview of current themes and research on second language learning in the early school years.* Oxford: Oxford University Press.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice.*Studies in Higher Education*, *31*(2), 199-218.

Nikolov, M. (1999). 'Why do you learn English?''Because the teacher is short.'A study of Hungarian children's foreign language learning motivation. *Language Teaching Research*, *3*(1), 33-56.

Nikolov, M., & Mihaljević Djigunović, J. (2011). All shades of every color: An overview of early teaching and learning of foreign languages. *Annual Review of Applied Linguistics*, *31*, 95-119.

Nisbet, J. &J. Shucksmith (1986). *Learning Strategies.* London: Routledge and Kegan Paul.

Nunan, D. (1987). Communicative language teaching: making it work, *English Language Teaching Journal*, 41: 136–45.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge University Press.

Ohta, A. (1995). Applying sociocultural theory to an analysis of learner discourse: learner-learner collaborative interaction in the zone of proximal development. *Issues in Applied Linguistics*, *6*(2), 93–121.

Oliver, R. (1998). Negotiation of meaning in child interactions. *The Modern Language Journal*, *82*(3), 372-386.

Oliver, R. (2000). Age differences in negotiation and feedback in classroom and pairwork. *Language Learning*, *50*(1), 119-151.

Oliver, R. (2002). The patterns of negotiation for meaning in child interactions. *The Modern Language Journal*, *86*(1), 97-111.

Oliver, R., & Mackey, A. (2003). Interactional context and feedback in child ESL classrooms. *The Modern Language Journal*, *87*(4), 519-533.

O'Malley, J. M., & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.

Owings, R. A., Petersen, G. A., Bransford, J. D., Morris, C. D., & Stein, B. S. (1980). Spontaneous monitoring and regulation of learning: A comparison of successful and less successful fifth graders. *Journal of Educational Psychology*, *72*(2), 250.

Oxford, R. L. (1999). Anxiety and the language learner: New insights. In J. Arnold (Ed.), *Affect in language learning*, (pp. 58-67). Stuttgart: Ernst Klett Sprachen.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. London: Sage.

Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, *8*(3), 265-289.

Piaget, J. (1969). *The psychology of the child*. New York: Basic Books.

Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning*, *44*(3), 493-527.

Pinter, A. (2007). Some benefits of peer–peer interaction: 10-year-old children practising with a communication task. *Language Teaching Research*, *11*(2), 189-207.

Pinter, A. (2011). Children learning second languages. London: Palgrave Macmillan.

Prodromou, L. (1995). The backwash effect: from testing to teaching. *ELT Journal*, *49*(1), 13-25.

Pryor, J., & Crossouard, B. (2008). A socio-cultural theorisation of formative assessment. *Oxford Review of Education*, *34*(1), 1-20.

Punch, F. K. (2004). *Introduction to social research: quantitative and qualitative approaches*. London: Sage.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science, 28*(1), 4–13.

Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, *18*(4), 429-462.

Rea-Dickins, P. (2006). Currents and eddies in the discourse of assessment: a learning-focused interpretation1. *International Journal of Applied Linguistics*, *16*(2), 163-188.

Rea-Dickins, P., & Gardner, S. (2000). Snares and silver bullets: disentangling the construct of formative assessment. *Language Testing*, *17*(2), 215-243.

Ridderinkhof, K. R., & van der Molen, M. W. (1997). Mental resources, processing speed, and inhibitory control: A developmental perspective. *Biological Psychology*, *45*(1), 241-261.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning*, *45*(2), 283-331.

Rosa, E., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness. *Studies in Second Language Acquisition*, *21*(4), 511-556.

Rubin. J. (1975). What the "good language learner" can teach us. *TESOL Quarterly, 9*(1), 4I -51.

Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, *37*(1), 15-24.

Ruiz-Primo, M.A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment*, *11*(3-4), 237-263.

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field methods*, *15*(1), 85-109.

Sadler, D. R. (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education*, *14*(3), 387-392.

Sandelowski, M. (1995). Qualitative analysis: What it is and how to begin.*Research in nursing & health*, *18*(4), 371-375.

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: reactions to tests. *Journal of Personality and Social Psychology*, *46*(4), 929.

Schmidt, R. (1992). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206-226.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, JW. Sew, T.

Suthiwan, & I. Walker, *Proceedings of CLaSIC 2010*, Singapore, December 2-4 (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.

Schmidt, R., & Frota, S.N. (1986). Developing basic conversational ability in a second language: a case study of an adult learner of Portuguese. In R. R. Day (Ed.). *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.

Schneider, W., & Pressley, M. (2013). *Memory development between two and twenty*. Philadelphia, PA: Psychology Press.

Schumann, J. (2001). Learning as foraging. In Z. Dörnyei and Schmidt, R. (Eds.). *Motivation and second language acquisition*. Honolulu: University of Hawaii, Second Language Teaching and Curriculum Centre.

Scriven, M. (1967). The methodology of evaluation. In *Perspectives of curriculum evaluation,* ed. R.W. Tyler, R.M. Gagne, and M. Scriven, 39–83. Chicago, IL: Rand McNally.

Shaughnessy, J. J., Zechmeister, E. B. and Zechmeister, J.S. (2003). *Research methods in psychology*. New York: Alfred A. Knopf.

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*(2), 63-75.

Shepard, L. A. (2006). Classroom assessment. *Educational Measurement*, *25*(4), 623-646.

Sidhu, G. K., Chan, Y. F., & Sidhu, S. K. (2011). Students' reactions to school-based ral assessment: bridging the gap in Malaysia. *Asian EFL Journal*, *13*(4), 300-327.

Siegler, R. S. (1994). Cognitive variability: a key to understanding cognitive development. *Current Directions in Psychological Science 2*(1), 1-5.

Skehan, P. (1996). Second language acquisition research and task-based instruction. In J. Willis, & D. Willis, (Eds.). *Challenge and change in language teaching*. Oxford: Heinemann.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, *49*(1), 93-120.

Sly, L. (1999). Practice tests as formative assessment improve student performance on computer-managed learning assessments. *Assessment & Evaluation in Higher Education*, *24*(3), 339-343.

Smith, P. K., Cowie, H., & Blades, M. (1998). *Understanding children's development*. Oxford: Blackwell.

Somekh, B. & Lewin, C. (2005). *Research methods in social sciences*. London: Sage.

Spadaro, K. (2013). Maturational constraints on lexical acquisition in a second language. In G. Granena, & M. Long (Eds.), *Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment* (pp. 43-68). Amsterdam: John Benjamins.

Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, *52*(1), 119–158.

Storch, N. (2009). *The nature of pair interaction. Learners' interaction in an ESL class: its nature and impact on grammatical development*. Berlin, Germany: VDM Verlag.

Stoynoff, S. (2012). Looking backward and forward at classroom-based language assessment. *ELT Journal*, *66*(4), 523-532.

Strauss, A. & Corbin, J. (1990). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Los Angeles: Sage.

Swaffield, S. (2011). Getting to the heart of authentic assessment for learning. *Assessment in Education: Principles, Policy & Practice*, *18*(4), 433-449.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J.P. Lantof (Ed.), *Socio-cultural theory and second language learning*, (pp. 97-114). Oxford: Oxford University Press.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, *82*(3), 320-337.

Szpotowicz, M., Mihaljevic Djigunovic, J., & Enever, J. (2009). Early language learning in Europe: A multinational, longitudinal study. In J. Enever, J. Moon & U. Raman (Eds.), *Young learner English language policy and implementation: International perspectives* (pp. 141-147). Reading, UK: Garnet Publishing.

Szpotowicz, M., & Szulc-Kurpaska, M. (2009). *Teaching English: To Young Learners*. Wydawnictwo Naukowe PWN.

Teasdale, A., & Leung, C. (2000). Teacher assessment and psychometric theory: A case of paradigm crossing? *Language Testing*, *17*(2), 163-184.

Thompson, M., and Wiliam, D., (2007). Tight but loose: A conceptual framework for scaling up school reforms. Paper presented at a Symposium entitled 'Tight but loose: Scaling up teacher professional development in diverse contexts' at the annual conference of the American Educational Research Association, April 9–11, in Chicago, IL. Retrieved March 5, 2010 from https://www.ets.org/Media/Research/pdf/RR-08-29.pdf

Torrance, H. (1995). Teacher involvement in new approaches to assessment. In H. Torrance, (Ed.). *Evaluating authentic assessment* (pp. 44–56). Buckingham: Open University Press.

Tsang, W. K. (2004). Feedback and uptake in teacher-student interaction: An analysis of 18 English lessons in Hong Kong secondary classrooms. *RELC Journal*, *35*(2), 187-209.

Tuckman, B. W., & Harper, B. E. (2012). *Conducting educational research*. Plymouth: Rowman & Littlefield Publishers.

Turek, A. (2013). Engaging young learners in L2 research. *Language Study Working Papers*, *5*, 32-40.

Upshur, J. A., & Turner, C. E. (1995). Constructing rating scales for second language tests. *ELT Journal*, *49*(1), 3-12.

van Lier, L. (2014). *Interaction in the language curriculum: Awareness, autonomy and authenticity*. London: Routledge.

Vandergrift, L. (2002). 'It Was Nice to See That Our Predictions Were Right': Developing Metacognition in L2 Listening Comprehension. *Canadian Modern Language Review/La revue canadienne des langues vivantes*, *58*(4), 555-575.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, *60*(2), 470-497.

Vann, R. J., & Abraham, R. G. (1990). Strategies of unsuccessful language learners. *TESOLQuarterly*, *24*(2), 177-198.

Victori, M., & Lockhart, W. (1995). Enhancing metacognition in self-directed language learning. *System*, *23*(2), 223-234.

Vilke, M., & Vrhovac, Y. (1995). *Children and foreign languages II*. Zagreb: Faculty of Philosophy, University of Zagreb.

Vurpillot, E. (1968). The development of scanning strategies and their relation to visual differentiation. *Journal of Experimental Child Psychology*, *6*(4), 632-650.

Vygotsky, L. S. (1987). *Mind in society: The development of higher psychological processes*. Harvard: Harvard University Press.

Walsh, S. (2006). *Investigating classroom discourse*. London: Routledge.

Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language teaching research*, *11*(2), 121-142.

Watkins, C. (2003). *Learning: A sense-maker's guide*. London: Association of Teachers and Lecturers.

Wen, Z. (2012). Working memory and second language learning. *International Journal of Applied Linguistics*, *22*(1), 1-22.

Wen, Z., & Skehan, P. (2011). A new perspective on foreign language aptitude research: building and supporting a case for "working memory as language aptitude" *Journal of English Language, Literatures in English and Cultural Studies*, (60), 15-44.

Wenden, A. (1987). Conceptual background and utility. In Wenden, A., & Rubin, J. (Eds.). (1987). *Learner strategies in language learning*. London: Prentice-Hall International.

Wilden. E. & Porsch, R. (2014). Children's receptive EFL competences at the end of primary education: Evidence from the German "Ganz In" In J. Enever, Lindgren, & E. Ivanov, S. (Eds.) *Conference Proceedings from Early Language Learning: Theory and Practice 2014.* Umea: University of Umea.

Wiliam, D. (2001). An overview of the relationship between assessment and the curriculum. In D. Scott (Ed.) *Curriculum and Assessment* (pp. 165-181). Westport: Albex Publishing.

Wiliam, D. (2009). Assessment for learning: why, what and how? An inaugural professorial lecture by Dylan Wiliam. *Institute of Education University of London*.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation*, *37*(1), 3-14.

Williams, J. (2001). The effectiveness of spontaneous attention to form. *System*, *29*(3), 325-340.

Wilson, J. T., Scott, J. H., & Power, K. G. (1987). Developmental differences in the span of visual memory for pattern. *British Journal of Developmental Psychology*, *5*(3), 249-255.

Winer, G. A., Craig, R. K., & Weinbaum, E. (1992). Adults' failure on misleading weight-conservation tests: A developmental analysis. *Developmental Psychology*, *28*(1), 109-120.

Winter, G. (2000). A comparative discussion of the notion of validity in qualitative and quantitative research. *The qualitative report*, *4*(3-4). Retrieved December 12, 2011 from http://www.nova.edu/ssss/QR/QR4-3/winter.html

Wood, D. (1998). *How children think and learn: The social contexts of cognitive development*. London: Blackwell Publishing.

Yung, B. H. W. (2002) Same assessment, different practice: professional consciousness as a determinant of teachers' practice in a school-based assessment scheme, *Assessment in Education,* 9(1), 97–117.

Zhang, L. J. (2001). Awareness in reading: EFL students' metacognitive knowledge of reading strategies in an input-put environment. *Language Awareness, 11*(4), 248–268.

Appendices

Appendix 1: Ten principles of Assessment for Learning

The 10 principles are quoted from a leaflet produced by the Assessment Reform Group in 2002. The title of the leaflet was: *Assessment for Learning: 10 Principles. Research-based principles of assessment for learning to guide classroom practice.*

'Assessment for Learning is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there.

Assessment for learning should be part of effective planning of teaching and learning

A teacher's planning should provide opportunities for both learner and teacher to obtain and use information about progress towards learning goals. It also has to be flexible to respond to initial and emerging ideas and skills. Planning should include strategies to ensure that learners understand the goals they are pursuing and the criteria that will be applied in assessing their work. How learners will receive feedback, how they will take part in assessing their learning and how they will be helped to make further progress should also be planned.

Assessment for learning should focus on how students learn

The process of learning has to be in the minds of both learner and teacher when assessment is planned and when the evidence is interpreted. Learners should become as aware of the 'how' of their learning as they are of the 'what'.

Assessment for learning should be recognised as central to classroom practice

Much of what teachers and learners do in classrooms can be described as assessment. That is, tasks and questions prompt learners to demonstrate their knowledge, understanding and skills. What learners say and do is then observed and interpreted, and judgements are made about how learning can be improved. These assessment processes are an essential part of everyday classroom practice and involve both teachers and learners in reflection, dialogue and decision making.

Assessment for learning should be regarded as a key professional skill for teachers

Teachers require the professional knowledge and skills to: plan for assessment; observe learning; analyse and interpret evidence of learning; give feedback to learners and support learners in self-assessment. Teachers should be supported in developing these skills through initial and continuing professional development.

Assessment for learning should be sensitive and constructive because any assessment has an emotional impact

Teachers should be aware of the impact that comments, marks and grades can have on learners' confidence and enthusiasm and should be as constructive as possible in the feedback that they give. Comments that focus on the work rather than the person are more constructive for both learning and motivation.

Assessment should take account of the importance of learner motivation

Assessment that encourages learning fosters motivation by emphasising progress and achievement rather than failure. Comparison with others who have been more successful is unlikely to motivate learners. It can also lead to their withdrawing from the learning process in areas where they have been made to feel they are 'no good'. Motivation can be preserved and enhanced by assessment methods which protect the learner's autonomy, provide some choice and constructive feedback, and create opportunity for self-direction.

Assessment for learning should promote commitment to learning goals and a shared understanding of the criteria by which they are assessed

For effective learning to take place learners need to understand what it is they are trying to achieve - and want to achieve it. Understanding and commitment follows when learners have some part in deciding goals and identifying criteria for assessing progress. Communicating assessment criteria involves discussing them with learners using terms that they can understand, providing examples of how the criteria can be met in practice and engaging learners in peer- and self-assessment.

Learners should receive constructive guidance about how to improve

Learners need information and guidance in order to plan the next steps in their learning. Teachers should: pinpoint the learner's strengths and advise on how to develop them; be clear and constructive about any weaknesses and how they might be addressed; provide opportunities for learners to improve upon their work.

Assessment for learning develops learners' capacity for self-assessment so that they can become reflective and self-managing

Independent learners have the ability to seek out and gain new skills, new knowledge and new understandings. They are able to engage in self-reflection and to identify the next steps in their learning. Teachers should equip learners with the desire and the capacity to take charge of their learning through developing the skills of self-assessment.

Assessment for learning should recognise the full range of achievements of all learners

Assessment for learning should be used to enhance all learners' opportunities to learn in all areas of educational activity. It should enable all learners to achieve their best and to have their efforts recognised.' (ARG, 2002b, p.2)

Appendix 2: Diagnostic behaviours of teachers identified by Edelenbos and Kubanek-German (2004)

'11 potential diagnostic activities were identified:

1) checking whether the material=content of the lesson has been understood (e.g., observing the desired behaviour, checking the students' language production);

2) use of an observation schedule or checklist;

3) administering and interpreting the results of a test;

4) questioning (e.g., posing questions, offering stimuli for students to evaluate, add to or complete)

5) checking homework;

6) informal, enquiry-focused observation of the whole class;

7) informal, enquiry-focused observation of individual students;

8) goal=task oriented observation of the whole class;

9) goal=task oriented observation of individual students;

10) assisting students in interpreting feedback and undertaking appropriate action;

11) monitoring systematic errors.' (Edelenbos & Kubanek-Germna, 2004: 264-5)

Appendix 3: Notes from an interview with the school director

1) School – 75 years of experience worldwide
2) Opened in Poland – 1996; four branches in <CITY> and <CITY> plus in-company teaching
3) 1996-2000 – mostly adults and older teens; 2000-2004 – as young as 12; from 2004 7-11 year olds; and from 2005 – 5-11 year olds.
4) Current numbers: about 2000 students including about 300 7-11 year olds.
5) Curriculum: a set of Can Do statements based on course books for each course, Primary courses (7-9) - <NAME OF COURSEBOOK>, Pre-Teen courses (10-11) - <NAME OF COURSEBOOK>;
6) Assessment – no explicit guidance in the curriculum documents, teachers are asked to use their professional judgement in how they assess. There are, however, expectations to produce reports for parents with Polish grades, on a scale from 1-fail to 6-exceeds expectations, twice a term, mid-way and at the end of each term.
7) Staffing of courses for 7-11: all teachers have YL certificate; a number of specialists recruited since 2004; additionally since 2004 YL teaching constituted a significant part of the in-service training plan for all teachers.
8) From its own marketing research the language school knows that parents enrol children mostly because they want to supplement the provision at day school, want children to have better grades at day school and/or want children to learn at a higher level than the day school offers.

Appendix 4: Consent Form (Teachers)

NB: The current study was initially based for two years at London Metropolitan University before moving to the University of Reading

<div align="right">City, <date></div>

Dear Teacher,

In the coming year, I would like to conduct a small scale study which is a part of my PhD project and investigates how Assessment for Learning works in the context of teaching English to learners aged 7-11.

As you have been timetabled to teach classes of learners at that age, I am writing to you to ask if you would kindly agree to participate in the study.

As a part of this research, I would like to ask you to complete a questionnaire, participate in one to one interview with me and sometime later a discussion group (focus group) with other teachers.
I would also like to ask your permission for some of your lessons to be observed and video recorded.

All data will be recorded anonymously, numerically coded and I will never use your name in observation notes, the questionnaire, the interview or the focus group. Confidentiality of the collected data will be maintained at all times: all hard copies of the data will be kept in a safe, locked cupboard, all digital data will be kept on a password protected computer and when in transfer they will be saved on an encrypted and password protected memory stick.

You will be free to withdraw your permission at any time and without having to give a reason. If you decide to withdraw your permission, please inform me about it in writing.

All data will be managed according to 29th August 1997 Data Protection Act (Full text: DZ. U. 2002 R. NR 101 POZ. 926, ZE ZM.)

Please sign the reply slip below and return to me at your nearest convenience.
Should you have any questions about the research or procedures involved please do not hesitate to contact ma at aga.turek@op.pl or talk to me in person. (See overleaf)

Kind regards,
Aga Turek

_ _
I agree to participate in the research.

I have been informed that this will include completing a questionnaire, participating in an interview and a discussion group (focus group), having some of my lessons observed and video recorded.

I understand that all data will be managed according to 29th August 1997 Data Protection Act (Full text: DZ. U. 2002 R. NR 101 POZ. 926, ZE ZM.)

I understand that I am free to withdraw my permission at any time and without having to give a reason and if I choose to do so I will do it in writing.

Teacher's name: ....

Teacher's signature: ....

Date: ....

Appendix 5: Consent Form (Parents)

NB: The current study was initially based for two years at London Metropolitan University before moving to the University of Reading

City, <date>

Dear Parents and Carers,

In the coming semester we would like to conduct a small scale study to investigate how Assessment for Learning is implemented in the context of teaching English to learners aged 7-11. This will constitute a part of a PhD study of one of our teachers, Aga Turek. The study is based at London Metropolitan University in the UK.

As your child is in this age category, we would like to kindly ask your permission to include your child in this study. We would also like to ask your permission to:

a) For some of your child's lessons to be observed and notes made
b) For the observed lessons to be video recorded

All data will be recorded anonymously, numerically coded and we will never use your child's name. Confidentiality of the collected data will be maintained at all times: all hard copies of the data will be kept in a safe, locked cupboard, all digital data will be kept on a password protected computer and when in transfer they will be saved on an encrypted and password protected memory stick.

You will be free to withdraw your permission at any time and without having to give a reason. If you decide to withdraw your permission, please inform us about it in writing. All data will be managed according to 29th August 1997 Data Protection Act (Full text: DZ. U. 2002 R. NR 101 POZ. 926, ZE ZM.) .Please sign the reply slip below and return to your child's teacher by <deadline>.

_ _ _ _ _ _ _ _ _ _ _ _ _ ____ _

I agree for my child to:
a) Participate in lesson that will be observed and notes taken
b) Participate in lessons that will be video recorded

I understand that all data will be managed according to 29th August 1997 Data Protection Act (Full text: DZ. U. 2002 R. NR 101 POZ. 926, ZE ZM.)

I understand that I am free to withdraw my permission at any time and without having to give a reason and if I choose to do so I will do it in writing.

Child's name: ....

Parent's name: ....

Parent's signature: ....Date: ....

Appendix 6: Letter granting access to the school for the purposes of conducting the study (scanned).

Please note that the white rectangles cover information which identifies the school and have been added to ensure anonymity.
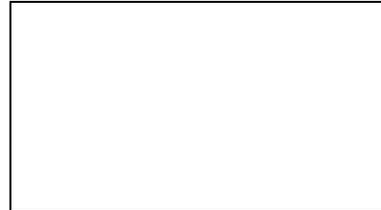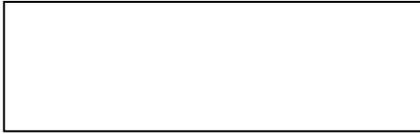
Warsaw, 29th September 2010

London Metropolitan University
166-220 Holloway Road
N7 8DB London

To whom it may concern,

Re: Access to Warsaw teaching centre

I can confirm that access to the Teaching Centre will be provided to Aga Turek for the purposes of conducting her PhD research. Such access will include opportunities to observe lessons, check records, conduct interviews with students and teachers and approach their parents for the purposes of conducting surveys and interviews. Prior approval will be given to all letters to parents, teachers and learners regarding written permission to conduct this study.

Yours sincerely,

Appendix 7: Comparison between the study reported by Colby-Kelly and Turner (2007) and the present study

Area of comparison	Colby-Kelly & Turner, 2007	The present study
Context	Pre-sectional English for Academic Purposes course in a Canadian university	Teaching English to Young Learners in a language school in Poland
Age group	Adults	Children 7-11
Participants	9 teachers 42 students	8 teachers 148 students
Duration	Three months	29 weeks
Study design	A descriptive study	A descriptive and exploratory study
Research questions	1. What are teacher and student perceptions of formative assessment in a second language (L2) classroom setting? 2. What is the nature of formative assessment in a second language classroom setting? 3. What evidence can be found that formative assessment benefits learning?	1. How do teachers understand AfL after receiving a limited amount of training and being encouraged to use AfL techniques for at least one academic year when teaching English to young learners aged 7-11? 2.1 How do teachers' translate their understanding of AfL into classroom practice in a TEYL context with students aged 7-11 in a private language school in Poland? 2.2 Do teachers report any changes in their practice of using AfL over time? 3. What is the observable impact of AfL on classroom interactions in a TEYL context?
Methods	Mixed-method	Mixed-method
Research Tools	Observation schedules: 'adapted from the study described in Turner (2001, 2006) to catalogue assessment episodes; duration; teacher or student initiation; number of students involved; origin; skill (listening, speaking, writing, reading, or general); and focus (grammar, vocabulary, pronunciation, pragmatics, or meaning-based)' (Colby-Kelly & Turner, 2007 p. 19)	Observation schedules: based on the observation schedule used in the ELLiE study (Enever, 2014 personal communication) and informed by Colby-Kelly & Turner (2007) and literature review cataloguing AfL techniques; duration; classroom interactions; number of students actively engaged in a task; skills and interactions.
	Teacher Questionnaire based on literature review and analysis of curriculum documents	Delayed teacher questionnaire based on the findings from the analysis of lesson observations, conducted sixteen months after the observations.
	-	Teacher Interviews and a focus group

	Curriculum Documents (CDs)	Record of Work Done documents (RoWDs)
	Student interviews	Student perspective was not researched in the current study but would constitute a valuable focus for future research.
Data Sets	16 audio recorded, non-participant lesson observations	28 non-participant lesson observations: 26 video recorded and 2 audio recorded[31]
	9 questionnaires completed by teachers; each with 51 items in 4 categories: assessment and students, assessment and teachers, assessment and learning, and course assessment needs, 4-point Likert used in teacher questionnaire plus six open ended questions asking for comments	8 delayed teacher questionnaires completed by teachers; each with 18 Likert-type items plus four open-ended questions. The questionnaire focused on use and impact of AfL.
	No detailed information published regarding the data set comprising curriculum documents	14 sets of RoWDs from a 28-week period collected, in total 448 lessons recorded
	12 audio recorded, student interviews, field notes taken	n/a
	n/a	8 audio-recorded teacher interviews and 1 focus group, field notes taken
Data Analysis	Interpretational analysis including: Quantitative: frequency counts Qualitative: content analysis, comparative analysis Triangulation of data	Interpretational analysis including: Quantitative: frequency counts Qualitative: content analysis, comparative analysis Triangulation of data

---

[31] Two lessons were not video recorded due to the lack of parental consent. These lessons were audio recorded. Consent was granted for audio recordings.

Appendix 8: Demographic Form

Assigned teacher code (for the researcher's use only, please do not complete)	
Name	
Gender	
Age – please choose one category	21-29 30-39 40-49 50-59 60+
Please indicate which qualifications you hold.	
How many years of experience in teaching do you have?	
How many years of experience of teaching English as a foreign language do you have?	
How many years of experience of teaching English as a foreign language to students aged 7-11 do you have?	
How many years of experience of using AfL have you got?	
In what context did you gain experience of using AfL?	
Which of the above courses has had the greatest impact on you teaching practice?	
What is your role at the school?	
Which language(s) do you consider you mother tongue(s)?	
To be completed in September 2011 before the cross-sectional data collection stage commenced but a year after the initial AfL training took place) List any training courses you have attended since September 2010?	
To be completed in June 2012 (the end of the cross-sectional data collection stage) List any training courses you have attended since September 2011?	

Appendix 9: Teacher Interview Aide Memoire

Introduction: Thank you for agreeing to participate in this study and for taking the time to participate in this interview. In your answers, please think about classes with learners aged 7-11 only. Please remember that I am interested in your views and I would appreciate it if you could share all your thoughts and opinions about AfL with me. Please, do not feel that you are in any way expected to know answers to my questions or that there is something in particular that I am hoping to hear from you. In fact, I would prefer to call the questions, prompt as this is what they are; they are prompt for our discussion. I am interested in your perspective on the topics which I will bring up through the questions. As you know from the consent form, this is a strictly confidential conversation and your name will never be associated with the content of this interview. If at any point during the interview, you feel like adding additional comments or returning to a previously discussed issue, please feel free to do so. I would also like to inform you that the draft findings from the analysis will be shared with you during the focus group discussion which will take place later this year. Of course, as I said before, anonymity will be ensured. Are there any questions that you like to ask me before we start? Can we start?

How would you explain what Assessment for Learning is to a teacher that does not know anything about it?

How would you explain what AfL is to a parent of one of your students?

How do you implement AfL in your lessons with learners aged 7-11?

Which part of the lesson do you use AfL in? Why?

Have you adapted the AfL techniques that you learnt about in the training sessions? How? Why?

What do you think of AfL as a means of assessment?

Would you recommend using AfL to other teachers who haven't used it before? Anything in particular? Why? Anything to be mindful of?

How would you describe your attitude towards using AfL in your YL classes?

How would you describe the attitude of your learners towards AfL?

How important, relevant and helpful is using AfL to delivering good quality teaching of English to young learners? Please comment on the teacher's and learners perspectives.

Have you noticed anything that you'd describe as positive or negative impact of AfL in your lessons?

What do you see as the most important influence of AfL on your practice, if anything?

Is there anything you'd like to add?

Appendix 10: Focus Group prompts

1: In the interviews some teachers were saying that they started using or are planning to 'use more AfL' with age primary (7-9) and pre-teens (10-11). Could you comment on what the phrase 'use more AfL' could refer to? What might the main reasons for such choices be? What outcomes might be expected from that? What characteristics of AfL make it useful in the lessons with children?

2: In the interviews, you also talked about AfL being helpful in learning. I would like to find out more about this area. What do you think about this? How can AfL support learning?

3: It seems from the initial analysis of the data that AfL is generally perceived as a form aiding continuous assessment that provides the teachers with ongoing knowledge of how their students are coping with tasks. Would you agree with this? How does this knowledge affect classroom practice?

4: Some research into AfL suggested that AfL might lead to improvement in students' achievement. These studies were carried on large samples of students and over a period of many years so I am not looking here to confirm or refute these. What I would like to understand better though is why the improved achievement might happen? In what way could AfL improve achievement? What is your opinion about this issue?

5: Has AfL changed anything in your teaching?

6: Any other comments?

Appendix 11: Examples of Records of Work Done (ROWDs)

Example 1: T1's ROWDs from six lessons. The bolded black-rimmed rectangles have been superimposed digitally to cover the teacher's name.

Example 2: T5's Records of Work Done from six lessons. The bolded black-rimmed rectangles have been superimposed digitally to cover the teacher's name.

Appendix 12: Lesson Observation Schedule

Lesson Observation Template Date: Group code:					Teacher code: Student codes: Lesson duration:				
Section 1: Classroom Practice					Section 2: Observable Impact				
Column 1. Time (beginning – end /duration)	Type of activity e.g. speaking = class survey Is the AfL technique used as an integral part of the lesson (IP) or does it seem to be a separate mini-stage of an activity (Sm-S) or a separate activity (SA)?	AfL technique used e.g. success criteria	What is the purpose for using this technique? Is it made explicit to students? If so, how?	Record in workbooks/ notebooks e.g. Yes -SC written in	Reaction to the AfL technique e.g. refer to SC/ look refer to SC/ talk about	Engagement with the task	Classroom interactions L, L-L, T-1L, T-C, IND,	Can any other impact on the lesson/students /teachers be observed?	Additional Comments

The empty cells were used to record field notes.

Appendix 13: Delayed Questionnaire for Teachers

Thank you for agreeing to complete this questionnaire. Please read the questions and instructions carefully. It will take about 10 minutes to fill in the questionnaire.

Q1:
Read each item and then write in the number of hours that you spent in the following types of training in the academic year 2012/13.

	How much time was devoted to training about any form of assessment.
	How much time was devoted to training about Assessment for Learning.
	The total amount of time that you spent observing other colleagues teach.
	What percentage of the time that you spent observing other colleagues in 2012-13 focused on/was devoted to the use of AfL. E.g. if you observed 5 60min lessons and 3 of them included use of AfL, enter 60%

Q2:
Read each item and then insert the relevant score in the left column to indicate how often you have used the following AfL techniques with Young Learners since the beginning of Spring Term 2013 (February 2013) <u>with students aged 7-11 ONLY</u>. If there are other AfL techniques that you have used, please add them in the space provided and score them as well. Use the following scale to record your answers.

1	2	3	4	5
never or almost never	Rarely	sometimes	often	every or almost every lesson

	Circle mistakes without explanations
	Colour Coding
	'I can' statements
	Increased thinking time
	Learning Partners
	Mind Maps
	Next steps/ Next time...
	Peer-assessment
	Perfect Purple/ Red to Remember
	Self-assessment
	Sharing good and bad model
	Sheriff's star
	Smiley Faces
	Star charts
	Success Criteria
	Traffic Lights
	Two Stars and a Wish
	WALT

Q3: In Q2 you indicated AfL techniques which you have used since February 2013. Why did you choose to use these techniques over other techniques?

Q4: What was the reason(s) for not using the techniques which you scored 1 in Q2?

Q5: Please think about your three most frequently used techniques. Which activity types were the techniques mainly used with and/or why were these techniques useful techniques with these activities?

Q6: Please think about your three most frequently used techniques again. What was the impact of using them on the teaching and learning processes?

Appendix 14: A Completed Observation Record

Lesson Observation Template Date: 10th February 2011 (Lesson 3- L3) Group code: G4 (age: 7-9)					Teacher code: T7 Student codes: Present: S33, S34, S35, S37, S38, S39, S40, S41, S42, S43, S44, Absent: S36 Lesson duration: 60min (start time: 16:30 end time: 17:30)				
Section 1: Classroom Practice					Section 2: Observable Impact				
Colum n 1. Time (begin ning – end /durati on)	Type of activity e.g. speaking = class survey Is the AfL technique used as an integral part of the lesson (IP) or does it seem to be a separate mini-stage of an activity (Sm-S) or a separate activity (SA)?	AfL technique used e.g. success criteria	What is the purpose for using this technique? Is it made explicit to students? If so, how?[32]	Record in workbooks/ notebooks e.g. Yes -SC written in	Reaction to the AfL technique e.g. refer to SC/ look refer to SC/ talk about	Engage ment with the task	Classroom interactions	Can any other impact on the lesson/st udents /teachers be observed ?	Additional Comments
16:30-16:32 /2min	Hello chant	-	-	N	-	E:10 DE:0 1 child late	T-WC	-	-
16:32 – 16:35	Vocabulary revision: instructions	-	T: *'Now I want you to show me if*	N	-	E: 9 DE: 2	T-WC	-	Revision of vocabulary from

---

/3min	for an activity with flashcards		*your partner is right. Is it a good action?'* (to assess peer's performance)				TR: instructions giver SR: instructions receivers		the previous lesson.
16:35 – 16:38 /3min	Pair work	PM (IS): clap if action correct 'bravo'	-	N	PER to correct answers R: react to peer's claps	E: 11 DE: 0	S-S TR: monitor SR: the assessors and the assessed	Attempts to improve performance based on peer feedback	In some pairs, children attempt to correct their answers when they don't get a clap. In others, children simply proceed to the next flashcard.
16:38- 16:41 /3min	Vocabulary revision: adverbs of place, pictures on board,	THUD (IP)	T mentions nothing Observable: WC responds and T targets SS who seem to misunderstand and rectifies misunderstan	N	R: SS look at peers when showing thumbs up or down	E: 8 DE: 3	T-WC TR: instructor, assessor, helper SR: demonstrators of own knowledge	-	Chn come up to WB with pictures and place them next to, opposite etc T focused on SS who are by IWB and the rest of the class are trusted to watch

			dings by questioning.						but 3 children become disengaged. There was no clear link between this and the previous task.
16:45-16:52 /7min	Introduction of a new theme/ set of vocabulary: transport	QQ (Sm-S): Mind map PM: THUD (Sm-S)	T: *'Let me see how many different words about transport you already know. If you think it's to do with transport say. We want many ideas.'* (T is establishing where children are in their learning.)	Y (mind map in notebooks – copied from the board after the whole class activity)	UL: one child asks a friend/rehearses before putting a hand up S: 'Helicopter?'	E: 9 DE: 2	S-SS TR: scribe, facilitator SR: share current knowledge	-	Chn work in pairs to brainstorm, t monitors

17:00-17:01 /2min	Instructions for Cool English 5 Activity Book p.49, a. 6	-	-	Y (corrections In the workbooks and the peer's signature)	-	E: 9 DE: 2	T-WC TR: source of instructions SR: instructions receivers	-	Lacked modelling how to PM, reward stickers
17:02 – 17:08 /7 min	AB p.49, a. 6	-	-	N	NER one upset child *'Teacher I not know what I do'* (T supports IND)	E: 9 DE: 1 Plus one student working with the T	SS: IND T-S TR: individual work with one student who didn't know what to do (1 out of the 9 in E:9 to the right) SR: individual work	-	
17:08-17:10 /2min	Check answers with a partner	LPs (IP) PM (SA)	T: *'You're the teacher now, mark*	N	PER children draw happy faces unprompted	E: 11 DE: 1	S-S TR: monitor SR: assessors	-	

			partner's work'		if they have correct answers				
17:10-17:15 /5min	Reading comprehension (True or False) Pair work – SS complete a task together	LPs (IP)	-	N	-	E: 8 - 9 DE: 3 – 2	S-S TR: monitor SR: complete the task in pairs	-	This section lacked clear instructions Distracted by interesting stationery that belonged to one of the students – no reaction from the teacher.
17:15-17:18 /3min	Reading comprehension feedback	THUD (SA)	-	N	-	E: 10 DE: 1	T –WC T- S TR: reads out sentences SR: individual students give answers, WC - THUD	-	Which qq? Indicate in SB how many qq to ask? (3)

17:19 – 17:23	Instructions for speaking: mini survey	SC: Scribed in IWB 1 – 3 questions 2 – 2 people 3 – 3 transport words (Sm-S)	*'When you ask you don't know what to do, when can you look? What will help you with this task?'*	N but SC on IWB	UL: 2 students confer to clarify instructions – R on the IWB	E: 11 DE: 0	T-WC S-SS TR: facilitates SR: co-create success criteria	Very good engagement level – all SS on task	-
17:23-17:29 /2min	Speaking: mini survey		-	N	R: children clearly look at IWB to check SC while performing the task	E: 11 DE: 0	S-S S-T TR: monitor and participant SR: individual completion of survey notes by asking questions to partner		-
17:29-17:30 /1min	Evaluation of own effort	Stars: SS give themselves stars out of	-	Y (stick in pieces of coloured paper with	UL: comparisons of number of stars	E: 11 DE: 0	IND T-S	Very good engagem ent level	SS individually suggest rating and T confirms – this was a little

315

		maximum three (SA)		stars, dated and signed by T)	PER congratulations for other students		TR: monitor/ assessor SR: assessors	– all SS on task	rushed due to lack of time. Quick decisions but not much formative discussion of next steps.
PM – peer marking SM – self marking SC – success criteria THUD – thumbs up/ down LPs – learning partners Other codes in this column: IWB – interactive whiteboard				R <n>– observable evidence of children referring to AfL techniques while completing tasks, e.g. success criteria are on the board PER – positive emotional reaction NER – negative emotional reaction Other codes in this column: <n> - a number UL – unsolicited liaison with peers related to the task at hand[33]		E: <n> – number of students who seem engaged with the task DE : <n> – number of students who seem disengaged with the task	Interaction patterns key: T – C: teacher – whole class T – 1L: teacher – learner T – xL: teacher – small group of learners L– L: learner- learner L –xL: learner – small group LL small group L– learner T – teacher C – whole class IND – independent work Other codes in this column: TR: - teacher role LR: - students role		

---

[33] UL – reflects what teachers mentioned in interview: increased cooperation between children.

Appendix 15: The transcribing convention

The transcribing convention was adapted from Walsh (2006). In order to represent the conversations that occurred and limit the impact of the process of transcribing the spoken conversations, two main considerations that guided the transcriptions are:

1. No corrections were made to the language;
2. Standard conventions of punctuation were not used;
3. If the transcriber was unable to understand what was being said, that fragment of the conversation is marked *unintelligible*.

Codes used in transcribing:

T – teacher

L – learner (not identified)

L1: L2: etc.,   identified learner

LL – several learners at once or the whole class

/ok/ok/ok/ - overlapping or simultaneous utterances by more than one learner

[do you understand?]   } overlap between teacher and learner

[I see]

= turn continues, or one turn follows another without any pause

(1)   pause of one second or less marked by number 1 in brackets

(4)   silence; length given in seconds

?   rising intonation – question or other

CORrect   emphatic speech: falling intonation

((4))   unintelligible 4 seconds: a stretch of unintelligible speech with the length given in seconds

Anna, Tomek – capitals are only used for proper nouns

T organizes groups - transcriber's comments (in bold type)

(...) in extracts quoted in the thesis, the ellipsis in brackets is used to signify that a fragment of the original transcript is excluded from the quotation

Appendix 16: A photo of one of the themes on an A3 piece of paper.
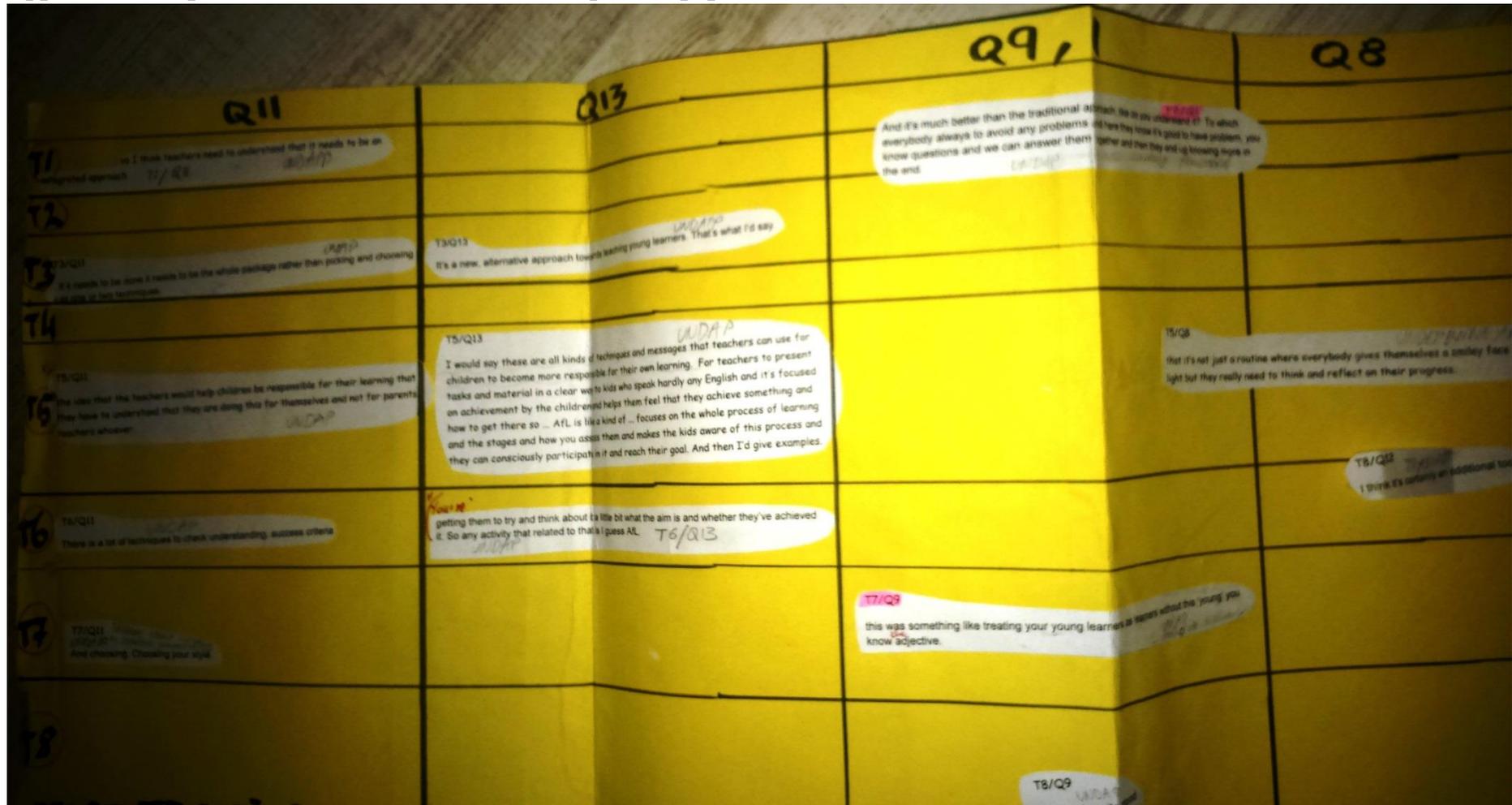
Appendix 17: The number of AfL techniques recorded in ROWDs in all groups by week, in the Autumn Term – demonstrating the lack of use of AfL during the period preceding summative reporting

Group codes	Teacher	The number of AfL techniques recorded in ROWDs each week of the Autumn Term													
		Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13	Week 14
G1	T3	2	3	0	1	0	2	0	0	2	0	0	0	(R) 0	2
G2	T1	2	4	4	4	4	3	3	3	3	3	3	(T) 0	(R) 3	1
G3	T2	0	0	0	0	0	0	0	2	1	0	(T) 0	0	(R) 0	0
G4	T5	2	2	1	3	4	3	3	4	2	3	4	(T) 0	(R) 2	3
G5	T7	1	1	2	1	1	2	2	2	2	1	1	(T) 0	(R) 1	1
G6	T4	0	0	0	0	1	2	1	0	1	0	(T) 0	0	(R) 0	0
G7	T6	0	0	0	0	0	0	0	3	2	4	1	(T) 0	(R) 0	0
G8	T8	0	0	0	1	0	1	0	0	0	0	(T) 0	0	(R) 0	0
G9	T2	0	0	0	0	0	0	0	1	0	0	(T) 0	0	(R) 0	0
G10	T3	0	3	0	0	3	4	0	0	3	0	3	(T) 0	(R) 3	0
G11	T1	3	2	3	2	2	3	3	3	3	3	3	(T) 0	(R) 3	3
G12	T8	0	0	0	1	0	0	0	0	0	0	(T) 0	0	(R) 0	0
G13	T5	2	3	4	4	2	2	3	3	3	4	3	(T) 0	(R) 2	2
G14	T6	0	0	0	0	0	0	2	3	3	1	(T) 0	0	(R) 0	0
Codes for other assessment-related events		Reports sent home – (R) Test – (T)													

319

Appendix 18: Types of AfL techniques

This appendix explains how each AfL technique was deployed by the teachers in the study. Examples are included to illustrate the implementation. The techniques are discussed in the same order as in Table 4.2 in Chapter 4 and in Appendix 20:

1) Success Criteria (SC)
2) Learning Partners (LP)
3) What are we learning today? - type questions (WALT)
4) Traffic Lights (TL)
5) Thumbs up or down (THUD)
6) Two stars and a wish (TSAW)
7) Sharing a good and a bad model (SGBM)
8) Smiley faces (SF)
9) I can statements (ICS)
10) Perfect Purple and Red to Remember (PPRR)
11) Next steps (NST)
12) Colour coding (CC)
13) Find the Fib (FTF)
14) Increased thinking time (ITT)
15) Star charts (SCH)
16) Indicate mistakes without explanations (IMWE)
17) Sheriff's star (SST)
18) Mind maps (MM)

1.Success Criteria (SC)

Types of tasks and skills used with: writing, arts and crafts, classroom instructions, speaking

Timing of use: before and during a task

Description: SC were either presented by the teacher or elicited from the students. They were often recorded as a short list which specified what was required from each student in order for their performance to be judged as successful. They specified either a sequence of steps to follow (i.e. process of how to perform the task) or qualitatively as a list of desired ingredients to identify what students' output should contain. During the task, learners were sometimes referred to the success criteria to remind them that they should

monitor their own performance. After the task, SC were often referred to when providing feedback on performance. SC could be elicited from children by using the SGBM technique.

Example 1: Written Success Criteria from different age groups

Success criteria for writing task from a class of 7-9 year olds
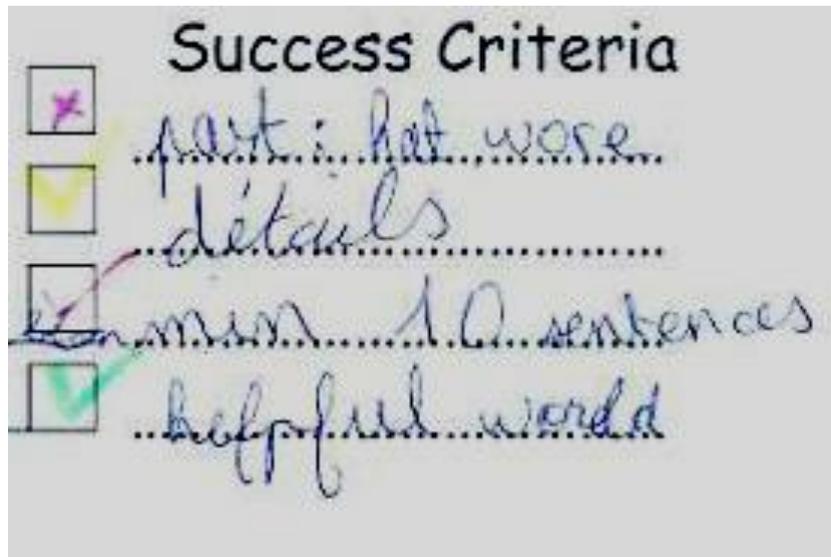
Success criteria for writing task from a class of 10-11 year olds

Success Criteria (in the photo above)

Use a title

Use present perfect in at least two sentences

Answer at least three of the questions from page 85

Join sentences together

Use the vocabulary from page 85 ex 2b

Example 2: A teacher eliciting SC for a speaking task from a group of 7-9 year olds

2.Learning Partners (LP)

Types of tasks and skills used with: speaking; vocabulary learning

Timing of use: throughout lessons and tasks; often in preparation for learners to work independently

Description: Learning Partners were pairs of small groups of students whose role was to provide mutual support in the learning process. Teachers either assigned partners or allowed learners to choose their own learning partner. Partners could change periodically, e.g. every lesson or every month. Learning Partners can be assigned randomly, for example by drawing names from a hat, or purposefully, e.g. based on friendship groups or proficiency levels. Learning Partners were often used when teachers implemented peer-marking or pair work.

Example: Learners aged 10-11 providing peer feedback to their learning partner

3. What are we learning today? - type questions (WALT)

Types of tasks and skills used with: learning objectives for the lesson

Timing of use: at the beginning of a lesson; often referred to throughout the lesson

Description: *WALT* stands for *What are We Learning Today*. Some teachers used funny characters (often animals) called *WALT* and, at the beginning of the lesson, asked children a question that signalled what the children would be learning in that session. The questions could start with 'Can you..?' as in, *'Can you tell me six things which you did yesterday?'* These served as initial prompts for establishing how students understood what their learning objective for the session was. Teachers often came back to these questions at the end of the lesson in order to raise the children's awareness of what they had achieved in that session and what needed to be improved. Although the WALT questions seemed to be *Yes/No* question, they are in fact much more open in nature because they serve as prompts for children to try and demonstrate their skills within the area defined by *WALT*. This was an opportunity for teachers to gather information about what to focus their efforts on in the lesson in order to ensure that by the end of that lesson, the learners could confidently answer 'yes' to the WALT question and demonstrate the skills that proved their answer to be correct. In other words, using *WALT* provided an opportunity for teachers to identify the gap between pupils' current levels of skill and the learning objective.

Example: WALT the Frog on the screen, before a lesson in a class of 7-9 year olds

WALT

4.Traffic Lights (TL)

Types of tasks and skills used with: speaking; vocabulary learning

Timing of use: at the end of a lesson or a task; sometimes followed directly by Next Steps (NST)

Description: The principle of Traffic Lights is that children or teachers indicate the perceived degree of learners' achieving their learning objective with one of the three colours: red, amber or green. The colours are an analogy of traffic lights used for motor traffic. Red means – 'not achieved' but was often positively phrases, e.g. 'I will achieve it in the next lesson' or 'I need to work on this', 'amber' means 'partly achieved' often phrased as something like 'I can do it with some help' or similar and 'green' – 'objective achieved', often phrased in more accessible way as 'I can do it by myself' or 'I can do it well'. Traffic Lights could be recorded by colouring in one circle on a printed template or by learners' holding up pieces of paper in a given colour to show the teacher how they feel about their own achievement of the learning objective.

Examples: A Traffic Lights template with a learner's self-reflection

I can ask someone what they got for their birthday.

Green = I know this already.
Yellow = I know something about this.
Red = I don't know this yet.

5.Thumbs up or down (THUD)

Types of tasks and skills used with: classroom instructions; arts and crafts

Timing of use: throughout lessons; often during and after giving instructions;

Description: This technique requires all learners simultaneously to react by signalling with their thumbs up ('Yes'/'Correct'/'I agree'), thumbs in the middle ('Not sure'/'Don't know'/'Almost correct') or thumbs down ('No'/'Incorrect'/I disagree'). This technique was used: 1) to provide peer feedback; 2) for learners to indicate how well in their opinion another learner had performed; 3) to express agreement and disagreement; or 4) to indicate if the children understood the instructions that they had been given for a task.

Example: teachers' descriptions of how they implement THUD technique

'I also use thumbs, thumbs up for good things, and feeling confident and moving thumbs around for different levels of confidence' (T1, INTERVIEW)

'We use thumbs up, thumbs down, and across to indicate if they know what they are about to do.' (T7, INTERVIEW)

6.Two stars and a wish (TSAW)

**Types of tasks and skills used with:** writing

Timing of use: after an activity; often when Success Criteria were used before the activity

Description: When using this technique, teachers indicate two positive things about a piece of child's writing and identified one area for development. The observed lessons suggested that TSAW were often used to reflect criteria for success. This meant that

teachers indicated two criteria which a learner met and one which was not met or only partially met.

In the example below, the teacher indicated the correct use of 'There is" with singular nouns and the fact that the learner managed to write about five photos. The area for development is to use 'There are' when talking about '2+' (two or more objects), namely with plural nouns.

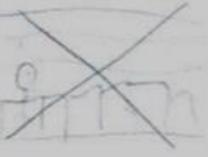**Example 1:** Two stars and a wish in 10-11 year group, marked by a teacher
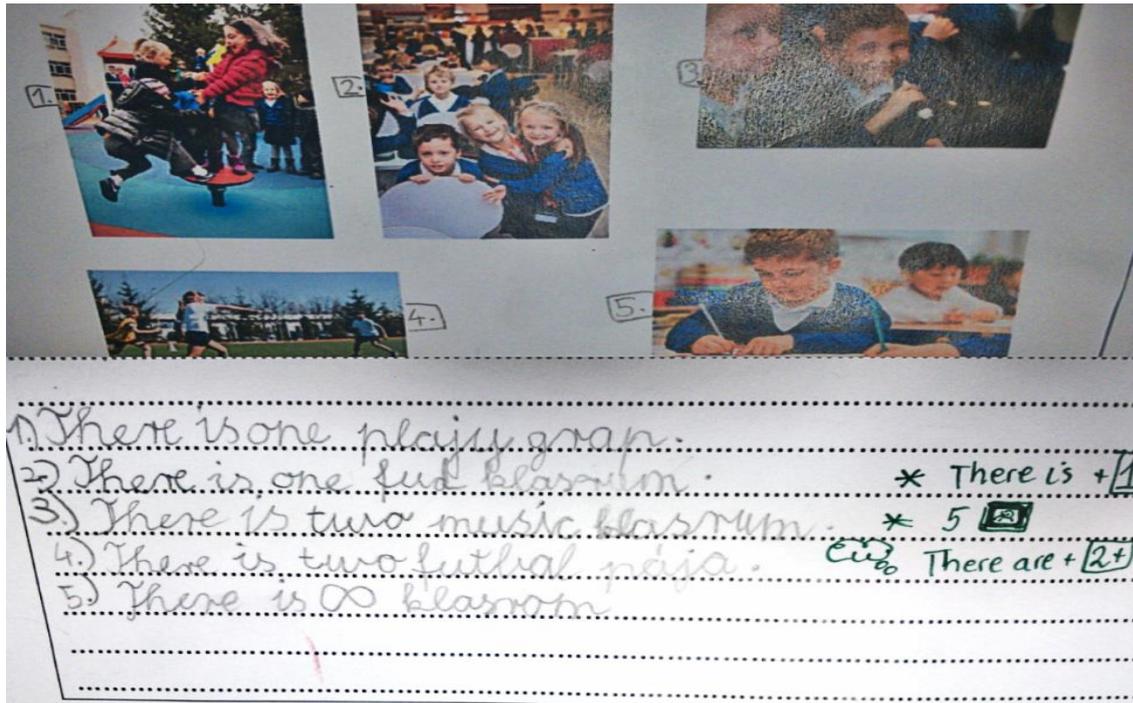
Example 2: Two stars and a wish in 10-11 year olds class, marked by a learner.

The pink comments read under the story template in the scan are:

* Eight good sentences,

* You used always, usually,

! Remember: I always go home. NOT: I go home always.

7.Sharing a good and a bad model (SGBM)

Types of tasks and skills used with: writing

Timing of use: At the beginning of the lesson; sometimes with reference to *I can statements* (ICS) or *WALT-type questions*

Description: This technique focuses on sharing two models of a completed task like the one that learners are about to embark on. Usually one good model is contrasted with a worse one to illustrate what makes a good performance. Teachers can elicit from the children what makes the exemplary piece of work better, hence arriving at the list of criteria.

Example: A teacher discussing two models of a completed newspaper report with a class of 7-9 year olds and supporting their understanding with pictures
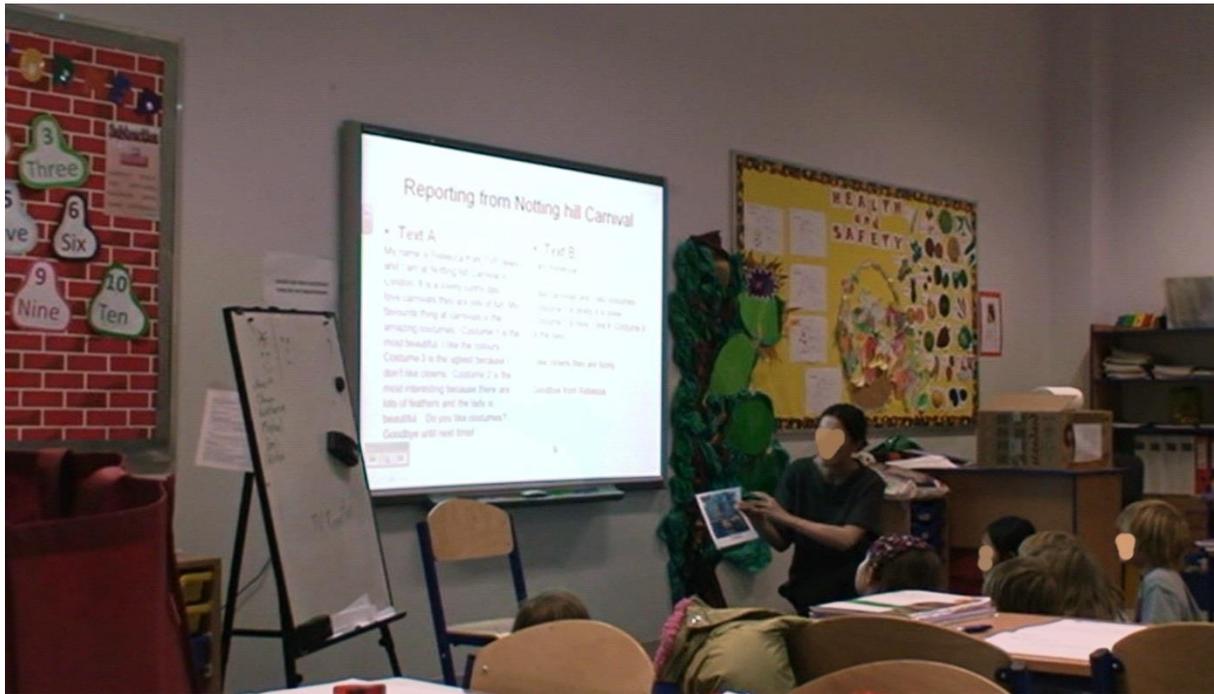
8.Smiley faces (SF)

Types of tasks and skills used with: writing; speaking

Timing of use: after a task; at the end of the lesson

Description: This technique serves the purpose of indicating how well a student has performed on a given task and provides a brief and quick way of gauging students' perceptions of their own achievement. Students are given a template of a face with eyes but no mouth and are asked to draw a mouth that would represent how well they think they have performed on the given task. Variations can include either a) one of the three types of faces: happy, sad or straight or b) a face that represents the emotions that the student feels about their progress in that lesson, e.g. excited, pleased, worried etc. Smiley Faces were often used in conjunction with other techniques, such as *I can statements* or *Success Criteria*.

Example 1 (below): Smiley Faces recorded next to *I can statements* from five lessons at the back of a learner's notebook

I can *remember* *of places in the town*

I can *talk about diferences:*

I can *write about my hero*

I can *describe people:*

I can sing the beach song.

Example 2 (below): Two learners discussing which smiley face to assign to an *I can statement*

9.I can statements (ICS)

Types of tasks and skills used with: writing; speaking

Timing of use: after a task; at the end of the lesson

Description: *I can statements* were often used in conjunction with a technique which demonstrated that learners could demonstrate the given skill. It could be with Smiley faces, which illustrated how confident students felt about their own skills; with Colour Coding, which highlighted evidence of success in learners' work; or with Traffic Lights, which demonstrated to what degree (fully, partially, not at all) learners could demonstrate the given skill. Once a task had been completed, the learners were asked to reflect or peer assess whether they can do something.

Example: Examples of ICSs are illustrated in Example 1 in the description of Smiley Faces above. *I can statements* were also sometimes used on their own and not accompanied by any other technique.
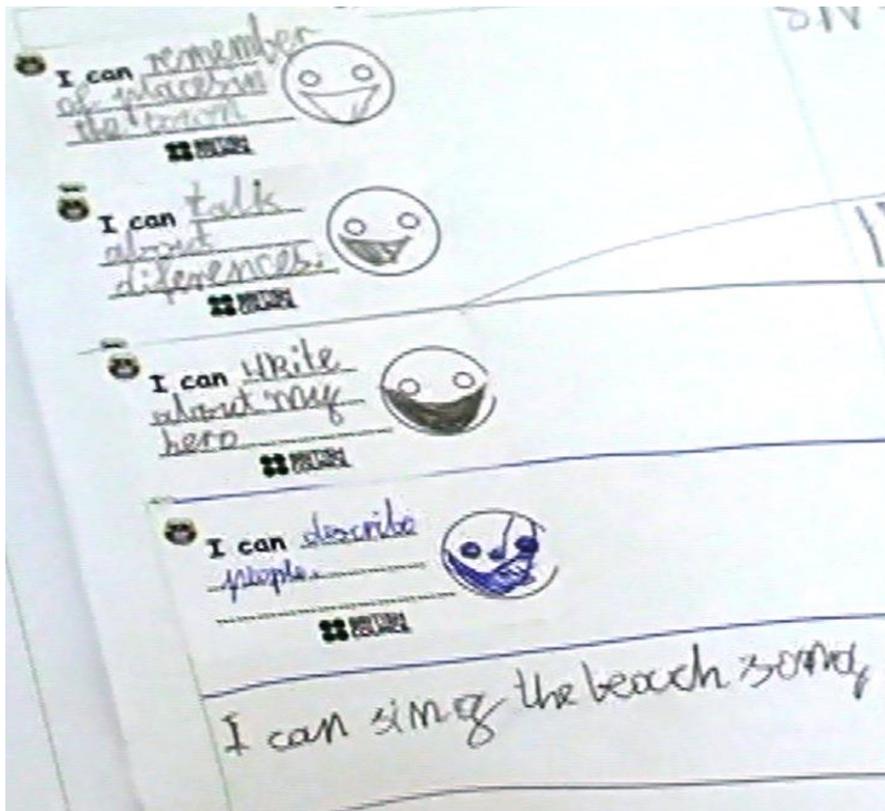
10.Perfect Purple and Red to Remember (PPRR)

Types of tasks and skills used with: writing; speaking

Timing of use: after a task; at the end of the lesson

Description: Two different colours were used, e.g. red and purple. They were given alliterated names which made them easier to remember and reinforce children's understanding of the colours' significance. These could be Red to Remember (or Reflection Red) and Perfect Purple (or Green to Go). Here red would mean *'There is a problem here and you need to fix this. What needs changing? What should you do differently next time?'* and purple or green would mean *'This is where you showed that you can do well what you have been asked to do and you can move on to learning new skills now'*. In the writing tasks, the colours were used similarly to circle pieces of the students' written work when they were demonstrating the skills that the tasks was designed to practice: purple or green for indicating good work and red for indicating where improvement was necessary. The good work as given a record of success while the students' whose work required improving had to reflect on what exactly needed to be improved. Other colour codes used by the teachers who participated in this research included 'Improvement Indigo', 'Perfect Purple', 'Oh-Dear Orange' and 'Brilliant Blue'. In speaking tasks, the technique was enacted by showing the appropriate code on a

laminated piece of paper especially when a child made a mistake in the target language and needed to reflect on it.

Example: A photo of Perfect Purple in a child's notebook
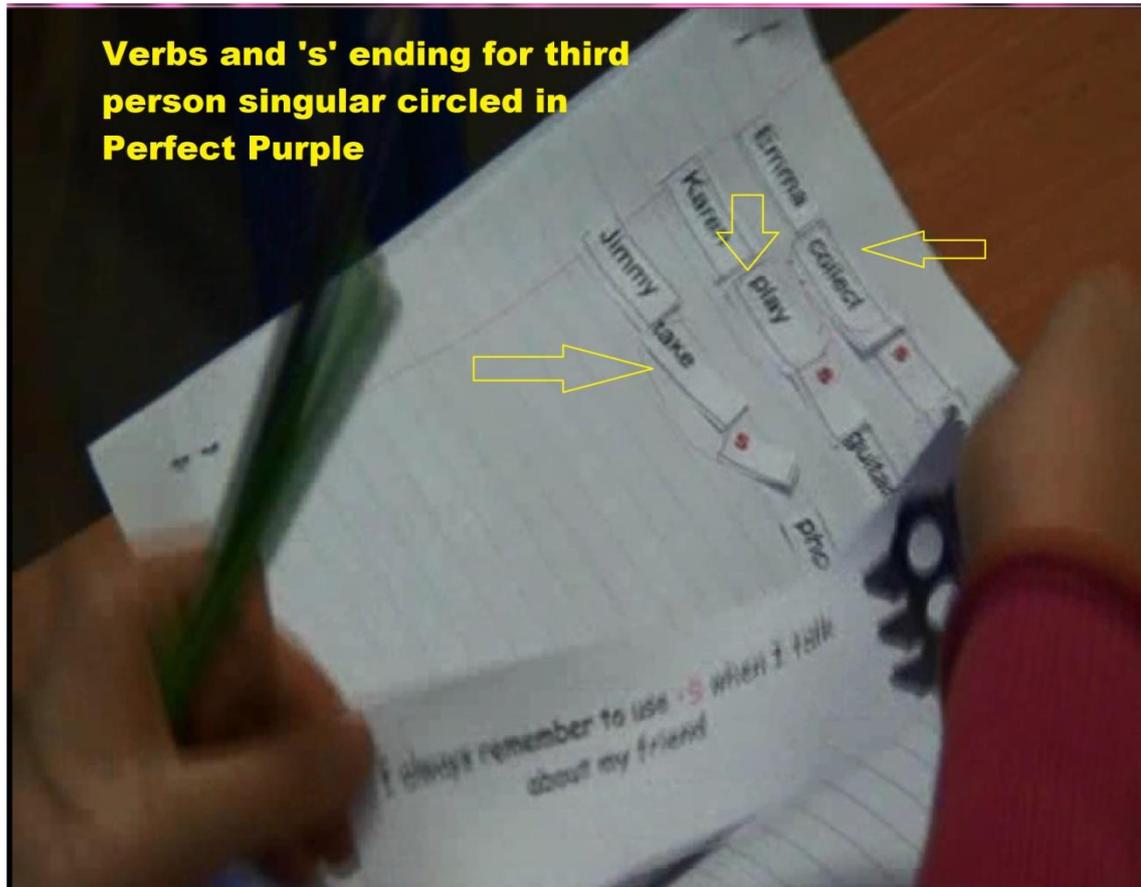
11.Next steps (NST)

Types of tasks and skills used with: writing; vocabulary learning; grammar

Timing of use: after a different AfL technique: most commonly after Traffic Lights and Perfect Purple and Red to Remember

Description: This technique was used as a means of closing the gap between current and desired performance. It could be used on its own or in conjunction with other techniques such as Traffic Lights or Success Criteria. The provision of an additional mini-activity to close the gap between the current and expected levels of performance is an important characteristic of the Next Steps technique.

Example: In the example presented in the following section (Colour Coding, 10), the teacher would provide a Next Step in the form of a piece of advice as to what the student should think about the next time with a short activity to ensure that the learner could understand the Next Step and be encouraged to remember the correct grammar form. The

Next Step comment reads: 'Next time remember to use past. Which is past? Circle it. had or has'.

12.Colour coding (CC)

Types of tasks and skills used with: writing

Timing of use: after an activity had been completed to be revisited by a student; often followed by Next Steps or combined with Success Criteria; often recorded in notebooks.

Description: Colour coding might also be used when a child's work was marked against success criteria (for description of success criteria, see Point One of this appendix). The feedback provider i.e. the child themselves, a peer or the teacher would assign a colour to each criterion and then look for fragments of the text that demonstrated that the criterion was fulfilled. When such an extract was identified, it was highlighted with a circle in the colour assigned to that criterion. If there were a certain colour lacking or not enough of it to demonstrate that the learner had met the given criterion for success, that child could be asked to perform a short task following the main task to upgrade their work (i.e. Next Step).

Example: A character description, written by an 8 year old with colour coding (CC) in response to the success Criteria (SC) presented in technique Number 1 and Next Step (NST) written by the teacher.
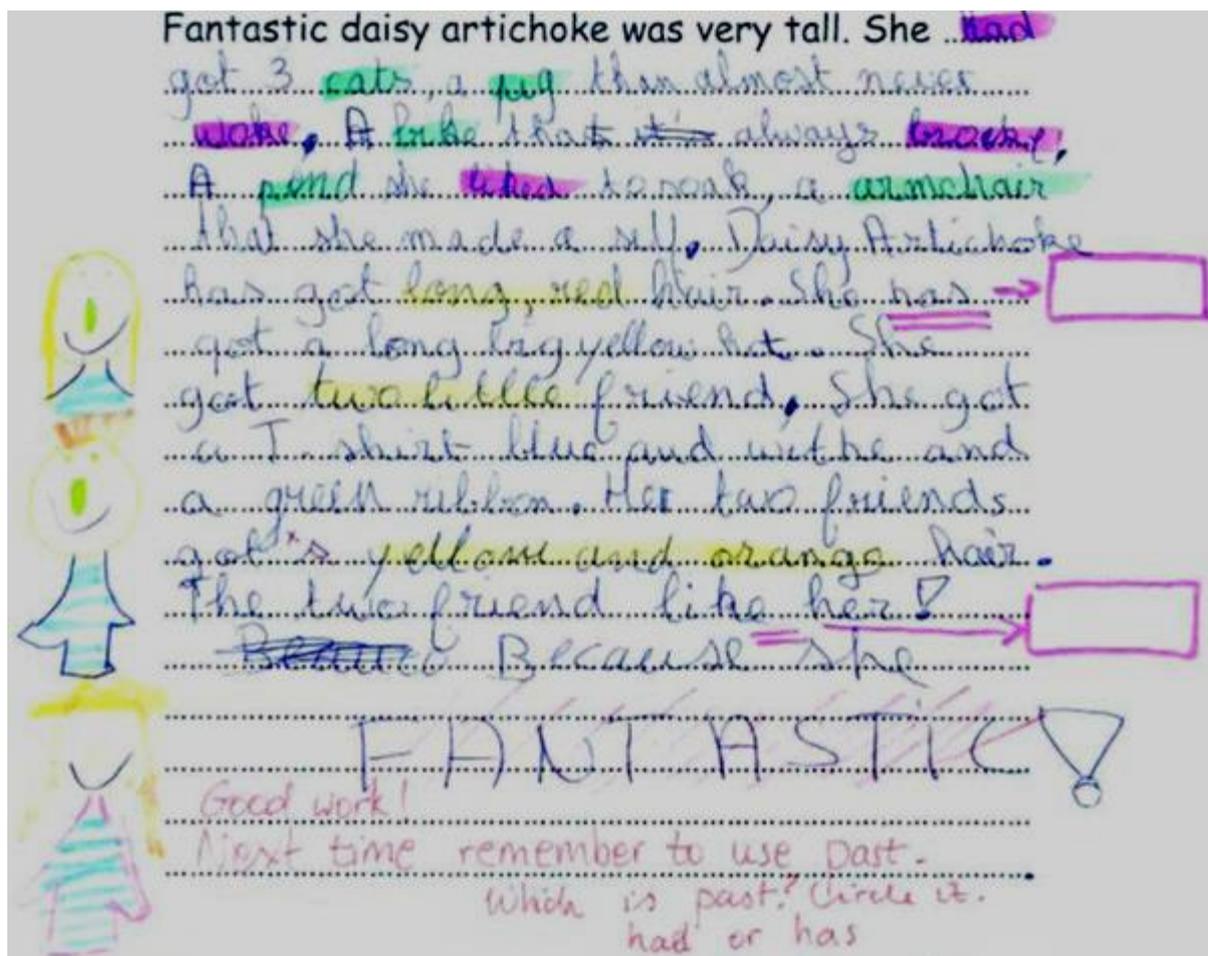
13. Find the Fib (FTF)

Types of tasks and skills used with: speaking; vocabulary learning; grammar

Timing of use: after a new grammar rule or set of vocabulary had been introduced and practiced.

Description: This technique would be enacted in the form of a short activity in which students would be given a few a possible solutions or answers and would have to find one that was 'the fib': i.e. not acceptable. The aim of that technique was to identify gaps in understanding and rectify them.

14. Increased thinking time (ITT)

Types of tasks and skills used with: speaking; reading comprehension tasks

Timing of use: throughout lessons or activities; often after asking a question

Description: The focus of this technique is to allow learners time to think, in silence. This time would be given after asking a question and before accepting any answers or after giving instructions and before allowing learners to start the task. Teachers often

explicitly informed the students that there would be thinking time and requested that nobody should put their hand up or offer any answers but everyone should try to think about the answer.

15. Star charts (SCH)

Types of tasks and skills used with: learning objectives for the lesson

Timing of use: at the end of a lesson

Description: A Start Chart is a type of a poster that the teacher could display in the classroom and which contained the aim(s) for the lesson(s). Learners who could demonstrate that they had met that learning objective could stick a star sticker next to their name. Objectives could be added to the poster to build up a picture of achievement across the semester and to indicate areas that required further work for individual students.

16. Indicating mistakes without explanations (IMWE)

Types of tasks and skills used with: grammar; spelling; writing; reading comprehension tasks

Timing of use: recorded in notebooks; while a task is being completed; often with reference to *I Can Statement* or *Success Criteria*

Description: This technique aims to indicate a mistake by circling or underlining it. When using this technique, teachers would not provide explanations of the mistake or corrections. This had to be discovered by the student themselves.

Example 1: In the example of writing presented under technique Number 10 (Colour Coding), there are two instances in which verbs in the present tense are underlined and space is provided to write verbs in the past tense. The task that the learner was given was to write a description of a book character based on pictures from the previous weekend. Hence, all the writing should have been done in the past tense. This was also clarified through the Success Criteria. The IMWE AfL technique was used to draw the learner's attention to the wrong form of the verb and encourage them to reflect on what the correct form was.

Example 2: Another example is provided below. The teacher simply circled a syllable and annotated it with a question 'What's wrong here?' The learner needed to notice the spelling mistake and correct it.

17.Sheriff's star (SST)

Types of tasks and skills used with: the learning objective for a lesson

Timing of use: at the end of the lesson or an activity

Description: This technique entailed the use of physical stars which looked like a sheriff's badge. The students would receive a badge each when they arrived at the class. They could take the star off when they were feeling lost and 'not in charge of their learning'. When the teacher noticed that a child took a star off, (s)he would know that the child needed individual support during the lesson. This was a safe way for children to indicate that they were having difficulties without making it public to the whole group of learners.

18.Mind maps (MM)

Types of tasks and skills used with: projects

Timing of use: at the beginning of a project or a lesson; referred to at the end

Description: Mind maps are used as an assessment technique by producing a mind map at the beginning of a project or unit of work and then referring to it at the end. A map could be created with two colours; one meant – things that we already know about this topic; and the other – questions that we have about it. As the work progressed, learners would use a third colour to add extra information and new vocabulary that they had learnt or any additional questions that they might have about the topic. This third colour demonstrates visually the amount of learning that would be taking place. Mind maps would be done in groups and displayed in the classrooms.

Appendix 19: Extended version of Extract 4.1 from the Focus Group discussion.

EXTRACT 4.1

(Focus group discussion on what AfL is, recorded in May 2012)

[1]   T3: *the whole thing is about that getting them to understand what to do and*

[2]   *then to figure out how well they have done it and it doesn't matter which*

[3]   *techniques you use, right?*

[4]   T5: *I see what you're saying but I also think that they need to know what they*

[5]   *need to improve you know* (1) *like they need to know why they are getting an*

[6]   *amber light, yes?*

[7]   T3: *yeah ok* (1) *yes* (1) *yes* (1) *that too*

[8]   T1: *it does perhaps depend on the groups or how you present it but I think*

[9]   *my students would get a bit bored if it was just success criteria and traffic*

[10]   *light I feel like they are more with me if I vary it a bit I mean I keep the*

[11]   *same focus as you said it is about sharing the purpose of this lesson and*

[12]   *then if they achieved the goal and maybe getting them to think how they can*

[13]   *get better but I feel that it needs variety*

[14]   T8: *I'll support that actually both of you I think (1) some variety is needed*

[15]   *but in fact it's the purposes that make it all meaningful and worthwhile*

[16]   T2: *so I would say that AfL is a kind of philosophy that involves measuring*

[17]   *your students' progress minute-by-minute of the lesson on the ongoing basis*

[18]   *and checking that they have the understanding of something and it also*

[19]   *involves them knowing what to do in order to achieve goals*

[20]   T5: *yeah I guess*

Appendix 20: The scoring process for organising AfL techniques in order of most-least commonly used

Name of AfL technique	Tech-nique Code	Points from ROWDs analysis	Points from lesson observations field notes analysis	Total number of points	When was it used within a lesson?	Types of activities and skills used with
Success Criteria	SC	18	16	34	Before and during a task	Writing, Arts and crafts, Classroom instructions
Learning Partners	LP	14	18	32	Throughout lessons and tasks Often in preparation for learners to work independently	Speaking Vocabulary learning
What are we learning today? - type questions	WALT	17	14	31	At the beginning of a lesson Often referred to throughout the lesson	Learning objectives for the lesson
Traffic Lights	TL	15	12	27	At the end of the lesson or a task Sometimes followed directly by NST	Speaking Vocabulary learning
Thumbs up or down	THUD	10	17	27	Throughout lessons Often during and after giving instructions;	Classroom instructions Arts and crafts
Two stars and a wish	TSAW	16	11	27	After an activity Often when SC were used before the activity	Writing
Sharing good and bad model	SGBM	12	15	27	At the beginning of the lesson Sometimes with reference to ICS or WALT	Writing
Smiley faces	SF	11	13	24	After a task At the end of the lesson	Writing Speaking
I can statements	ICS	13	8	21	At the end of the lesson	Learning objectives for the lesson
Perfect Purple and red to Remember	PPRR	9	5	14	After an activity has been completed to be revisited by a student Often followed by NST Often recorded in notebooks	Grammar Writing (often with grammar focus)

Next steps	NST	7	6	13	Often after a different AfL technique Most commonly after TL and PPRR	Writing Vocabulary learning Grammar
Colour coding	CC	6	7	13	After an activity has been completed to be revisited by a student Often followed by NST or combined with SC Often recorded in notebooks	Writing
Find the Fib	FTF	8	3	11	After a new grammar rule or set of vocabulary has been introduced and practiced	Speaking Vocabulary Grammar
Increased thinking time	ITT	0	10	10	Throughout lessons or activities Often after asking a question	Speaking Reading comprehension
Star charts	SCH	5	4	9	At the end of a lesson	Learning objectives for the lesson
Indicate mistakes without explanations	IMWE	0	9	9	Recorded in notebooks While a task is being completed Often with reference to ICS or SC	Grammar Spelling Writing Reading comprehension tasks
Sheriff's star	SS	4	0	4	At the end of the lesson or an activity	Learning objectives for the lesson
Mind maps	MM	0	2	2	At the beginning of a project or lesson Referred to at the end	Projects

Appendix 21: Inter – lesson frequency of using AfL by individual teachers

No. Of AfL techniques	T1			T2		
	Lesson frequency	Percent	Cumulative Percent	Lesson frequency	Percent	Cumulative Percent
0	6	10.7	10.7	46	82.1	82.1
1	1	1.8	12.5	8	14.3	96.4
2	11	19.6	32.1	2	3.6	100.0
3	25	44.6	76.8			
4	13	23.2	100.0			
Total	56	100.0		56	100.0	

No. Of AfL techniques	T3			T4		
	Lesson frequency	Percent	Cumulative Percent	Lesson frequency	Percent	Cumulative Percent
0	35	62.5	62.5	43	76.8	76.8
1	5	8.9	71.4	9	16.1	92.9
2	16	28.6	100.0	4	7.1	100.0
3						
4						
Total	56	100.0		56	100.0	

No. Of AfL techniques	T5			T6		
	Lesson frequency	Percent	Cumulative Percent	Lesson frequency	Percent	Cumulative Percent
0	5	8.9	8.9	39	69.6	69.6
1	1	1.8	10.7	1	1.8	71.4
2	10	17.9	28.6	8	14.3	85.7
3	34	60.7	89.3	8	14.3	100.0
4	6	10.7	100.0			
Total	56	100.0		56	100.0	

No. Of AfL techniques	T7			T8		
	Lesson frequency	Percent	Cumulative Percent	Lesson frequency	Percent	Cumulative Percent
0	28	50.0	50.0	53	94.6	94.6
1	21	37.5	87.5	3	5.4	100.0
2	7	12.5	100.0			
3						
4						
Total	56	100.0		56	100.0	

Appendix 22: Professional development activities which teachers participated in May 2012 - October 2013

Teacher Code	Time devoted to training about assessment	Time devoted to training about AfL	The time spent observing colleagues teach	Time spent observing AfL in lessons
T1	1hr	0 hrs	2 hrs	1hr
T2	2 hrs	2 hrs	1.5 hrs	1.5 hrs
T3	2 hrs	0 hrs	2 hrs	0 hrs
T4	8 hrs	4 hrs	10 hrs	4 hrs
T5	1 hr	0 hrs	15 hrs	3 hrs
T6	14 hrs	0 hrs	4 hrs	0 hrs
T7	2 hrs	0 hrs	5 hrs	2 hrs
T8	1 hr	0 hrs	3 hrs	0 hrs

Appendix 23: Applying the Stroch (2002) model to interactions which occurred during the use of AfL

EXTRACT No.	TURN	TRANSCRIPT
4.3 T-1L 7-9 year olds	[1] [2] [3] [4] [5] [6] [7]	T1: what else makes it good (1) L14: the text A is bigger= T1: = yes it's bigger (1) what do you mean by bigger (1) there are lots of (1) L14: words= T1: =there are longer (1) L14: sentences T1: sentences (1) T underlines a sentence on the board
equality – low to medium (T1 initiated and led the conversation) mutuality – medium to high (student's suggestions were acknowledged by the teacher, e.g. in turn 3, L14 reacted appropriately to T1's requests for clarification, providing alternative answer when in Turn 6, the answer from Turn 4 did not seem satisfactory for the teacher) Outcome: Q4 expert/novice		
4.4 T-1L 7-9 year old	[1] [2] [3] [4] [5] [6] [7] [8] [9]	T5: ok (1) which one are you going to give yourself? L13: what? T5: you think you're green (1) yellow or red? L13: looking at the teacher for 2s T5: do you know eight words? (1) do you know eights words [NAME]? L13: what? T5: do you know eight of these expressions? do you know EIGHT? yes? L13: yes T5: ok then (1) do green light (1) that's good
equality – low (T5 initiated and controlled this interaction) mutuality – medium to low (T5 offered guiding questions but seemed to fail to allow time for L13 to reflect on them, L13 did not contribute a suggestion, a question or an independent answer) Outcome: Q3 dominant/passive		
4.5 T-1L 10-11 year olds	[1] [2] [3] [4]	L11: how is throwing? T7: give me the correct question (2) L11: how (1) do we spell (1) the word throwing T7: excellent stuff T writes 'throwing' on the whiteboard
equality – medium to low (T7 controlled the conversation by requesting a correct question before answering the initial question, T7 knew and finally provided the correct answer) mutuality – medium to high (L11 initiated conversation, responded to the teacher's request, T7 and L11 both had their requests met by the interlocutor) Outcome: Q4 expert/novice		

4.6 T-1L 10-11 year olds	[1] [2] [3] [4] [5] [6] [7] [8]	L12: too MUCH fans? T6: yes (1) too many fans? (1) yes (1) L12: too many (1) of much? T6: because 1 fan (1) 2? L12: fans T6: so that means that much or many? L12: many T6: many fans (1) so 1 fan (1) 2 fans (1) many fans
equality – medium to low (T6 controlled the conversation by guiding the student towards revising a grammar rule) mutuality – medium to high (L12 initiated the conversation and asked for confirmation in Turn 2, T6 and L12 both had their requests met by the interlocutor) Outcome: Q4 expert/novice
4.7 L-L 7-9 year olds	[1] [2] [3] [4] [5] [6] [7]	L9: cinema L10: let's go to see Harry Potter at half past eight am L9: am?(1) or pm?(1) L10: evening? L9: yeah pm L10: and we will meet at half past eight pm= L9: =pm
equality – medium to high (both learners participated actively); mutuality – medium to high (L9's initial suggestion was accepted and developed (time proposed) by L10, L9 and L10 engaged in some negotiation of meaning of phrases am and pm to finally reach an agreement). Outcome: Q1, collaborative
4.8 L-L 7-9 year olds	[1] [2] [3] [4]	L7: can we put on it a black? L8: it's green for christmas tree (1) L7: hmm (2) L8: I think you know it green Both Ls reach for green colouring pencils
equality – medium to high (both learners participated actively); mutuality – medium to high (L8 responds to the suggestion made by L7 with a counter suggestion, L7 acknowledged the counter suggestion somewhat hesitantly, to which L8 responded by reinforcing it). Outcome: Q1, collaborative
4.9 L-L 10-11 year olds	[1] [2] [3] [4] [5] [6] [7] [8] [9]	L1: has (1) he (1) tidied the kitchen yet? L2: yes he has L2 draws a tick on the corresponding picture (2) have he taken the rubbish out yet? (1) L1: has he (1) L2: ok (1) he (2) has (2) L2 draws a tick on the corresponding picture L1: no (1) but has he taken the rubbish yet (1) L2: no (1) this now (1) L2 points to a different picture L1: but this one is has he taken the rubbish yet L2: yes he has (1) now you
Storch's (2002) model: equality – medium to high (both learners participated actively); mutuality – medium to low (corrections by L1 not understood by L2, L1 did not change correction giving strategy to indicate lack of understanding but continued to provide implicit corrections). Outcome: Q2, cooperative

4.10 T-1L 7-9 year olds	[1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14] [15] [16] [17]	T3: This is Stas[34] (1) can he do it (1) can you talk about beach activities using like/love and ing words T points to the board which has WALT written on it (2) the same statement (1) T points to a speech bubble in the activity L15 (reads from the speech bubble): I like to sleeping T3: What do you think (1) is that a happy face (1) medium face (1) or a frown (1) L15: medium= T3: = medium= L15: =medium T3: yeah (1) this is not so good (1) to sleeping (1) do people sleep on the beach (1) L15: shakes head T3: sometimes (1) but what's better (1) L15: sleeping T3: what can he do to get a happy face (3) can he cross something here (1) L15: I like sleeping T3: yes (1) this is better

equality – low to medium (the conversation was initiated and controlled by the teacher; especially evident in turns 10-11 when T3 does not discuss the student's suggestion that people never sleep on the beach but moves the conversation towards the grammatical focus.)
mutuality – medium to high (overall T3 and L15 agreed the judgment of correctness of Stas' sentence and the improvement needed; both interlocutors offered suggestions and responded to one another.)
Outcome: expert/novice

4.11 L-L 10-11 year olds	[1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13] [14]	L3: what time (2) L4: when (2) L3: did you (2) L4: what time when did you (2) L3: nie dobrze (1) [Eng. *not good*] L4: nie no co ty? (1) może być [Eng. *no what are you saying? (1)it's ok*] L3: kiedy ty rano wstałeś? (1) [Eng. *when did you get up in the morning?*] when did you get up in the morning L4: chyba o której rano wstałeś? (1) [Eng. m*aybe what time did you get up in* *the morning?*] L3: what time (2) what time did you get up in the morning= L4: =what time (1) did you co? [Pol. co = Eng. *what*] L4 writes on TSAW template. L3: get up in the morning

equality – medium to high (both learners participated actively);
 mutuality – medium to high (L3 and L4 made suggestions and responded to their interlocutor's suggestion, finally agreement was reached.)
Outcome: Q1, collaborative

---

[34] The name of a fictional character in the picture that T3 was referring to

Appendix 24: Summary of findings on teachers' belief about AfL

Theme 1: AfL implies a degree of integration of assessment with teaching and learning by:
- a)   making learners more aware of what they are learning;
- b)   focusing learners on achieving the learning objectives;
- c)   teachers continuously building AfL opportunities into lessons;
- d)   being a framework for giving meaningful explicit feedback that helps to move learning forward (from teacher, peers and own reflection);
- e)   motivating learners to learn and helping them enjoy the process;
- f)   providing structure and focus to teachers' lesson planning.

Theme 2: AfL is compatible with communicative language classroom methodology but tensions exist between AfL and summative assessment.
- a)   AfL was considered a new name for a concept that was already present in the teachers' TEYL practice.
- b)   AfL was considered easy to implement in communicative TEYL classrooms.
- c)   there existed some tension between AfL and summative assessment in a TEYL context, which seemed to be connected to reporting to parents.

Theme 3: AfL could be implemented in a TEYL context using a number of different AfL techniques.

Theme 4: AfL techniques could be used to serve the purposes of giving and clarifying instructions, sharing aims, criteria for success, feedback and measuring learners' confidence about their learning.

Theme 5: When AfL was used, learners were able to sustain independent and pair work for longer, without the need for support from the teacher, because:
- a)   AfL techniques provided scaffolding.
- b)   Using AfL helped learners become more responsible for own learning.

Theme 6: Using AfL created opportunities for a larger number of one-to-one interactions between teachers, learners and peers.
- a)   When students were able to work without a teacher's help, the teacher could spend time on monitoring work more effectively and providing individual support (T-1L interactions).
- b)   Using AfL enabled teachers to introduce more pair work in the lessons (L-L interactions).
- c)   Students collaborated rather than competed when working together (L-L interactions).

Appendix 25: A model of self-regulated learning and the feedback principles that support and develop self-regulation in students reproduced from Nicol and Macfarlane-Dick (2006)
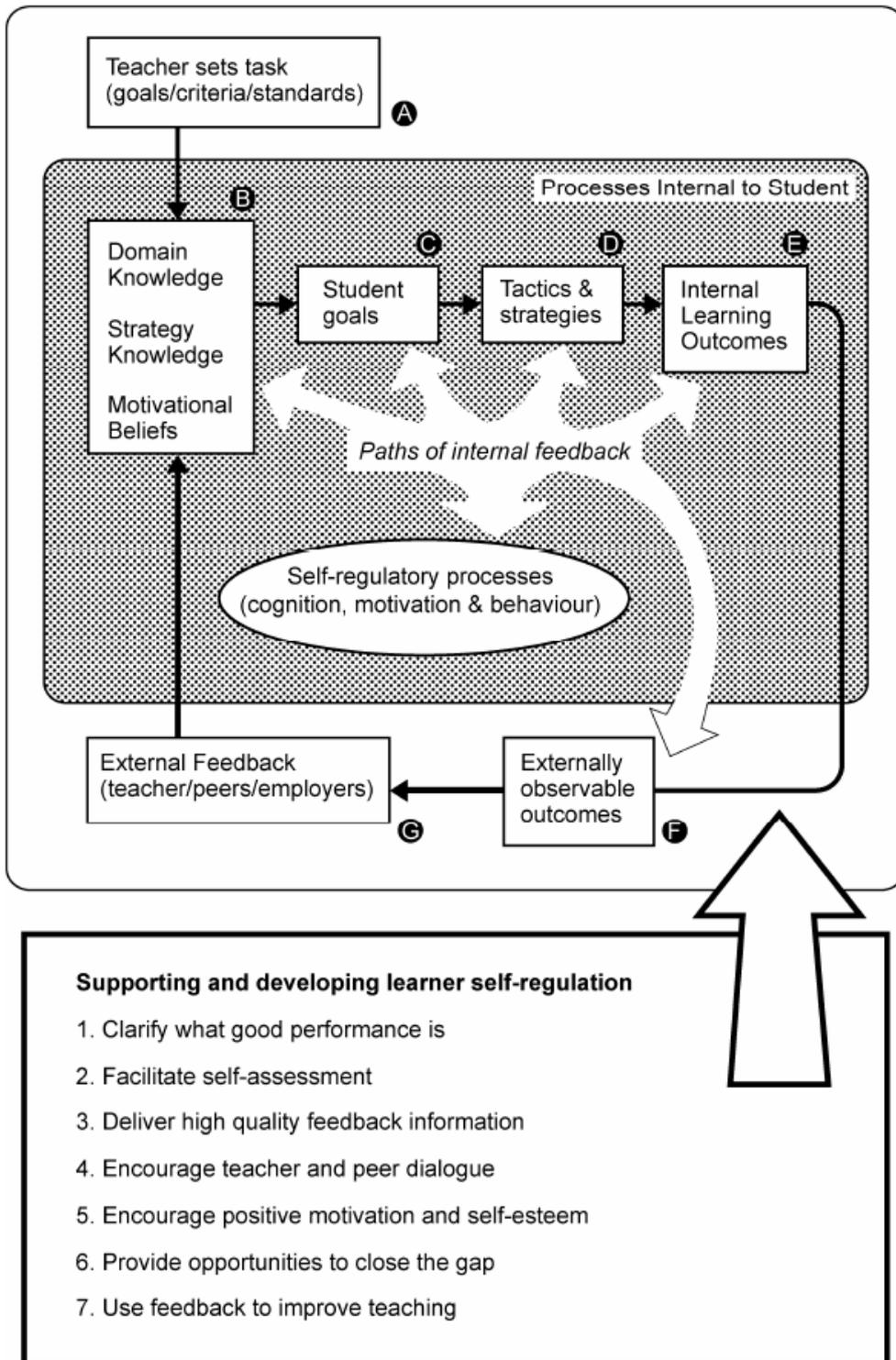

Figure 1. A model of self-regulated learning and the feedback principles that support and develop self-regulation in students.

Appendix 26: Teacher roles in L2 feedback provision (Furneaux et al. 2007)

Initiator:

'Alerts by providing a specified (lexical, grammatical, stylistic, semantic, discoursal, mechanical) or unclassifiable (dotted lines, circle, question mark etc.) alert. The alert may take the form of a question or an explanation, provided there is no actual correction.' (p.76)

Supporter:

'Responds positively to the text with either symbols (++) or comments.' (p.77)

Advisor:

'Identifies areas where the student needs to do further work, either on this particular text or in general. The teacher may offer to help the student work on a problem area. The note is clearly intended as advice.' (p.77)

Suggester:

'The teacher indicates advice by suggesting a better alternative in brackets (where elsewhere, for example, items are crossed out), or writes alternative(s) above the uncorrected original. (This contrasts mainly with Provider behaviour. In the case of Suggester the teacher does not indicate that what the student has written is actually wrong. This category was used only when the teacher clearly contrasted between Provider behaviour and Suggester behaviour and used a different system of correction for each.)' (p.78)

Provider:

'Provides the correct form by substitution, addition, deletion or reordering of an item of language or punctuation. Such corrections do not change the meaning. May be accompanied by an explanation or identification of the problem.' (p.78)

Mutator:

'The teacher alters the text by deleting, adding or rewriting. Such alterations change the meaning.' (p.79)

Appendix 27: Extended version of Extract 4.2 from the Focus Group discussion.

Extract 4.2

(Focus group discussion, recorded in May 2012)

[1]   T2: *with time input from teacher should be smaller =*

[2]   T7: *=mhm=*

[3]   T2: *=because they know their success criteria how to do it or they should be aware*

[4]   *of what is expected but I think at that point monitoring becomes more effective*

[5]   *to make sure that they're actually doing it properly*

[6]   T4: *I think this is a very good point=*

[7]   T1: *=yes more time to monitor better*

[8]   T4: *I think* [T2] *is probably right and when you're doing it you can give them*

[9]   *some individual advice and talk to them when they're stuck on something*

[10]   T1: *what* [T2] *seems to be saying is that when they can get on with a task with*

[11]   *their success criteria etcetera, then you as a teacher have more time to monitor*

[12]   *and help them*

[13]   T2: *yes and how you would do that is an important element of this discussion*

[14]   *when you monitor and how you monitor and what you say to individual students*

Appendix 28: The use of AfL techniques - comparisons between the two age groups

(* The techniques marked with an asterisk were used for providing feedback when monitoring performance)

Purposes for using AfL (reported by teachers in interviews)		AfL techniques observed in both age groups	Techniques observed solely or mostly with 7-9 year olds	Techniques observed solely or mostly with 10-11 year olds
To give and clarify instructions		n/a	Smiley Faces Success Criteria Traffic Lights Thumbs up/down	n/a
To share learning intentions	To ensure that students understand what outcome is expected	Success Criteria Sharing good and bad model	n/a	Mind maps
	To ensure that students know what they are learning	Traffic Lights Colour Coding Sharing good and bad model	What are we learning today?-type questions	Two stars and a wish Mind Maps Perfect Purple/Red to Remember 'I Can' Statements
To provide feedback	To facilitate peer-feedback and teacher feedback	Learning Partners* Traffic Lights Colour Coding	Thumbs up/down Find the Fib Smiley Faces Sheriff's Star Star Charts	Two stars and a wish Increased Thinking Time*
	To facilitate self-assessment and reflection on learning and areas for improvement	Traffic Lights Next Steps	What are we learning today?-type questions* Smiley Faces	Next Steps Indicating Mistakes without Explanations* Increased Thinking Time* Mind Maps Two stars and a wish Perfect Purple/Red to Remember*
To measure students' confidence		Traffic Lights	Thumbs up/down Smiley Faces	n/a
To communicate with parents		n/a	Smiley Faces	n/a
N.B. No techniques were observed to be used for the following purposes reported by teachers: To keep records, To set up homework				

Appendix 29: A coded interview with a teacher

(Codes: Teacher – T7; Researcher - R)

Please note that the transcribing convention is reported in Appendix 15

R: Ok, so the first question is how would you explain what Assessment for Learning is to a teacher that does not know anything about it?

T7: hmm (1) that's a good question (1) training learners to learn and helping them see the advantages of learning helping them enjoy learning motivating them (1) this is what I think it could be' (1) for me when I first started doing it **MOT** in a class then I had to do some research and read about it and it was fun because I had two classes at the same level and in one I used AfL in the other I didn't (1) so that showed me that definitely young learners respond really well to this so I like it (1) and when Tom[35] came to observe my lesson last semester and he was really impressed with how (1) how they can say so much and they do it out of their own will and how did you do it (1) and I didn't do anything special (1) **SCAFF** I was just using the techniques and he saw something after like two months of doing this so it was for me some kind of very good feedback that you know this is working so that's the experience I would share with them and then they need to go away and read about AfL

R: and how would you explain what AfL is to a parent of one of your students?

T7: it's trickier with parents because I'm not sure if they would agree with me if I told them that this is assessment (1) I think parent was tests really and this is a little sad **NCOMPSA**

R: I see so could you tell me how you implement AfL in your lessons with learners aged 7-11?

T7: I used smiley faces for behaviour and classroom management, for showing well they think they have done but also smiley faces for assessing their work so last semester I wanted to get some more information about this so I went on to **SF** the websites somebody suggested and there I observed real schools in England **MESCONF** how kids assess their work and also other people's work so this is what I used (1) WALT to introduce the lesson's aims and you know (1) and then somehow skeleton for the lesson so they knew what's expected so this was just you know (1) maybe learner training or something (1) I think so this might be this (1) and **SHAIM**

---

[35] Pseudonym of a different teacher

350

success criteria this was for completing tasks and also success criteria for finishing lesson to get game time (1) so you know if we got through the success criteria then they had longer game time (1) so they had WALT so this was the aim of the lesson and then on the next flip we had success criteria to complete the lesson you know we had like a plan of the lesson (1) revise vocabulary related to the cinema (1) do an exercise using the vocabulary watch a clip and discuss and prepare a plan for a review (1) so this was like a lesson menu but looked like criteria for success (1) and (1) learning partners of course (1) I used this with primaries mainly to involve more pair work and it sort of helps them get what pair work and there were two or three students who were a bit weaker from the other students so just you know I used the learning partners to you know (1) because nobody wanted to work with them so by doing the learning partners they somehow got on well

SC

GCINT

WALT

SC

SHAIM

LP

PAIRLL

R: thank you for this and could you describe which part of the lesson you use AfL in? Why?

I think you know (1) when they have to do some group work then use it (1) when they have to do writing then of course success criteria WALT at the beginning of the lesson (1) class rules reminding of class rules (1) just you know also at different times of the lesson and we also use thumbs up thumbs down and across to indicate if they know what they are about to do (1) so for example if I give instructions and explain what to do and then I say 'show me if you know what to do' and this is an indicator and if somebody puts this this is ok and if somebody puts this then they are quite confused and I know who to help and this is completely lost (1) this shows me if they understand because for example my instructions might have not been clear so if see that majority have this then I know that the instructions must have been clear but if I see that many people have this or this then this is feedback to me on my instructions

CFS

THUD

GCINS

R: I see and have you adapted the AfL techniques that you learnt about in the training sessions? How? Why?

T7: just the one I mention earlier success criteria for the lesson (1) we discussed what they meant and I think they got interested I think basically the more they are involved in what's going to happen in the lesson the better their behaviour I think they basically like doing things like that (1) this worked better with more

advanced groups of kids and with younger I tried just success criteria for tasks and WALT and it was ok (1) went down well **MOT**

T7: I think that in our reality you know in our situation in Poland parents are very (1) they're very (1) attached to tests (1) they like test because they thing this is the assessment so I think the test could be somehow summary of what they've done something like this (1) AfL it is assessment because you know when students assess their own work so it's assessment yeah (1) so for me it is assessment but you know some parents or some students as well might like to **NCOMSA** see something like this so we should use both

T7: yes definitely I would because it makes (1) in terms of planning at the beginning it is quite time consuming to find things to be preparing this but once you start using it you'll see that the lessons flow much more smoothly (1) I think it's good to observe somebody using this read something like you know (1) have some training but also observe somebody and then think of your own way how to implement it because you know it's good to see your learners and think which of these would be suitable for this group and choosing (1) choosing **LPL** your style

T7: for me it was really something (1) eye openers on using lots of stuff I haven't used (1) I used stuff for classroom management all the techniques but this was something like treating your young learners as learners without this **EASIM** young you know adjective (1) for me it was very good I really enjoyed doing that yeah

T7: hmm (1) to be honest I haven't asked them but I think it would be a good idea to just ask them how they like it (1) do they like using WALT? do they like using this? so I think that I'll ask the maybe at the end of the year (1) but **RESP**

also in the classroom I definitely have fewer problems with behaviour and they are more willing to do things (1) that's what I noticed

R: ok so the next question is how important, relevant and helpful is using AfL to delivering good quality teaching of English to young learners? Please comment on the teacher's and learners perspectives.

T7: it is important because I think that for the young learners it's very important to see that learning is part of them so it's a long term investment for the learners **MOT** especially for the kids who are not so keen on learning (1) they may get into the routine of constantly looking at what they do not only during English lesson but also at school so it might help them outside you know (1) and through AfL they can also get bits of information from me about how they've done

R: ok thanks let's move on have you noticed anything that you'd describe as positive or negative impact of AfL in your lessons?

T7: I haven't noticed anything negative. It's had a positive impact on me **CONTAS** because I(1) my planning has now (1) includes this so definitely the planning is more thorough and also it's very positive for the students because they see you know (1) they are not receivers of teaching but they are part of the teaching and they have their say by talking about WALT what it means by assessing **FEED** their work by showing if they understand they are actively engaged in this process so I think only positives

R: What do you see as the most important influence of AfL on your practice, if **LPL** anything?

T7: hmm (1) most important for me I think is a that I've become aware of a very big importance of what kids can bring into the classroom they can decide **RESP** and they are responsive and they and I mean I also notice how much they can **FOCLO** learn (1) and even if they come to a class in the morning and they are sleepy and so on they got involved and they catch on you know lots of things so I think it is good for teaching and for learning (1) with my years o experience it was good to do something new so this had impact on me (1) and for the kids I think, they also notice that they can say something and assess and they like doing this and that shows them that they are responsible for something (1) it influences what you do because it focuses on you knew (1) kids focus on the learning and they see the reasons why they do things and then they see how well they've

done something and they can do it themselves (1) so that's I think the main

focus

R: thank you and is there anything you'd like to add?

T7: not really just that I have enjoyed doing it and I am happy that I have these   CONTAS

tools at my disposal really   RESP

                                                                                        FOCLO

KEY TO CODES: FEED

INTAFLTL		**Data concerning the teachers' beliefs about the integration of AfL with teaching and learning processes in TEYL classrooms.**
	AWA	making learners more aware of what they are learning;
	FOCLO	focusing learners on achieving the learning objectives;
	CONTAS	teachers continuously building AfL opportunities into lessons;
	FRFEED	being a framework for giving meaningful explicit feedback that helps to move learning forward (from teacher, peers and own reflection);
	MOT	motivating learners to learn and helping them enjoy the process;
	LPL	providing structure and focus to teachers' lesson planning.
COMPTM		**Data concerning the teachers' descriptions of whether and how AfL was compatible with TEYL methods.**
	NEWN	AfL was considered a new name for a concept that was already present in the teachers' TEYL practice.
	ASEIM	AfL was considered easy to implement in communicative TEYL classrooms.
	NCOMSA	there existed some tension between AfL and summative assessment in a TEYL context, which seemed to be connected to reporting to parents.
AFLTECH		**Data concerning accounts of how the teachers implemented AfL techniques. Names of techniques:**
	LP	Learning Partners
	SF	Smiley Faces
	SC	Success Criteria
	THUD	Thumbs Up and Down
	WALT	What are we learning today type questions
PURUSE		**Data concerning descriptions of the purposes of using AfL in TEYL classrooms.**
	GCINS	giving and clarifying instructions,
	SHAIM	sharing aims,
	CFS	criteria for success,
	FEED	feedback
	MESCONF	measuring learners' confidence about their learning
IMPIPG		**Data concerning accounts of the teachers' beliefs about the impact of AfL on sustaining individual, pair and group work**
	SCAFF	AfL techniques provided scaffolding.
	RESP	Using AfL helped learners become more responsible for own learning.
IMPINTER		**Data concerning the teachers' reports of the impact of AfL on interactions in TEYL classrooms**
	MONITTL	When students were able to work without a teacher's help, the teacher could spend time on monitoring work more effectively and providing individual support (T-1L interactions).
	PAIRLL	Using AfL enabled teachers to introduce more pair work in the lessons (L-L interactions).
	COLLAB	Students collaborated rather than competed when working together (L-L interactions).

Appendix 30: The transcript from the focus group discussion with teachers

(Codes: Teacher – T; Researcher - R)

Please note that the transcribing convention is reported in Appendix 15

R: ok I've switched it on so let's start ok? (1) so I will ask you some in the interviews some teachers were saying that they started using or are planning to 'use more AfL' with age primary (7-9) and pre-teens (10-11). Could you comment on what the phrase 'use more AfL' could refer to? What might the main reasons for such choices be? What outcomes might be expected from that? What characteristics of AfL make it useful in the lessons with children?

T2: the easy answer is that is seems to work very well so goals easily established everyone knows what they need to do and there is no concept of what did we do or what haven't we done and everyone seems to have a clear idea of what's going on the simplest level it works (1)

T3: yeah and it also shows that kids can get involved and take responsibility for their own learning and that applies to young children too (1) I've certainly used it with my young ones

T7: Yes and I have also used it with kids almost every lesson for writing which is sort of obvious and they came up with success criteria for the task and they had also something like that for the speaking and it worked so I think you know=

T1: =yeah, I've used it as well for writing because I think the model works really well and identifying success criteria for writing=

T3: =and in a way with AfL bits like the success criteria I think things like this already exist like something like with the writing task when they have to write the letter they have three bullet points and that's exactly the same as success criteria really

T1: =yeah it does (2)

R: ok, thank you for this (1) in the interviews you also talked about AfL being helpful in learning. I would like to find out more about this area. What do you think about this? How can AfL support learning?

T5: I think like [T3] said just now it gives them the criteria so they know exactly what to do and then you can see if they can do it or if you can maybe help with something actually

T6: I agree [T5] that it makes them more informed what you want them to show you otherwise it's a bit unfair to give them a grade when maybe they didn't know what you

wanted them to do in a task in the first place and here with AfL you say to them that this is Afl and they=

T4: *=I'm not sure about it (1) I'm not sure that we need to identify to the students that this is AfL this technique (1) Although I am always very keen on students knowing why they're doing something I'm not really sure if you'd have to label it*

T3: *especially at that age as long as in the lessons it's enjoyable and it's motivating and they're actually learning something*

T2: *I'm thinking about my classes this semester if I was to label AfL in any obvious way I think they would quickly get bored of that and I think that the key issue for using AfL at that age is that it has to be done in a sort of hidden way for students at that level if possible=*

T6: *=yeah yeah yeah that's what I mean you don't have to tell them this is AfL but you can tell then that these are you success criteria sure that's what I meant*

T7: *I asked one or two classes after we've done something like this I mean success criteria for a lesson I showed them a flipchart and I asked them what they thought of this because we knew how the lesson was going to continue and because we had a discussion of what was going to happen in the lesson and they could understand what they'd be doing so they enjoyed it (1) so maybe finding out from them if they've find it useful was a good idea*

T2: *I would go with that I would say that when I've done AfL and the sorts there was established order either explicitly defined or not it's much better than in a kind of ad hoc way of teaching when kids after a certain period of time start looking towards you either for that order and if they don't have it they get a lost in the lesson (1) so I think it is a very important thing to have that order and kids get used to it very quickly*

T1: *yeah so maybe that's quite important but if I'm honest initially I was a bit unsure if this would work but with time I have seen exactly that they are more focused on learning because maybe they know exactly what it is they need to do or even maybe how to do it=*

T7: *=yeah*

T3: *the whole thing is about that; getting them to understand what to do and then to figure out how well they have done it and it doesn't matter which techniques you use, right?*

T5: *I see what you're saying but I also think that they need to know what they need to improve you know (1) like they need to know why they are getting an amber light yes?*

T3: *yeah ok (1) yes (1) yes (1) that too*

T1: *it does perhaps depend on the groups or how you present it but I think my students would get a bit bored if it was just success criteria and traffic light I feel like they are*

more with me if I vary it a bit I mean I keep the same focus as you said it is about sharing the purpose of this lesson and then if they achieved the goal and maybe getting them to think how they can get better but I feel that it needs variety

T8: *I'll support that actually both of you I think (1) some variety is needed but in fact it's the purposes that make it all meaningful and worthwhile*

T2: *so I would say that AfL is a kind of philosophy that involves measuring your students' progress minute-by-minute of the lesson on the ongoing basis and checking that they have the understanding of something and it also involves them knowing what to do in order to achieve goals*

T5: *yeah I guess*

R: *would anyone like to add anything else to this? (2) no? (1) ok in that case my next note says the following it seems from the initial analysis of the data that AfL is generally perceived as a form aiding continuous assessment that provides the teachers with ongoing knowledge of how their students are coping with tasks (1) would you agree with this? (1) how does this knowledge affect classroom practice?*

T7: *young learners like routines and when routines are established they feel safe=*

T2: *=and if a routine is established or they make up their own routine they are happier that when there were no routines=*

T8: *if you go to something as simple as drilling, which is not AfL I know but the teacher often thinks are we in danger of being boring but it has been said that drilling makes students at certain levels feel very comfortable very safe so there is one thing to do with it*

T7: *=and fun is also important but fun but the key is to have a variety fun element and learning and I think in my experience at least they enjoyed learning things and they were curious=*

T5: *=yes I agree with that=*

T7: *exactly and I really enjoyed looking at those kids when they were assessing their writing as peers and this actually helped because you know what kids are like (1) and improved dynamics of the groups*

T3: *yeah (1) generally (1) yeah that's a good point*

T1: *I found that with mine it depended on the AfL technique because at the start of the lesson I always started with can you so by the end of the lesson will you be able to answer this question yes I can and throughout the whole year they were excited (1) they came in they sat down and they were (1) can I read it I want to read it but success criteria for writing I did find that they kind of lost the fun (1) the first semester they were really oh I*

can see why that's good that's fun and by the second semester they kept saying we already know we already know so I wondered if it needed more variety because of that technique and of how I was presenting it rather than the technique being flawed (1) it was my presentation so I now have the star system now and they need to use that instead

T2: with time input from teacher should be smaller =

T7: =mhm=

T2: =because they know their success criteria how to do it or they should be aware of what is expected but I think at that point monitoring becomes more effective to make sure that they're actually doing it properly

T4: I think this is a very good point=

T1: =yes more time to monitor better

T4: I think [T2] is probably right and when you're doing it you can give them some individual advice and talk to them when they're stuck on something

T1: what [T2] seems to be saying is that when they can get on with a task with their success criteria etcetera, then you as a teacher have more time to monitor and help them

T2: yes and how you would do that is an important element of this discussion when you monitor and how you monitor and what you say to individual students

T6: they might think they know what they are doing but they may have the basics established but actually without the monitoring in place they could then go off down different roads and do different things and you know

T3: students are always inclined to start doing an activity and when they think they know what they have to do (1) they will start doing it and as a teacher you have to stop them and say look, this is what you have to do look at success criteria include them all don't just do it quickly=

T6: =exactly=

T3: =so I think the teacher always has to be there=

T8: =yeah they just sometimes want to be first

T4: I mean people tend to get trendy and say ok we're going to hand it all over to students' but the teacher still has to be there doing the job and monitoring and making sure that they are doing it in a right way=

T1: =yeah I don't think anyone would question the need for teacher to be there but to monitor and help them

T2: yes and you can do it more effective or less effective ways to make sure that what you wanted to be happening is in fact what is happening but then if you have to stop them half

way through and sometimes even go back to the beginning AfL is like a path you can follow and make sure you don't get lost because they know what to do next to complete the task=

T8: *=a path how poetic=*

T2: *=you want them to complete (1) teachers should be aware of that and it's probably not the easiest skill to pick up*

T8: *yeah ok but monitoring doesn't' have to be sitting over somebody's shoulder like a vulture and =*

Ricky: *=well yeah this is the thing=*

T8: *=and I've had debates with people who observed me in lessons before because I can be on this side of the room monitoring students speaking on that side of the room=*

T7: *=yes exactly=*

T8: *= yeah and I can do it because I've trained myself to do it or I've been trained to do it*

T5: *yes it comes with the age I would say I can do this=*

T6: *so you can monitor unobtrusively. And maybe AfL is what helps them monitor themselves a little too*

T7: *=yeah=*

T3: *I'm just listening and all I want to add to this is that I still give the students and impression that they are full steam on their own because I don't want them to rely on me so much (1) but they can rely on success criteria sort of thing (2)*

R: would anyone like to comment on this?

T6: *it's just going back to what we already said really isn't it about monitoring*

R: ok then let's move on to my next note (1) some research into AfL suggested that AfL might lead to improvement in students' achievement. These studies were carried on large samples of students and over a period of many years so I am not looking here to confirm or refute these (1) what I would like to understand better though is why the improved achievement might happen? in what way could AfL improve achievement? what is your opinion about this issue?

T4: *from one point of view, with exams we know what the exam format is and when you drill the students through various means to practice that formula before they do the exam then when they do the exam there are no surprises and everyone knows what is going on (1) preparation is key*

T6: *well (1)*

T1: *I think for me for the young learners you're teaching them skills of (1) before they start something to consider what will make it successful so instead of launching into something*

and just doing it they pause and my primary groups especially have learnt to think before they do a thing=

T7: =mhm=

T1: = which is a skill they didn't have before

T7: yeah that's true=

T3:=yeah=

T7: =generally raising their awareness that they have to think before they say it and they really have to find these bullet points on paper or in their heads so I also think that generally using those techniques (1) learners enjoy learning more because they can see the progress from point A to point B and somehow throughout the classroom

T3: I guess one potential danger or pitfall is that if they become accustomed to the idea of bullet points it might set a limit to what they would be doing in the free natural world and it becomes like Callan when you're doing sentence sentence sentence

T6: it might encourage this way of thinking as well but I'm not sure if that happens or not (1) well I've not seen it happen <unintelligible>

T7: I just had another reflection that you know if you balance the lessons where you go through the success criteria with the lessons where students themselves decide what to do you'll notice that they implement these without even being asked to do it so I think it's not a matter of training like parrots or Pavlov's dog but it's about teaching them how to think

T4: I like that

T1: and how about (1) I'm thinking about this particular student of mine [NAME] he was this very messy writer very quick very fast very messy and he has realized that that's a weakness and that was his own success criteria and by the end of the year you can see that students can identify their own strengths and weaknesses and what they have to address as well as like the whole level

T7: yeah it really helps you to do some sort of learner training with them even the younger ones

T1: -yeah=

T7: =I've observed that too (3)

[somebody laughs]

T2: are we done?

R: well no I was just waiting to see if anyone would like to add anything (1) anyone? (2) no? ok so next question from me and please feel free to say whatever you think even if you

suddenly remember something for two questions ago still say it please (1) ok next one (1) has AfL changed anything in your teaching?

T7: *if you see that they are doing well you can proceed to the next task and if they are not doing well then maybe you will have an extra task in between or more scaffolding (1) maybe this yeah (1) that you can take decisions within one lesson AfL makes that easier*

T5: *actually yeah I agree with that I think it does make it easier*

T7: *yeah in order to give them something more complicated or fine-tune something they are still working on*

T2: *my initial thoughts to that questions were taking me back to the idea of monitoring because you've established the framework and they know what's going on and it's time to do a writing exercise success criteria whatever they now have to create their own piece of writing and then there is a question if ok (1) let's now monitor that they are going down the right tracks or if they are not (1) and if they not then it's kind of getting back there or re-teaching but suggesting and helping and hinting in various ways different thing*

T3: *it's a question of helping them see what they are not doing quite right yet and pointing out with AfL how to improve this*

T7: *yeah definitely (1) if you have things like you know how well you did with the task and then starting to ask them which things needed improvement and how they could do it so working out some sort of action plan you know and making them aware of what needed to be improved or something like that*

T1: *you mean in response to their individual needs rather than us deciding about that?*

T7: *yeah it could be but it could also be us deciding because sometimes they just need that guidance*

T2: *I'm looking at all your questions and all your thoughts and a lot of my work and I'm kind of thinking all the way through this oh, how would I put AfL into this or that and I with certain aspects you could do it and I think that it seems to work very well as a group with a small group or a big group but on the one to one level I think that it has its prons and cons and perhaps you can see them see a little bit more clearly with a group but it could be just down to one practitioner and how he or she applies AfL to that situation*

T1: *I used AfL with my one to one student who was 11 and especially the focus of the question within the next hour you'll be able to do this it's really (1) it worked for him that element has definitely worked*

T7: *And [T2] what age group were you thinking about?*

T2: *pre-teens really*

T7: *So in this case I can see a lot if you like because you could work out a learning path for this person and =*

T2: *=that's really it a lot of students at that age do not really know what they want and parents want them to come to English and at that point you have a lot of work to do to get them on board but with AfL maybe it can be this is this target =*

T6: *you see I love it with pre-teen I always feel so good after the lesson with AfL I go out of the lesson and I feel that went well you know so I'm happy with using it in that age group for sure*

T7: *which works for me too but also with primaries I like with primaries too I mean*

T3: *I have a similar experience like when I prepare some AfL and I use it and they really sort of switch off but then I don't have the time or I have to do a test or something and then I forget and when I remember again it's like (1) it really work for them then*

T8: *yeah I think you know we need to remember that it's not just AfL but a test also serves a purpose you know I mean I don't test them all the time but at least I know you know for when the parents ask me what the id has learnt then I can show them it's there on the test*

T5: *you see I stopped doing tests actually (1)I mean I still do the material but it's not like a big deal test but we just do it and then like use traffic lights to mark it together or something*

T8: *I see*

T2: *Yeah, in this case =*

T3: *= I still do the tests but I sort of feel less sort of attached to them not sure why*

T1: *and the children don't like tests either*

T7: *it's like an escape for them (1) I mean AfL (1) escaping from testing* **[laughs]**

T4: *it's a phenomenon of TEFL teaching (1) it really is(2)*

R: what is?

T4: *testing (1)*

R: what do you mean?

T4: *we test so much (1) like all the course are for some sort of exam ultimately*

T8: *yeah what you gonna do? tough s**t you know*

T5: *no but with kids you can sill teach them I think like without testing all the time*

T1: *I agree with that I think kids are lucky really (2)*

R: *any other comments?*

T2: *I like how it works*

T7: *mhm I like how it works for kids*

T2: *yes it seems like a very basic approach but it works really effectively and I think it's a good thing and I think it's definitely a way forward*

T5: *=oh yeah=*

T2: *for a lot of people and the way they teach and maybe we should come up with success criteria for England football team*

T8: *Yeah can dos*

[a number of people laugh]

R: *anything alse?*

T5: *no I think we need coffee now*

R: *ok thanks everyone END OF RECORDING*