Designed to fail:

Evaluating Grade 1-3 English First Additional Language textbooks as preparation for English medium Grade 4 Natural Sciences and Technology in South Africa

Gené Barnard McAravey
MCRGEN001

A minor dissertation submitted in partial fulfilment of the requirements for the award of the degree of Master of Education

School of Education
University of Cape Town
2021

Declaration

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature: Signed by candidate Date: 2 February 2022

Dedication

This dissertation is dedicated to my parents.
To my father, Cyril, who taught me about writing and remains the editor inside my head,
and to my mother, René, who believed I could anything.
Thank you for raising me with so much love.

Acknowledgements

To my supervisor, Carolyn, thank you for your expertise, insightful feedback, and consistent, gentle support. Your patience and encouragement has been a balm in this season.

To my sister, Camren, thank you for your endless patience, love and practical support during the writing of this thesis. Thank you for always listening to my thoughts and giving me valuable feedback. Thank you for your guidance in the use of Excel and for reading through endless drafts. Perhaps most importantly, thank you for thousands of cups of tea placed next to my computer day after day. I know I wasn't always the easiest to put up with, but I hope you know how much I love you and value your support. Without you this thesis would not have been possible.

Finally, to my Lord and Saviour, "With man this is impossible, but with God all things are possible."- Matthew 19:26. Thank you for your abundant blessing on my life. May you watch over the children of South Africa and create beautiful possibilities for their futures as well.

Abstract

Despite official school language policies mandating English as the medium of instruction from Grade 4 onwards, nearly half of South Africa's English Language Learners (ELLs) are obtaining less than 50% in their final (matric) English examinations. These same students are writing all of their other high school examinations in English. It seems clear that the majority of students do not have an adequate mastery of written English.

What is the nature of the English language instruction these children are receiving? Especially in the earliest years of schooling? Are children being adequately prepared to learn through the medium of English? To investigate these issues, the research question for this study asked: 'Does children's learning of English in the subject English First Additional Language (EFAL) in Grades 1-3 prepare them to make sense of a Grade 4 Natural Sciences and Technology (NST) textbook?'

This study employed content and discourse analysis, specifically relating to school textbooks and curriculum documents. The focus was on evaluating the readability of one Grade 4 NST textbook, in relation to the language resources provided by a Grade 1-3 English Language Teaching (ELT) scheme. The methodology employed was mixed methods, with elements of both quantitative and qualitative data analysis.

Findings revealed that for all three key areas relating to text accessibility, namely Vocabulary, Genres and Syntax, the subject EFAL in Grades 1-3 does not adequately prepare learners to make sense of a Grade 4 NST textbook. Based on this, I recommend that learners no longer be made to transition to English as the Language of Learning and Teaching (LOLT) in Grade 4. Instead I recommend that children receive bilingual instruction throughout primary school in all learning areas, and that LTSMs and teacher training be updated to support this approach. I also recommend that EFAL and other home and additional languages be consolidated into a single language arts period and also taught bilingually. Alternatively, I recommend that learners be given two to three additional years to study EFAL before using it as a LOLT and that the EFAL course of study be revised to align more closely with the language and literacy demands of content subjects.

List of acronyms and abbreviations

AWL	Academic Word List
BICS	Basic Interpersonal Communication Skills
BNC-COCA	British National Corpus and Corpus of Contemporary American English
CALP	Cognitive Academic Language Proficiency
CAPS	Curriculum and Assessment Policy Statement
CEFR	Common European Framework for Languages
DBE	Department of Basic Education
EFAL	English First Additional Language
ELLs	English Language Learners
ELT	English Language Teaching
GSL	General Service List
FAL	First Additional Language (a.k.a. L2)
FP	Foundation Phase (Grades 1-3)
HFWs	High Frequency Words
HL	Home Language (a.k.a. L1)
K-1, K-2…	1st thousand most frequent words, 2nd thousand most frequent words…
LFWs	Low Frequency Words
L1	Home Language / First Language
L2	First Additional Language / Second Language
LB	Learner's Book
LOLT	Language of Learning and Teaching (a.k.a. MOI)
LTSMs	Learning and Teaching Support Materials
MOI	Medium of Instruction (a.k.a. LOLT)
MRC	Medical Research Council
NST	Natural Sciences and Technology
MS	Microsoft
R	Reader
SFL	Systemic Functional Linguistics
SLA	Second Language Acquisition
TOEFL	Test of English as a Foreign Language
TOEIC	Test of English for International Communication

List of figures

List of tables

Contents

Chapter 1: Introduction and Rationale

1.1. Context

South Africa is a multilingual nation, with twelve official languages enshrined in its constitution. These include nine African languages, namely isiZulu, isiXhosa, Sepedi, Setswana, Sesotho, Xitsonga, Siswati, Tshivenda and isiNdebele, as well as Afrikaans, English and South African Sign Language. In sharp contrast to this linguistic diversity however, the exit examinations for the country's public schooling system are offered in only two of the official languages: English and Afrikaans.

What makes this particularly problematic is that only about 20% of the country would describe one of these two as the language they speak most often at home – 12% for Afrikaans and 8% for English (Statistics South Africa, 2018). This means that for the vast majority of South African children, success in their final school examinations (with future education and employment prospects at stake) hinges critically on an ability to demonstrate knowledge and skills using a language not spoken often in their homes. Overwhelmingly, the language chosen for this is English (Plüddeman, 2015).

If we examine the results of these exit examinations, particularly for the subject English First Additional Language (EFAL) which is written by non-home language speakers of English, a disturbing picture emerges. Of the 489 072 candidates who wrote the National Senior Certificate examination for EFAL in 2019, 42% did not meet the minimum criteria of eligibility (a grade of 50%+) to study at any of South Africa's tertiary institutions. An additional 30% of the cohort only met the criteria for some of South Africa's tertiary institutions (above 50% but below 60%). Only 28% of the candidates obtained an EFAL result that would enable them to study at any tertiary institution in South Africa (60%+) and of these, only 9% achieved a result above 70%. (DBE, 2019:16-17, own calculations)

These results are problematic given their implications for students accessing (or rather, failing to access) tertiary education. However, what is far more serious is the fact that these same students are writing all of their other high school examinations (Physics, Biology, History, Geography, Mathematics, Economics, to name only a few) in English as well.

We also know that the students who manage to make it as far as these exit examinations are only 60% of the cohort that originally enrolled in Grade 1 – the other 40% drop out somewhere between Grade 1 and Grade 12 (Van Broekhuizen et al., 2016). Although these poor educational outcomes can be attributed to any number of systemic failures, it must be asked: for how many of these students, unwilling or unable to complete twelve years of schooling, was the language of learning and teaching a contributing factor?

While South Africa is considered to have a very progressive language in education policy which promotes additive bi/multilingualism (Department of Education, 1997), in practice South African schools follow the 'early-exit' model of bilingual education (Ouane & Glanz, 2011; Plüddeman, 2015). The Department of Basic Education (DBE) encourages and supports home language education for the first three of years of formal schooling (Grades 1-3) while English is studied as an additional language. After these three years children are expected to make an abrupt transition to English as the medium of instruction (MOI) in Grade 4, with African home languages relegated to one subject among many for the remainder of a child's school career.

While schools have the legal power to offer home language instruction in African languages or even bilingual instruction beyond Grade 3, almost none of South Africa's schools have been willing or able to do so (with the exception of a limited number of schools involved in pilot projects – see Mini & Botha, 2020, and Mbude, 2020, for more details). And yet, as the results discussed earlier show, the majority of learners have not developed the proficiency required to successfully write their Gr12 exit examinations in English.

While South Africa's educational challenges are not limited to language alone, we need to understand why, despite having received eight (or sometimes more) years of continuous instruction through the medium of English, nearly half the learners are still obtaining less than 50% in their final EFAL examinations (not to mention their results in their other content subjects also written in English). What is the nature of the English language instruction these children are receiving? What is being taught? How is it being taught? With what kinds of resources? Especially in the earliest years of schooling, are children being adequately prepared to learn through the medium of English? It is these questions that have motivated me to embark on the research that is the focus of the rest of this paper.

1.2. Research Question and Objectives

The research question for this study is as follows: Does children's learning of English in the subject English First Additional Language in Grades 1-3 prepare them to make sense of a Grade 4 Natural Sciences and Technology (NST) textbook?

The aim of this study is to gather evidence establishing whether children are being adequately prepared in their first three years of schooling to undertake the remaining nine years using English as the medium of instruction, or the Language of Learning and Teaching (LOLT) as it is officially called in South Africa. It is important to recognise, however, that the medium of instruction in South Africa (as is the case in many other countries) is a highly contentious and politicised issue.

There is a great deal of debate surrounding the merits of mother tongue education versus learning in English as well as discussions around what grade it is best to transition languages in. It is not self-evident that the policy currently being implemented in the majority of South African schools is the most effective or desirable of the choices available. Indeed, it seems highly unlikely that the 2-4 hours of English a week recommended by the Grade 1-3 EFAL Curriculum and Assessment Policy Statement (CAPS) is sufficient preparation for learners to use English as a LOLT in Grade 4.

Ultimately, however, whatever the policy chosen, the onus rests on the planned curriculum and materials approved by the DBE to fulfil the objectives given. One way to determine if this is being done is to undertake a detailed analysis of the contents of the approved learning and teaching support materials (LTSMs) to establish what kind of learning and teaching is meant to be taking place in the classroom. This is especially important since in many of South Africa's public schools, the DBE's supplied LTSMs are often the only teaching materials teachers make use of in the classroom.

I would like to emphasise here that the chief concern of this study is the extent to which possibilities for learning are created by the curriculum and materials approved by the DBE. This evaluation does not extend to how the materials are actually used in practice. To borrow the words of another researcher in this field:

> My concern…is with the analysis of materials 'as they are', with the content and ways of working that they propose. This, it must be emphasised, may be quite distinct from what actually happens in classrooms. Analysing materials…is quite a different matter from analysing 'materials-in-action'. Precisely what happens in classrooms and what outcomes occur when materials are brought into use will depend on numerous further factors, not least of which is the reinterpretation of materials and tasks by both teachers and learners. (Littlejohn, 2012:181)

It is beyond dispute that teachers have an essential role to play in mediating the use of learning materials in classrooms. However, we must first scrutinise what teachers are being asked to teach (and how the DBE proposes they do so) to determine whether the goals being set are realistic and appropriate.

If the planned curriculum and learning materials have been created in such a way that, after three years, they do not adequately prepare children to study other subjects in English, then the DBE is perpetrating a gross injustice against learners and teachers alike. If this is indeed the case, it would not be an exaggeration to say that South Africa's public education system has been designed to fail.

1.3. Chapter Outlines

The remainder of this paper is organised into six chapters. Chapter 2 will review available literature related to two questions: 'What is language?' and 'What does it mean to know a language?' as well as exploring literature related to readability and school textbooks. Chapter 3 will provide a brief overview of the analytical framework that will be employed in this study as well as addressing some methodological issues.

Chapters 4, 5 and 6 will focus on my analysis of textbooks, exploring the topics Vocabulary, Genre and Syntax respectively. Each of these chapters will be subdivided into a Framework, Methodology and Results section. I have chosen to structure the chapters in this way because each of these topics constitutes a field of study in its own right. The literature pertaining to each as well as the methodology employed in my analysis will be distinct for each of the three topics. As such, the analysis and results will be easier to follow if the topics are dealt with thematically. Finally, Chapter 7 will revisit my research questions and bring together the findings from Chapters 4, 5 and 6. I will then discuss some recommendations and conclude.

Chapter 2: Literature Review

In order to answer questions around the teaching of language as well as the use of language in teaching materials, I must first be able to answer two more fundamental questions: 1) What, exactly, is language? and 2) What does it mean to know a language? To this end, my literature review will begin by developing a definition of language and then move towards a discussion of what constitutes knowing a language. From this point I will proceed to a detailed examination of school language and literacy and finally to discussing research around textbooks and their readability which will be of particular relevance to this study.

2.1. Theorising Language

What is language? Garcia and Wei (2014:6) explain that, for most people, a common sense understanding of language is "what we speak, hear, read or write in everyday life. And we speak, hear, read and write in what are considered different languages, such as Arabic, Chinese, Spanish and Urdu." The limitations of this view of language quickly emerge when considering the existence of sign languages, which make use of gestures, as well as traffic signs, which make use of symbols. Neither of these fit neatly into the definition mentioned above and yet both are well-developed systems of communication.

An alternative view of language, put forth by linguist Ferdinand de Saussure, is that language is a system of signs, where each sign consists of a signifier (in the examples above, the sound, gesture or symbol) and a signified (the concept or idea to which the signifier refers). Saussure also made a distinction between what he called *langue* and *parole.* He used the term *langue* to refer to the abstract system of rules and conventions which govern specific languages while the term *parole* was used to refer to actual instances of language use by individuals. (Garcia &Wei, 2014:6)

The scholar Noam Chomsky reformulated the *langue*/*parole* distinction into the categories competence and performance, where the former refers to language as an abstract system and the latter, real instances of language use by people (Garcia & Wei, 2014:6). Historically, it has been the study of competence that has dominated the field of linguistics and consequently informed the teaching of language in classrooms. This view of language privileges knowledge of vocabulary and grammar and ignores all the complicating factors that arise in actual

communication contexts, regarding these as obscuring our understanding of language as a system.

Since the 1970s a different school of thought has emerged which posits that it is exactly these contextual factors which are central to understanding how language functions. Contextual factors may include things such as the cultural and social environment of the communicators, their backgrounds and relationship with one another, their shared knowledge, their aims and purposes for communication, their desires, preferences, sense of identity as well as the linguistic repertoires they command.

To recognise the important role that each of these factors play is to widen the scope of knowledge required to be a competent language user to an almost infinite degree. This problematises the approach to language teaching which focuses almost exclusively on the skills of listening, speaking reading and writing. It calls into question the assumption that language knowledge can be decontextualised, universalised and taught by schools in ways that will enable students to operate as competent users across the wide array of possible communication contexts.

One prominent scholar who argued for the social approach to the study of language was the sociolinguist Dell Hymes, who critiqued Chomsky's concept of competence and countered it with the enlarged concept of 'communicative competence'. Hymes (1972) argued that when a child learns language, their learning extends beyond mere knowledge of the grammatical, into the realm of appropriate language use. Hymes' concept of communicative competence involves knowing when to speak and when not to, what to talk about with whom, when, where, and in what manner.

James Gee, another theorist who subscribes to the view of language as social practice, combined linguistic knowledge with knowledge of these other social elements of communication to form his concept of 'Discourse: with a big D'. He defines the latter as "[c]ombinations of ways with words and ways with "other stuff" (bodies, clothes, objects, tools, actions, interactions, values, and beliefs) that can get people recognized as having certain socially significant identities" (Gee, 2015:2). Gee explains that understanding both oral and

written language necessarily involves understanding the discourses that people use to assert and contest their views of the world and their own positions in it. (Gee, 2015)

This approach is not limited to understanding spoken communication and is increasingly gaining traction as a way of approaching the study of literacy as well, notably in Brian Street's characterisation of what he calls autonomous versus ideological models of literacy. Autonomous models of literacy are those which "conceptualise literacy in technical terms, treating it as independent of social context, an autonomous variable whose consequences for society can be derived from its intrinsic character" (Street, 1992:5). Street contrasts this with an ideological view of literacy which recognises "the variety of cultural practices associated with reading and writing in different contexts" and which sees literacy as inseparable from power structures in society (Street, 1992:5). A great number of scholars, including Gee, have done work in this field, often referred to as New Literacy Studies, some of whom I shall highlight as the discussion proceeds.

To return once more to the questions outlined at the start of this section, in answer to 'What is language?', this paper subscribes to a view of language and literacy as social practice. I maintain that language and literacy are always and everywhere embedded within particular sociocultural contexts and are used by people with particular identities for particular purposes, both of which may depend on the situation and change over time. All of this together shapes the production and reception of language. In answer to the second question, 'What does it mean to know a language?', knowing a language is not limited to simply knowing its vocabulary and grammar, but extends to understanding and being able to produce the various kinds of language required by different communication contexts.

The sociocultural approach to language and literacy necessitates a further important recognition, namely that it is not possible to ever know a language in a complete sense (Blommaert & Backus, 2011). Rather, an individuals' knowledge of language is dynamic – it changes over time as the position and needs of its user changes. Language knowledge is always only ever partial – a fraction of the totality of linguistic resources available. Some types of language we only acquire as we enter new communicative terrains (formal employment, parenthood, membership of a sports team) while others fall away once we no longer consider them necessary or appropriate (teenage slang, high school French) (Blommaert & Backus,

2011). The acquisition of language resources is less like accumulating chips at a poker table and more like the development of muscles in a living organism. Language resources can expand or shrink, shift and reshape themselves, strengthen or atrophy over time depending on how they are put to use.

This has important implications for teaching and evaluating language abilities. Blommaert and Backus (2011) use the example of the Common European Framework for Languages (CEFR) to illustrate this point. They explain how the CEFR has become a widely circulated standard for measuring language competence, but state that:

> such measuring instruments are a form of science fiction. They have only a tenuous connection with the real competences of people, the way they are organized in actual repertoires, and the real possibilities they offer for communication. This is because they measure only part of language knowledge, a part that is privileged for socio-economic reasons, not for inherent linguistic ones. (Blommaert & Backus, 2011:24)

Using the example of one of the authors, they sketch a picture of how the same person's English proficiency might be ranked very differently depending on the contents of the test they are assessed with. Were the test to focus on academic genres of text and talk (the author's area of expertise), he would almost certainly score a C2 – the CEFR's most advanced level of proficiency. If, however, the test were based on how the author interacts with a medical doctor, a plumber, an IT helpdesk operative or an insurance broker, he would likely score an A2 – the most elementary level of proficiency. They conclude by asking "So, 'how good is his English' then? Let it be clear that this question can only be appropriately answered with another one: 'which English?'" (Blommaert & Backus, 2011:24).

This concluding question of Blommaert and Backus is key for the purposes of this study. A sociocultural view of language highlights the vast array of linguistic and sociolinguistic resources that people can and do deploy across different communication contexts. If one wishes to evaluate English language instruction in South African schools, it must be acknowledged that English language skills are not monolithic. It becomes necessary to examine exactly which resources schools are giving children the opportunity to learn (Gee, 2003), how these relate to the knowledge and resources children bring with them from home, as well as what knowledge and resources they will need going forward in their schooling. In order to answer these

questions, I shall now proceed to a discussion of the types of language and literacy typically found within classroom contexts.

2.2. School Language and Literacy

In 1982, Shirley Brice Heath published a book entitled 'Ways with words: Language, Life and Work in communities and classrooms' containing a now famous language and literacy study on three different communities in the South-eastern United States. The chief finding of Heath's study was that not all communities socialise their children to use language and literacy in the same way, or in ways that schools value and reward. When children struggle to meet the language and literacy demands of the classroom, the issue is often that they have grown up in environments that have afforded them different types of knowledge and nurtured different language and literacy practices. Naturally, this shapes how these children engage with formal schooling. (Heath, 1982)

Gee (2015) conceptualises this phenomenon as the gap between children's primary and secondary discourses. According to him, almost all human beings acquire a discourse from the group or community that socialises them in early life, which he calls the primary discourse. He goes on to explain that all the discourses we acquire later in life become our secondary discourses. These can be the discourses of religious communities, businesses, government institutions, video games, or of schools and more. The discourse of school involves using certain kinds of school language and being a particular kind of student. "This identity and these forms of language can, at points, conflict with the identities, values and ways with words some children have learned at home as part of their primary discourse. For other children there is a much better fit." (Gee, 2015:4)

The problem goes further than simply a mismatch between home and school literacies, however. Even within schooling, different subjects make different language and literacy demands upon learners. Blommaert and Backus (2011:19) explain that language resources "cluster around particular social arenas and become generative in those arenas". That is to say, the growth of linguistic resources as a result of exposure and use tends to be limited to within particular discourses, and thus do not necessarily translate into competency in a different area.

Gee (2003) highlights this issue in his paper entitled 'Opportunities to Learn: a language-based perspective on assessment'. He states that:

> [r]eading is not such general and obvious matter as most people, including test makers assume…After all we never just read 'in general', rather, we always read or write *something in some way*. We don't read or write newspapers, legal tracts, essays in literary criticism, poetry, or rap songs, and so on and so forth through a nearly endless list, in the same way. Each of these domains has its own rules and requirements. Each is a culturally and historically separate way of reading and writing, and, in that sense, a different literacy. (Gee, 2003:28)

This, argues Gee, is one of the chief reasons children obtain poor educational outcomes, not the economic background of learners or a failure to teach phonics well in the early years. "[T]he truth of the matter is that a great many more children fail school because, while they can decode print, they cannot handle the progressively more complex demands that school language makes on them as they move up in the grades and on to high school."(Gee, 2003:43)

To relate this more directly to what happens in classrooms, the knowledge and practices involved in answering a comprehension activity for a science text differ from those required by a piece of creative writing assigned in English and differ again from those needed to critique a historical account of events. Children tend to learn the resources needed for each of these tasks separately, usually by being exposed to examples of such texts and then being given an opportunity to produce them with guidance and feedback. Though some literacy knowledge and skills are transferable (such as some aspects of grammar or punctuation), different texts still require different vocabulary, have different registers and may be structured very differently at both a text level (compare the organisation of a fairy tale with that of a report on a science experiment) and a sentence level ('Once upon a time…' vs 'Step 1: Place the beaker on the table…'). Schools cannot simply teach one type of literacy in the home or first additional language as a subject and then assume that children can perform a very different literacy skill in a content subject.

This limitation of language resources is also reflected to some extent by the popular distinction made by Cummins (2008) between Basic Interpersonal Communication Skills (BICS) and Cognitive Academic Language Proficiency (CALP). Cummins specialises in the field of bilingual education and created these distinctions in order to account for the discrepancies he observed among immigrant students in Canada who, despite being able to communicate

conversationally (BICS) in English still lagged behind peers in the academic tasks expected of them at school (CALP). Cummins' work serves to illustrate that schools often fail to recognise what children are actually able to do with language and how this relates to what is expected of them.

In South Africa, the treatment of language and literacy within the curriculum demonstrates a failure to recognise the diverse nature of linguistic resources within learners and across subjects. An autonomous view of language is prevalent where literacy is perceived as a unified, transferable skill rather than a diversified, socially situated one. Different languages are kept in silos, each relegated to its own period in the timetable. Within these periods, languages are taught as specialised disciplines, with curricula emphasising language structure, literature, poetry and creative writing skills, rather than focusing on language as a communicative tool. Mixing languages is frowned upon and in many cases considered to be a sign of incompetence on the part of teachers (Probyn, 2009; McKinney & Tyler, 2019).

This monolingual bias (McKinney, 2017; Garcia & Wei, 2014) has meant that the transition to English as the LOLT in Grade 4 is an abrupt and disconcerting one, both for learners and teachers. It is assumed that what is taught in the limited hours of EFAL instruction in the Foundation Phase (FP) is sufficient to equip learners to deal with the English language and literacy demands of a variety of specialised academic disciplines within the Intermediate Phase. It is this assumption that I wish to interrogate.

2.3. Readability and Textbooks

The aim of this study is to develop a nuanced understanding of the kinds of English language and literacies found within South Africa's Grade 1-4 classrooms. Of particular interest is the LTSMS learners encounter within schools. While a great deal of research has been done on language policies and the transition to English within South Africa and the broader sub-Saharan context (Desai, 2016; McKinney, 2017; Heugh, 2013; Probyn, 2015; Ouane & Glanz, 2011), there has been limited research on the appropriateness of textbooks for learners learning in a second language within the region (Milligan et al. 2016:330).

In the South African context, the Threshold Project conducted by Carol Macdonald and her research team in the mid-1980s focused on this issue to some extent. One of the areas Macdonald and her team investigated was the gap between learning English as a subject in the

first four years of school and how this prepared children for using English as a medium of instruction from year five onwards (at the time, learners transitioned to English in Grade 5). As part of this investigation, the project evaluated school textbooks and found that the initial years of studying English were woefully inadequate at preparing children to cope with the kinds of English texts they were later faced with in the other subject areas of school. (Macdonald, 1990; Macdonald & Burroughs, 1991)

There are very few recent studies focusing on textbooks and language learning that have been conducted in South Africa and other sub-Saharan countries. I shall highlight some in my discussion below focusing on two related areas of scholarship, namely the "readability" of texts and the development of language supportive learning materials.

Regarding readability, Gunning (2003:175) explains that the central goal of this area of study is to ensure that teachers are providing materials of the appropriate level to students. This is because students who are given materials that are too easy are not challenged and do not grow, while students provided with materials that are too difficult also fail to progress, go off task, may disrupt lessons, or may become so frustrated that they give up entirely. (Gunning, 2003:175)

According to Gunning (2003), we can assess difficulty of materials either quantitatively, qualitatively or through a combination of both methods. Quantitative approaches make use of what are known as readability formulas. While there are a number of different formulas, almost all of the most widely used ones tend to focus on two main variables – vocabulary difficulty and sentence length. The idea is that the greater the number of difficult words and the longer the sentences, the more challenging the text. The mechanics of measuring these two are elaborated below:

> Sentence length is universally measured by the average number of words per sentence. Difficulty of vocabulary is assessed in a variety of ways: number of words not on a list of words tested and found to be known by most students at a certain grade level, number of words not on a list of high frequency words, grade levels of the words, number of syllables in the words, number of letters in a word, number of words having three or more syllables, and frequency with which words appear in print. All of these measures are highly correlated. (Gunning, 2003:176)

Because there are low frequency words which are easy and high frequency words which are difficult, and likewise for short and long words, Gunning argues that readability formulas which use word lists that have been tested on students are likely the most valid.

If one considers the view language and literacy discussed so far, the limitations of using formulas to establish the readability of texts becomes apparent. Gunning himself acknowledges that readability formulas do not assess "physical features of text such as the size of print, type, quality of illustrations, and number of words on a page...neither is the density of concepts or appeal of the materials [assessed]" (2003:180). To this we could add the topic of the text, the degree to which learners are familiar with it, as well as the vast number of other sociolinguistic factors as discussed earlier.

In response to these limitations, teachers from New Zealand developed a different framework to evaluate the difficulty of texts. The factors they considered included content, children's assumed background or experience, story structure and style, language structure, illustrations, size and type of print, and amount, placement and balance of illustrations with text (Gunning, 2003:180). The reading levels found in the popular international reading programmes, Reading Recovery as well as Fountas and Pinnell are based on this work. These are examples of qualitative approaches to assessing readability.

According to Gunning (2003), qualitative systems have the advantage of being able to assess factors ignored by quantitative approaches. However, since many qualitative factors are subjective, such systems may be prone to judgement error. Gunning concludes that the best approach is to combine both quantitative and qualitative elements to determine what he reconceptualises as the accessibility of texts. In his shift from readability to accessibility he incorporates many of the factors highlighted by the sociocultural approach to language. He explains:

> Difficulty of text should be judged in terms of students' backgrounds, capabilities, and interests. Context also needs to be considered. Will students need an in-depth grasp of information contained in the text or just an overall sense of it? Will they be reading this on their own or will they be given help? Teachers need to ask such questions as: Can my students understand this text? If not, what can I do to make it accessible? Or is it so difficult that I can't make it accessible? Research suggests that students achieve the most when all but 2% to 5% of the words are known. (Gunning, 2003)

It is exactly in response to the sometimes overwhelming gap between learner abilities and text demands in schools that language supportive textbooks have been developed. Barrett and Bainton (2016:393) define these as learning materials that make explicit the language learning that is integral to learning curriculum subjects. According to Milligan et al. "[l]anguage supportive pedagogy is reflected in textbook design in textual characteristics, the range of activity types, the use of vocabulary, the use of visuals and the inclusion of bilingual practices" (2016:332).

Barrett and Bainton (2016) conducted a study where they examined the impact of language supportive textbooks in Tanzania's secondary schools. In Tanzania, children learn for the first seven years in Kiswahili before transitioning to English as the medium of instruction in year eight (a late-exit transition model). Their study examined the impact of textbooks designed to help students transition from an existing primary school knowledge base encoded in Kiswahili, to being able to engage around secondary school science concepts in English conversation and writing.

The authors describe two of the key strategies employed by the language supportive textbooks. The first was the inclusion of glossaries which provided translations of key English words into Kiswahili. The second involved sequencing learning activities to ensure topics were introduced with informal talk activities to enable students to recall their prior knowledge. "Each topic included a reading activity, which was followed by an exploratory talk, giving learners an opportunity to experiment with talking about new concepts. Towards the end of a topic, structured support was given for producing writing in English... Informal and exploratory talk could be in any language." (Barrett & Bainton, 2016: 399-400)

Milligan et al. (2016) conducted their own study on the use of language supportive textbooks in Rwanda at nearly the same time as the Tanzanian study. In Rwanda children study in Kinyarwanda for the first three years of primary school before transitioning to English as the medium of instruction in their fourth year of schooling. Milligan et al.'s research describes how the textbooks they designed were targeted at Year 4 learners with low levels of English language proficiency. The books focused on providing reading passages which were short, with grammatically simple sentences. Academic words and subject specific words were kept to the

bare minimum and vocabulary was contextualised. Learners were also supported to use new concepts in English by Kinyarwanda or bilingual glossaries as well as visuals. Similar to the Tanzanian materials, the activities included in the textbooks were also developed with an eye to supporting the children to speak, read and write about subject concepts in English. (Milligan et al., 2016:332) Both these studies highlight ways in which textbooks can be specifically designed to support children who are transitioning from learning in an African language to learning in English. This begs the question: to what extent do the textbooks used in South African schools support this language transition?

One paper, entitled 'The readability of two Grade 4 natural sciences textbooks for South African schools', published by Sibanda (2014), examines this issue. Sibanda used a mixed-methods approach to evaluate two Grade 4 science textbooks and concluded that both textbooks were above the intended reader's reading levels. Her findings highlight the problematic nature of English content subject textbooks in the South African setting. There are limitations to her research, however. In particular, the quantitative aspect of her study relied on the use of a very opaque *Text Readability Consensus Calculator* which combined a number of readability formulas (no details were provided on which formulas exactly, nor their relative weights). In terms of Sibanda's qualitative design, no clear methodology was provided for the basis on which she classified words as technical or not. Moreover, she did not account for how much and what type of English had been taught to learners prior to encountering the Grade 4 textbooks.

My study aims to address some of these gaps and add to the literature in this field by developing a robust and contextually appropriate framework for evaluating the accessibility of South African textbooks targeting English Language Learners (ELLs). I will now proceed to a discussion of this framework.

Chapter 3: Methodology & Analytical Framework

3.1. Framework Overview

I have drawn on a number of different sources (Gunning, 2003; Milligan et al., 2016; Van Rooyen, 1990; Nation, 2001; Laufer, 1992; Derewianka, 1990; Sheehan et al., 2014; Carney & Levin, 2002; Hall & Cook, 2012; Lorch, 1989; Kim & Clariana, 2017) to design an integrated, mixed methods framework which can be used to the analyse the factors affecting the accessibility of school textbooks from both the language and content subject areas, specifically for materials targeting ELLs. It consists of six areas of analysis, namely: Vocabulary (Vocab), Genre, Syntax, Visuals, Use of Home Language and Sign-posting, displayed in Figure 1 below:

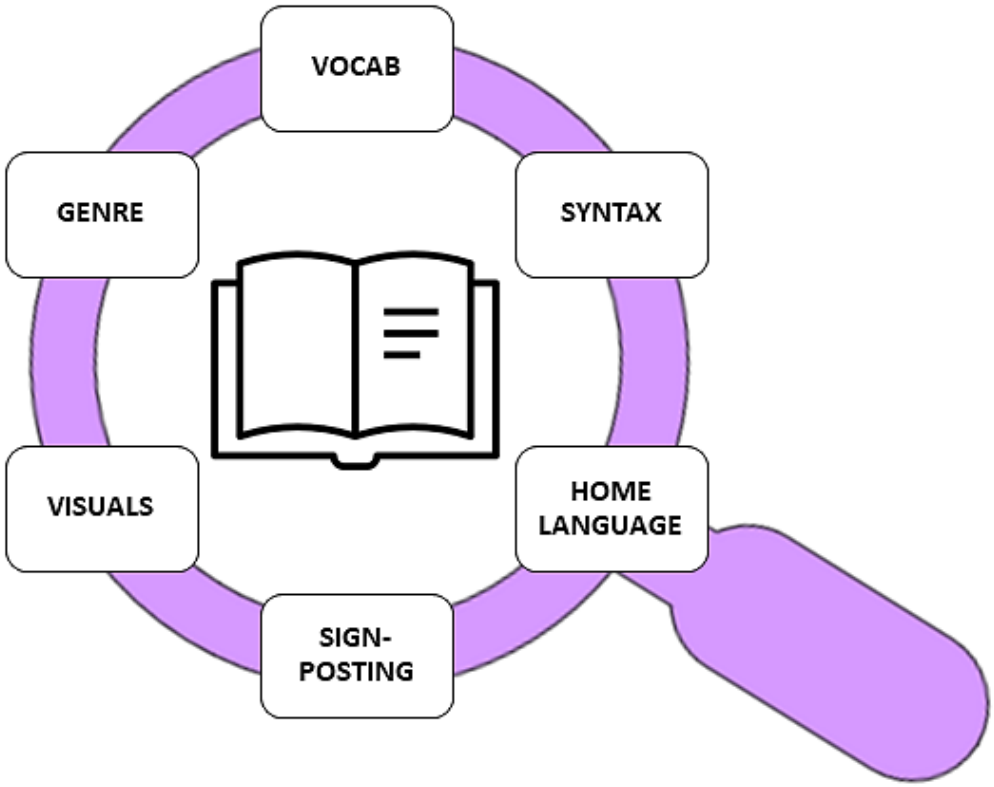


Figure 1: Framework for analysing learning and teaching materials targeting ELLs

Vocabulary refers to the amount and type of words found in a text (Nation, 2001). Genre refers to recognisable structures and patterns of organisation within texts (Derewianka, 1990) and Syntax refers to the grammatical features of sentences (Fromkin, Rodman & Hyams, 2011). Visuals refers photographs, illustrations, diagrams, and other visual aids which often

accompany texts (Carney & Levin, 2002). Home Language refers to the presence of learners' own languages (L1) in or around a text, such as having bilingual glossaries, L1 explanations, etc. in the text (Hall & Cook, 2012). Sign-posting refers to the use of a number of written devices (e.g. headings, sub-headings, previews and summaries, etc.) which help readers organise and recall information about the text (Lorch, 1989).

Of the six areas of analysis displayed, an important distinction can be made between the first three (Vocabulary, Genre and Syntax) and the last three (Visuals, Home Language and Sign-posting). Vocabulary, Genre and Syntax are all intrinsic features of a text – they cannot be separated out from it. Visuals, Home Language and Sign-posting, however, can be considered distinct from texts and are generally added to clarify meaning or otherwise aid in comprehension. In this sense, they may be considered support features of texts. This distinction is displayed in Figure 2 below.

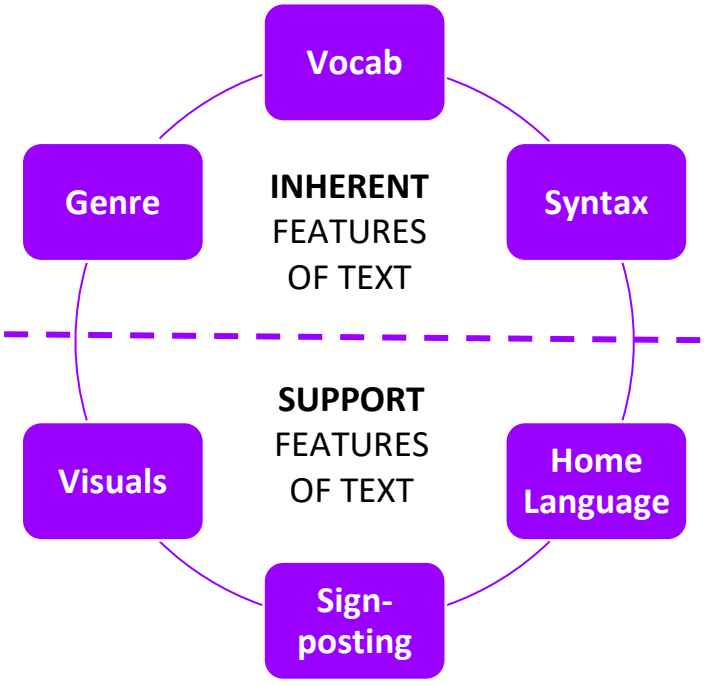


Figure 2: Inherent vs support features of texts

I will dedicate the remaining chapters to an in-depth evaluation of the three inherent features of texts – Vocabulary, Genre and Syntax. A review of the relevant literature and a detailed framework for each of these topics will be discussed in subsequent chapters. The three support features of texts – Visuals, Home Language and Sign-posting – will not be evaluated in this

study. However, a brief literature review pertaining to each of these can be found in Appendix A.

There is also a final group of factors not included in my framework, but which nonetheless have a significant impact on the accessibility of texts. We can group these factors together under the heading 'learner characteristics'. Learner characteristics refers to the natural diversity found among children which may mediate the extent to which a text is accessible for them. Different learners have different exposure to books and language, different amounts of prior knowledge regarding particular topics as well as different levels of interest in those topics, all of which affect their understanding of a given text. Dale and Chall (1949:22) explain that "when there is a strong interest in the subject matter, more effort is put forth and there is a greater amount of understanding."

I have not included the impact of learner characteristics as an area in my framework as it is not a feature of textbooks but rather a feature of readers. Nonetheless, the significance of learner-related factors should not be overlooked when evaluating text accessibility.

3.2. Methodology

This study employs content and discourse analysis. Content analysis is defined as "the study of recorded human communications, such as books, websites, paintings, and laws" (Babbie, 2014:341). Within this field, my study falls within a well-established tradition of analysing school textbooks and curriculum documents, where the aim is to evaluate the impact of such documents on educational outcomes. Typically such research focuses on particular subject areas and can make use of a variety of metrics for evaluation (Chu, 2017). Discourse analysis can be understood of the study of language above the sentence, or language in use, where the focus is on identifying "patterns (structure, organisation) in units which are larger, more extended, than one sentence" (Cameron, 2001:11).

The focus of my research is on evaluating the accessibility of one English Grade 4 NST textbook in relation to the language resources provided by a Grade 1-3 ELT scheme. The methodology employed in this study is mixed methods, with elements of both quantitative and qualitative data analysis featuring at different points.

Sampling

It should be noted that my study focused only on English language textbooks and is thus itself monolingual in its scope. To determine which school textbooks to include in the study, I made use of purposive sampling, "a type of nonprobability sampling in which the units to be observed are selected on the basis of the researcher's judgement about which ones will be the most useful or representative" (Babbie, 2014:200). The documents and books which formed part of content analysis are detailed in Table 1 below:

Category	Document/Book Name	Publisher	Scope of Analysis
ELT scheme	Platinum English First Additional Language Gr 1 Learner's Book	Maskew Miller Longman – Pearson	Entire book (111 pages, letter size)
ELT scheme	Platinum English First Additional Language Gr 2 Learner's Book	Maskew Miller Longman – Pearson	Entire book (107 pages, letter size)
ELT scheme	Platinum English First Additional Language Gr 2 Reader	Maskew Miller Longman – Pearson	Entire book (64 pages, A5 size)
ELT scheme	Platinum English First Additional Language Gr 3 Learner's Book	Maskew Miller Longman – Pearson	Entire book (124 pages, letter size)
ELT scheme	Platinum English First Additional Language Gr 3 Reader	Maskew Miller Longman – Pearson	Entire book (64 pages, A5 size)
Grade 4 Natural Sciences and Technology scheme	Platinum Natural Sciences and Technology Gr 4 Learner's Book	Maskew Miller Longman – Pearson	Entire book (202 pages, letter size)

Table 1: Sample of textbooks and readers used in analysis

In order to grasp the English language resources children are meant to develop in the Foundation Phase, it is necessary to understand how the course of study as envisioned by the DBE across Grades 1-3 (outlined only broadly in the CAPS documents) actually manifests in the Grade 1-3 EFAL LTSMs. These LTSMs (including the learner books for each grade as well as the accompanying readers available only for Grade 2 and 3) provide detailed examples of exactly the kind of English children are to be taught in the classroom as well as the manner in which this is to occur.

My rationale for selecting a Grade 4 NST textbook is based on the fact that a sociocultural approach to language requires a serious engagement with the varied nature of the discourses children are expected to gain proficiency in. I chose the content subject Natural Sciences and Technology because it requires learners to develop a large, specialised vocabulary and register to make sense of its content.

I chose the Platinum series of textbooks published by Maskew Miller Longman/Pearson based on the fact that the series is approved for use by the DBE and appears to be a popular textbook choice for many schools in both the Western Cape and the Eastern Cape. While this data sample can by no means be considered representative of all the textbooks available in schools, it is sufficient to begin exploring some of the language and text-related challenges EFAL learners contend with when they begin using English as their LOLT.

3.3. Limitations of Study

The limitations of this study are as follows. Firstly, I have selected only one ELT scheme and one NST textbook for analysis. There are a number of alternative ELT schemes and NST textbooks available on the market, approved by the DBE and in use in schools whose analysis may reveal different findings. Secondly, I have not conducted a detailed analysis of the kinds of activities and tasks that the materials require learners to perform. These activities and tasks are also an important aspect of the language and literacy demands made upon learners especially from a language production perspective. However, to evaluate these tasks and activities would have required fine-grained qualitative analysis of the materials that was beyond the scope of this study. Thirdly, I am analysing the language and literacy demands of only one content subject in Grade 4. This despite the fact that there are multiple content subjects taught in English in Grade 4, all of which make unique demands on the language and literacy resources of learners and all of which should be considered when evaluating the merits of the Foundation Phase EFAL course of study. Finally, as stated earlier, because of space constraints I will not be analysing any text support factors in the materials (such as visuals, use of home language or sign-posting) despite such factors having an important mediating influence on text accessibility.

The next three chapters will focus on the extent to which learners are prepared to cope with the vocabulary (Chapter 4), genres (Chapter 5) and syntax (Chapter 6) they will encounter in a Grade 4 NST textbook.

Chapter 4: Vocabulary

4.1. Framework

Vocabulary knowledge, also sometimes called *lexis,* refers to the knowledge of words in a language. Vocabulary plays a key role in successful reading comprehension. According to Stahl (2003:241) "[c]orrelations between measures of vocabulary and reading comprehension routinely are in the 0.90s. The correlations have been found to be robust almost regardless of the measures used or the populations tested." Laufer and Sim (1985) go so far as to say that vocabulary is the most pressing need of the foreign language reader. Because there are so many issues to consider when discussing vocabulary, this section will proceed by responding to a selection of questions which commonly occupy researchers in this field.

The first such question is 'How many words do learners need to know to understand English/ read in English/ communicate in English, etc.?' Before we can attempt to answer this question we must first deal with a more technical question, however, and that is, how do we count words?

There are four main ways that we can count words: as tokens, as types, as lemmas or as families. Nation (2001:7-8) provides a clear break down of each of these:

- **Tokens** ('running words')**:** Counting words as tokens is what we do whenever we produce a simple word count of a text. It involves counting every word in a text and if the same word appears more than once, then each occurrence of it is counted.
- **Types:** Counting words as types means we only count each word form once.
- **Lemmas:** Counting words as lemmas involves identifying a headword and combining it with some of the word's inflected forms (e.g. the plural, third person singular present tense, past tense, past participle, *-ing*, comparative, superlative and possessive forms) as well as the word's reduced (*n't*) forms, all of which are counted as the same word. Usually, all the items included under a lemma are the same part of speech.
- **Families:** Counting words as families is similar to counting words as lemmas in that a word family consists of a headword and its inflected forms. But in addition, word families also include a word's closely related derived forms, e.g. forms of the word with affixes such as *un-* and suffixes such as *-ly,* and *-ness*.

The rationale behind counting words as lemmas or families is that the inflections and derivations all share the same conceptual root and once learners master certain grammatical patterns they will automatically be able to understand and use words that are closely related to one another. Therefore, when trying to count how many words learners 'know' or need to learn, these related forms should be counted together.

The problem arises, however, when we try to determine what can be considered as 'closely related' forms of a word. Nation (2001:8) states that "[l]earners' knowledge of the prefixes and suffixes develops as they gain more experience of the language. What might be a sensible word family for one learner may be beyond another learner's present level of proficiency." The same can be said for inflections, especially when working with beginner language learners. Some inflections may be easy to recognise, such as the plural or regular simple past tense forms of a word. However, some verbs have irregular past tense forms. It then becomes a function of a learner's proficiency whether they will recognise the word 'went', for example, as the past tense form of the word 'go'. The danger of counting words as lemmas or families is that by doing so, one makes certain assumptions about the proficiency levels of those who will encounter these words which may not always be accurate.

Returning to the question of 'How many words do learners need to know?', I will focus on the number of words needed to be able to read in English. This question can be approached in two different ways. The first way is to ask, 'What percentage of words in a given text need to be known to be able to read that text?' The second approach would be to ask, 'How many English words are needed in absolute terms to be able to read in English?' The latter question, although intuitively more attractive to teachers and policy makers, is a problematic one given the socially situated nature of language and literacy resources, as discussed in Chapter 2. Nonetheless, it is one which has been asked and answered in the literature and merits attention. I will discuss both questions below.

According to Hu and Nation (2000), the majority of learners need to know around 98% of the running words (tokens) in a text to be able to read it unassisted, and this for texts that are not considered difficult (e.g. fiction texts with a strong chronological storyline). Knowing 98% of the tokens in a text equates to 1 in every 50 words of the text being unknown. Hu and Nation's

research indicates that as the density of unknown words increases, comprehension drops. They also found that while the reading skills and background knowledge of learners could to some degree compensate for limited vocabulary knowledge, this became impossible once learners knew less than 80% of the words in a text (i.e. 1 in every 5 words are unknown). A study by Laufer and Ravenhorst-Kalovski (2010) echoed some of these findings, stating that 98% coverage was the optimal threshold for text comprehension, but that 95% coverage (i.e. 1 in every 20 words is unknown) could be considered a minimally acceptable threshold, if one uses a more lenient definition of adequate comprehension. Furthermore, Nation (2001) found that if learners are reading a text where there is new vocabulary, at least 95% coverage is needed for learners to be able to use context clues to deduce the meaning of the new words. This information is summarised in Table 2 below:

Text Coverage	Unknown Words	Reading Comprehension
98% coverage	1 in 50 words unknown	Optimal threshold: majority of learners can read and comprehend the text unassisted
95% coverage	1 in 20 words unknown	Minimum threshold; also the minimum coverage required for learners to be able to work out the meaning of unknown words using context clues
90% coverage	1 in 10 words unknown	Only some learners can read and comprehend the text
80% coverage	1 in 5 words unknown	Impossible for learners to compensate for limited vocabulary knowledge with reading skills and background knowledge below this

Table 2: Thresholds of text coverage (% of words that must be known) to ensure text comprehension

With regards to vocabulary in absolute terms, Laufer (1992) estimated that knowledge of approximately 3000 word families (about 5000 word types) is needed for L1 readers to transfer their reading strategies to L2. "Until [learners reach this threshold], reading in L2 will be hampered by insufficient knowledge of vocabulary" (Laufer, 1992:130). Laufer's findings were based on an analysis of the reading comprehension and vocabulary test scores of 92 university first year ELLs whose home languages were Hebrew and Arabic. From the perspective of a social approach to language and literacy, it cannot be assumed that Laufer's findings will hold true for other language users reading different types of texts, especially given that my study focuses on early primary school learners of predominantly African home

languages. Nonetheless, Laufer's estimates provide a useful point of reference for comparison, both for my study and for the DBE's stated vocabulary targets.

According to CAPS (DBE, 2011), vocabulary acquisition targets for EFAL learners in South Africa are as follows: Grade 1 – 700-1000 words in context; Grade 2 – 1000-2000 words in context; Grade 3 – 1500-2500 words in context. If we understand words in context to be equivalent to word types, SA's ELLs will be below Laufer's threshold by between 2500 and 3500 words at the end of Grade 3, meaning that according to her predictions, they will be unable to read in a satisfactory manner in English. Moreover, this assumes that they meet their CAPS vocabulary targets. In reality, it is rare for students to have 100% mastery of all the material they have been taught. To establish the validity of these targets and also Laufer's estimates for the South African primary school context, I will compare both to the findings of my own study.

The next question which naturally emerges from the discussion around 'How many words do learners need to know?' is, 'What exactly constitutes knowing a word?' There are many things to know about a particular word. At the most basic level, to know a word we must know both its meaning and its form. Meaning can be understood as having a mental representation of what the word refers to (the underlying concept) but also extends to knowledge of which situations the word can or cannot be used in, and even to which other words a word is usually combined with (its collocations). For form, it is important to realise that a word's form differs for oral and written language (sounds versus symbols) and learners must know both to be able to read. To write and speak, they also need to be able to spell or pronounce the word. Furthermore, they must be able to manipulate word forms according to the grammatical rules of the language (which requires knowledge of a word's stem as well as its affixes) in order to be able to change words according to tense or number, or from noun into adjective, etc. (Nation, 2001; Carter, 2001)

Given the above, we cannot regard words as being simply known or unknown but instead need to see word knowledge as existing along a continuum. Waring and Takaki (2003:133) explain that word knowledge "could range from knowing only that you have seen or heard the word-form without being able to recall the meaning, to a full understanding of the word and its various nuances and use in a variety of contexts both receptively and productively."

In practice, most language teachers are interested in how best to teach word knowledge, and there are many debates surrounding this topic. Regardless of the strategy used, the most basic prerequisite of learning a word is encountering it, whether through hearing the word spoken or through reading the word. Such encounters may be incidental (i.e. occurring while a learner is focused on making sense of what they are hearing or reading) or they can be deliberate (where the goal is to acquire new vocabulary).

All types of encounters with words result in some degree of vocabulary acquisition. However, it is not easy to determine which types of encounters contribute most to learning. It is important to bear in mind that learning new words is an incremental process. Nation (2001:8) explains that "[a] word is not fully learned through one meeting with it, even if the meeting involves substantial deliberate teaching." The reason for this is at least partially attributable to the complexities inherent in learning words.

Not only are there many things to learn about each word, but there are many types of words and not all words are equally difficult to learn. Generally, concrete nouns are easier to learn than abstract ones, and are easier to learn than verbs (which have many inflections) or adverbs. Words will also pose different challenges to different learners. One aspect of this variability is described by Nation (2001) as a word's 'learning burden', i.e. the amount of effort required to learn a word. He states that learners with different language backgrounds will face different challenges. "For learners whose first language is not related to the second language, the learning burden will be heavy" (Nation, 2001:24). Other differences between individuals such as their level of language proficiency, motivation and aptitude can also affect acquisition. Word knowledge has also been found to decay over time (Waring and Takaki, 2003) so even if a learner knows a word today, that may not be the case several weeks, months or years from now.

Some researchers have focused specifically on vocabulary acquisition through reading. Research in this area has tried to determine how many times a reader must encounter a word in text to acquire it. The answers to this question have diverged greatly. Studies have reported as few as six encounters as sufficient to acquire a word (Saragi, Nation & Meister, 1978; Rott, 1999) while others report eight or more encounters are needed (Horst, Cobb & Meara, 1998), 10 or more are needed (Webb, 2007) or even as many as twenty to thirty encounters are needed

(Waring & Takaki, 2003). These varied findings may be attributable to differences in the ways the studies were conducted, such as the type and amount of exposure to words they allowed for, the type and timing of tests, the vocabulary being tested and differences between the populations of learners (their age, L1, English proficiency, etc.).

It is hard to adopt a specific position regarding the number of incidental exposures needed to acquire vocabulary for use in my own analysis given that none of the research used a population comparable to Foundation Phase learners with an African L1. However, Laufer and Rozovsky-Roitblat (2015:694) state that "literature provides sufficient evidence which shows that people acquire some vocabulary from reading only after six occurrences, particularly in the long term." This is likely to be an underestimation of the number of exposures needed for robust acquisition of words but is the position I will adopt.

Some researchers have contended that what matters most is not the number of times learners encounter a word but rather what learners do with words during those encounters. This is sometimes referred to as *depth of processing* (Craig & Lockhart in Laufer *&* Hulstijn, 2001). According to Schmidt (2001) at minimum, learners must *notice* words in order to learn them, where 'notice' refers to attending to a word with a degree of conscious awareness. This hypothesis has been verified by studies which found that "more vocabulary is acquired when the learners' attention is drawn to new words while reading than when attention is not drawn to it" (Laufer & Rozovski-Roitblat, 2015:690). This research implies that there is value in glossing or otherwise emphasising new vocabulary in texts to assist learners in acquiring words.

Baddeley (in Nation, 2001) maintains that words are more likely to be remembered if they are *retrieved* either receptively (hearing and seeing a word and having to retrieve the meaning) or productively (wishing to communicate the meaning of a word and having to retrieve the spoken or written form). It is important to note that retrieval does not occur if learners are presented with form and meaning simultaneously. Baddeley also argues that with repeated opportunities to retrieve a word, the pathway between form and meaning is strengthened, making subsequent retrievals easier. While based on depth of processing analysis, this argument supports the value of frequent encounters with words.

Generative processing has also been posited as an important factor in word acquisition (Joe, 1995). Generative processing takes place "when previously met words are subsequently met or used in ways that differ from the previous meeting with the word" (Nation, 2001:69). In this regard, words that appear frequently have a greater likelihood of being processed generatively than words which appear infrequently.

Despite the fact that there is no clear answer to the question 'Which types of encounters with words contribute most to acquiring them?' there are a few findings beyond what has already been discussed which are of particular significance to this study. The first is that the amount of vocabulary gained from incidental exposure is very modest in relation to the effort expended (Waring & Takaki, 2003; Zahar, Cobb & Spada, 2003). This means that vast amounts of reading or listening are needed to acquire a substantial amount of vocabulary, far more than seems feasible for ELLs in a classroom setting (Laufer & Rozovski-Roitblat, 2015). Furthermore, one's ability to acquire new vocabulary incidentally is a function of your existing knowledge base in that language, so for those who know very few words, such as beginner ELLs, it is very difficult to acquire new words incidentally (Waring & Nation, 2004).

Given the above, one might assume that deliberately teaching words to ELLs is a preferable strategy, and certainly this is true for especially beginner learners. However, the limitation of this method is that teaching even a single word in a classroom takes considerable time, and there are hundreds of thousands of words in the English language (Schmitt et al., 2017). So whether it is encountering words incidentally through reading or having teachers deliberately teach new words to learners, both strategies require a considerable time investment. Nation (2001) discourages teachers from relying too heavily on a single strategy and recommends a combination of incidental exposure, direct instruction, and a variety of support activities to encourage deep level processing in order to support vocabulary acquisition for learners.

Since it is clear that no matter which strategy is used, learning many new words will take time, it becomes important to ensure that the words learners do learn are the right kinds of words, i.e. words which will be useful to them. Which words are most useful will depend on the goals and purposes of the language user, but in general the most common way to assign value to a word is to establish the frequency with which that word is used in the language. This is determined by identifying which words occur most often in a carefully compiled corpus of

spoken or written language (usually a combination of both) which reflect a large variety of contexts. Words that appear often are known as High Frequency Words (HFWs) while words that appear seldom are known as Low Frequency Words (LFWs).

HFW lists traditionally consist of word families and are subdivided into the 1st thousand most frequent, 2nd thousand most frequent, 3rd thousand most frequent bands of words, and so on. It is usually the 1st and 2nd thousand most frequent bands that are of particular interest to researchers and educators as these usually account for the majority of words found in a text. There are many different frequency lists in circulation. Some of the most commonly used lists are the General Service List (GSL) which was compiled in 1953 (West, 1953), and the combined British National Corpus and Corpus of Contemporary American English (BNC-COCA) list compiled in 2012 (Nation, 2012). The latter corpus consists of approximately 660 million words compiled from a variety of spoken and written texts from the US and UK, including TV and radio shows, newspapers, popular magazines, fiction texts and academic texts.

There are also specialised frequency lists, such as the Academic Word List (AWL) compiled by Coxhead (2000). The AWL comprises 570 word families (outside of the two thousand most frequent words found in the GSL) that Coxhead found to be relatively well represented in academic texts across a number of subject disciplines. It should be noted that the AWL was developed using a corpus of texts aimed at university students (consisting of journal articles, university textbooks and lectures among other sources), and therefore the extent to which it represents vocabulary needed for academic study in primary and secondary schooling is unclear.

Whether these frequency lists are applicable to other English-speaking countries such as South Africa given their countries of origin (the US and UK) is also unclear. On the one hand, while English has regional variations, there are far more similarities in the language across regions than there are differences. Furthermore, a great deal of South Africa's English reading material is sourced from countries such as the US and UK, and so to some extent corpora like the BNC and COCA do reflect what South Africans are exposed to. However, from a spoken language perspective and for written material produced locally (such as for school textbooks), there are likely to be some divergences. A good example of this would be the word 'mum' which appears

in a British list of HFWs compiled using children's story books, but which would be unfamiliar to South African children who would instead be familiar with 'mom' or 'mama' (DBE, 2011). Since it cannot be assumed that such lists are applicable, their value needs to be investigated.

For the purposes of this study, I will examine how many words learners encounter in their EFAL materials in Grades 1 to 3, and also the frequency of these encounters. I will also compare the words that are encountered in the EFAL materials with the words found in the NST textbook, and both of these with the two thousand most frequent English words according to the GSL and the BNC-COCA list as well as the AWL. These comparisons will reveal how well the EFAL vocabulary prepares learners for their NST textbook and also to what extent international HFW are useful in determining the words that need to be taught to South African ELLs.

4.2. Methodology

In order to produce word lists for EFAL and NST, I transcribed the Grade 1, 2 and 3 Learner's Books and Grade 2 and 3 Readers published by Platinum for EFAL (there was no reader provided for Grade 1), as well as the Grade 4 NST textbook produced by Platinum into Microsoft (MS) Word documents. I did this without the use of specialised software. I excluded the front matter and contents pages of the books on the basis that young learners do not generally read these and began at the first page of the first lesson in each book. However, I did include glossaries and listening scripts provided at the back of the books if these were available.

After checking for any errors, I eliminated any parts of the books which had scrambled or incomplete words where learners had to rearrange or fill in blanks with letters. I also eliminated all numerals, inverted commas, ellipses and underscores, to reduce errors that would be generated by the software used to create the word lists. I left in all apostrophes. I then converted the documents into .txt files. The three learner books and two readers for EFAL were consolidated into a single .txt file (consisting of 34 743 tokens) while the NST textbook formed its own .txt file (consisting of 41 160 tokens).

I used the frequency software available on the Compleat Lexical Tutor website (available at https://www.lextutor.ca/freq/) to generate word lists for each .txt file. The frequency software generates lists of word types in order of the frequency of their appearance. One limitation of

the software is that it encodes punctuation (such as apostrophes) erroneously, but using a debugging chart, it is possible to correct these errors. The lists were pasted into MS Word documents to make these corrections. This means that my lists contained contractions such as 'I've' and 'didn't' as well as possessives (e.g. grandfather's). Another limitation of the software is that it is not possible to distinguish between homonyms. After debugging, the lists were then copied into separate MS Excel sheets.

From the excel sheets, I eliminated all single letter entries except 'a' and 'I'. I removed all proper nouns that were names and surnames of people, but left in the names of continents, countries, oceans, seas, mountain ranges, deserts, rivers, major cities and South Africa's provinces on the basis that these are words which will need to be read throughout a learner's school career. For the same reason, I kept the names of famous organisations (e.g. NASA), religious holidays (e.g. Hanukkah) as well as the names of languages and people groups (e.g. Xhosa, Zulu, San, etc.). I also retained several words which, although not English in origin, have become part of South Africa's English lexicon, such as 'gogo', 'rondavel', 'pap' and 'samoosas'. I eliminated made-up words in rhymes, poems and stories which were used for sound effects (e.g. inky pinky ponky, slishy, grrrrrr) but retained sound effects which have become conventionalised in their usage (e.g. ouch, oh). I ensured that where words were hyphenated (e.g. ice-cream), they were done so consistently throughout the text, and in both the EFAL and NST materials. I also retained abbreviations such as 'km', 'cm', 'vs' and 'etc.'. The resulting lists consisted of 2683 word types (33 304 tokens) for the EFAL and 2677 word types (39 857 tokens) for the NST materials.

I then used MS Excel operations to compare the EFAL and NST word lists and identify which of the EFAL word types appeared in the NST list. In order to calculate how many tokens in the NST textbook readers would be familiar with based on the vocabulary encountered in the EFAL materials (i.e. text coverage) I took the total number of tokens from the NST textbook and then calculated what percentage of these tokens consisted of words encountered in EFAL materials.

In order to conduct my analysis on HFWs, I needed to convert my word type lists into word family lists. This necessitated the elimination of the following previously retained items from the lists: all proper nouns, uniquely South African words, sound effects and abbreviations. This

was necessary because the software used to convert word types into word families would not recognise these categories of words. Contractions and compound words also had to be expanded into their composite words to be read by the software. The edited word lists for both EFAL and NST were then analysed using the Vocabprofile software on the Compleat Lexical Tutor website (available at https://www.lextutor.ca/vp/).

I analysed the edited lists using both the VP Classic programme and the VP Compleat programme. The distinction between the two is that VP Classic focuses on the first and second thousand most frequent word families according to the GSL, along with the 570 word families found in the AWL. VP Compleat focuses on the 25 000 most frequent words (in descending order of frequency) according to the BNC-COCA corpus. The benefit of conducting analysis using both the VP Classic and VP Compleat programmes is that this allows for a comparison of results using different HFW corpuses to allow for a degree of verification of the findings.

An additional area of the HFW analysis necessitated analysing the entire running text of the NST textbook using the same programmes. Unfortunately, the full NST textbook exceeded the software's word limit. To bypass this limitation, I divided the NST textbook into two halves and analysed each half separately. The results of my vocabulary analysis are detailed below.

4.3. Results

Word Types

Of the 2677 word types that appear in the NST textbook, 1398 word types do not appear in the EFAL materials at all. In other words, 52% of the word types found in the Grade 4 NST textbook are words learners have not encountered even once in three years of EFAL materials. Some examples of words which occur with high frequencies in the NST textbook but do not appear at all in the EFAL materials include *energy*, *heat*, *structure*, *habitat, system*, and *explain,* to name but a few. Beyond these unseen words, a further 668 word types found in the NST textbook (25% of the total number of word types found in the book) appear only 1 to 5 times in the EFAL materials while 445 word types (17%) appear between 6 and 29 times. Only 166

word types (or 6%) in the NST textbook appear 30 or more times in the EFAL materials. This information is displayed in Figure 3 below.

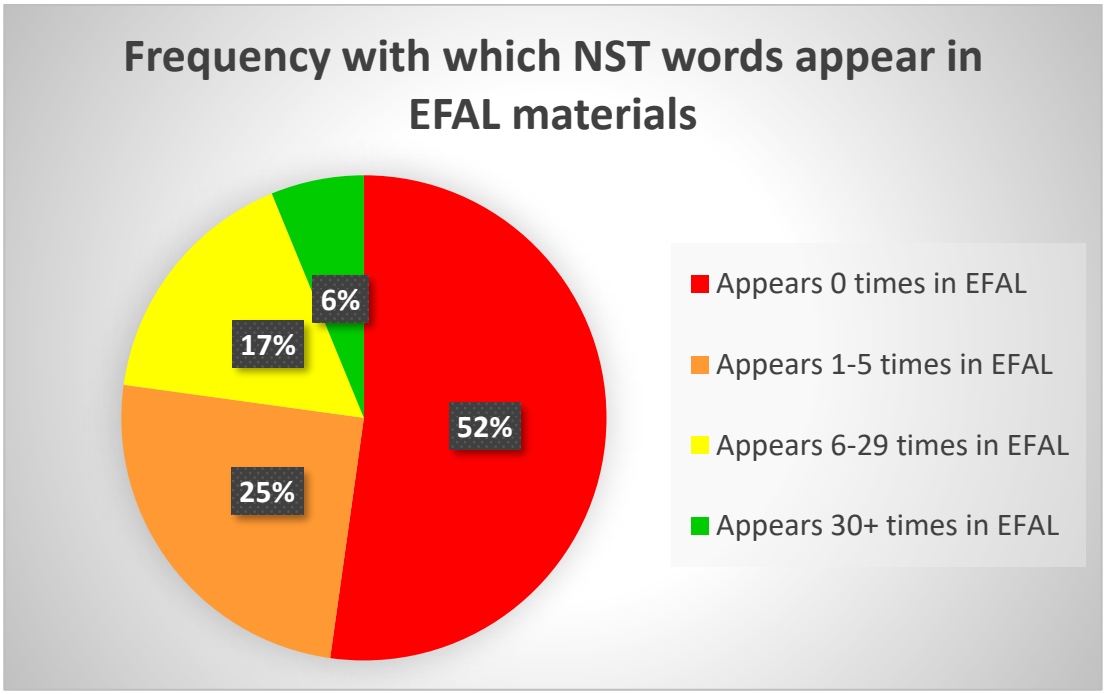


Figure 3: Frequency with which NST words appear in EFAL materials

The frequency brackets of 1-5 times, 6-29 times and 30+ times correspond to the literature regarding number of incidental exposures needed to acquire a word through reading. From the above we can see that (based on learners' exposure to words in the EFAL materials) 52% of the word types found in the NST textbook cannot be known, a further 25% have a very small probability of being known (together comprising 77% of all word types in the book), 17% of the word types may or may not be known (depending on what number of exposures you consider necessary for word acquisition) while only 6% of the word types will almost certainly be known regardless of which study's findings you support.

Text Coverage

As a percentage of the total running text in the NST textbook, unknown word tokens make up 18% of the book. This means that if learners learn every single word that appears in their EFAL materials (even those that appear only once or twice in the texts throughout three years of study), they would still only be able to read 82% of the running text in their NST textbook.

Another way of stating this is to say that the EFAL materials provide only 82% text coverage for the NST textbook. This is well below the threshold of 98% text coverage needed for independent reading. It is also below the 95% threshold needed for minimal comprehension, or to deduce the meanings of unknown words. Perhaps most importantly, this text coverage is alarmingly close to the threshold of 80% below which no learners can comprehend a text, regardless of how well-developed their reading skills.

If we do not assume that learners learn every word they encounter in EFAL (which is more probable) the figures become more problematic. Table 3 below provides a breakdown of text coverage (i.e. percentage of tokens from the running text that will be known by readers) as well as percentage of unknown text, based on the number of incidental encounters with vocabulary in the EFAL materials.

WORD TOKENS	NO. of TOKENS	% TEXT COVERAGE	% UNKOWN TEXT
Not seen in EFAL	**7043**	**82%**	**18%**
Seen <2 times in EFAL	8613	78%	22%
Seen <3 times in EFAL	9544	76%	24%
Seen <4 times in EFAL	10676	73%	27%
Seen <5 times in EFAL	11373	71%	29%
Seen <6 times in EFAL	**12086**	**70%**	**30%**
Seen <10 times in EFAL	14181	64%	36%
Seen <15 times in EFAL	15199	62%	38%
Seen <20 times in EFAL	16626	58%	42%
Seen <30 times in EFAL	**18590**	**53%**	**47%**
TOTAL TOKENS IN S&T	39857		

Table 3: NST coverage provided by EFAL vocabulary

The first row which relates to words 'Not seen in EFAL' displays the data already discussed. The second row displays how many word tokens in the NST textbook appear less than twice (i.e. once or not at all) in the EFAL materials. If we assume that learners will not know words encountered fewer than two times in their EFAL materials, then 22% of the running text in NST textbook will be unknown to them (a text coverage of 78%). The rows that follow display the same information for different numbers of encounters with words in the EFAL materials.

I have provided this breakdown because the number of encounters required to learn a word is contested. If we use the number of encounters recommended by Laufer and Rozovsky-Roitblat

(2015), namely six or more times (which is one of the more generous estimates in the literature), then 30% of running text in the NST textbook will be unknown to learners. That is to say, 3 out of every 10 words learners read in their textbook will be unknown. On the other end of the spectrum, the strictest estimate for number of encounters needed to learn a word is 30 or more times. By this estimate, approximately half the NST textbook (i.e. 1 in every 2 words) will be unknown to learners.

The data provided in Table 3 is averaged across all the text in the NST textbook. However, if we take a closer look at particular topics and units within the textbook (each with their own thematic focus and corresponding specialised vocabulary) the figures fluctuate somewhat. There are a total of 17 topics found in the NST textbook, each with between one and five units, spread across four school terms. The topics are as follows:

Term 1	Term 2	Term 3	Term 4
1. Living and non-living things	6. Materials around us	10. Energy and energy transfer	14. Planet Earth
2. Structures of plants and animals	7. Solid materials	11. Energy around us	15. The Sun and its effect on Earth
3. What plants need to grow	8. Strengthening materials	12. Movement and energy in a system	16. The Moon
4. Habitats of animals	9. Strong frame structures	13. Energy and sound	17. Rocket systems
5. Structures of animal shelters			

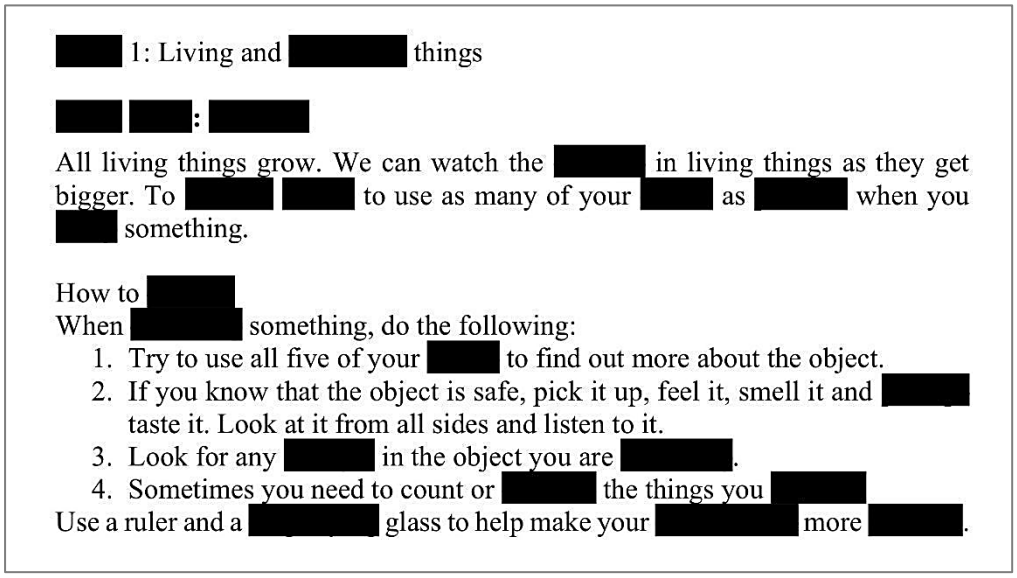
Table 4: Topics found in the NST textbook

Unit 1 of Topic 1 (studied at the start of Term 1 in Grade 4) is entitled ‘Living things’. For this unit, the EFAL vocabulary provides a coverage of 88% of the text. This is much better than the average coverage for the textbook, and in this sense Unit 1 of Term 1 may be considered a relatively easier unit for ELLs to comprehend. Even so, the text coverage provided by the EFAL vocabulary for this unit is still below the threshold needed for comprehension (95-98% coverage).

In comparison, Unit 1 of Topic 17 (studied towards the end of Term 4) is entitled 'Rocket systems'. For this unit, the EFAL vocabulary only provides a coverage of 74% of the unit's text. This is worse than the average coverage of the textbook and is below the cut-off threshold of 80% described earlier, meaning the text is incomprehensible for learners no matter how advanced their reading skills. The takeaway here is that the degree to which the NST textbook is incomprehensible to learners will vary based on the topic being taught, but the comparisons are between bad and worse.

Some of the unknown vocabulary found in the NST textbook is glossed in 'Key Word' boxes found in the margin at the start of certain units. Across the whole textbook, there are a total of 133 glossed terms. Although the literature indicates that glossing terms can aid learner comprehension of a text, this is only the case if the glosses themselves are comprehensible. Unfortunately, many of the glosses in the NST textbook contain unknown words in their definitions, severely limiting the extent to which they can aid learner comprehension.

To appreciate what low text coverage combined with unhelpful glosses looks like for a specific passage of text, I have included two short excerpts from Topic 1 and Topic 17, pictured below. I have blacked out the unknown vocabulary for the passages and for the second excerpt, highlighted the glossed term in grey. Figure 4 (pictured below) shows an excerpt with a text coverage of 84%, taken from Topic 1, while Figure 5 (see next page) shows an excerpt with a text coverage of 71% taken from Topic 17 .


[redacted] 1: Living and [redacted] things

[redacted] [redacted]: [redacted]

All living things grow. We can watch the [redacted] in living things as they get bigger. To [redacted] [redacted] to use as many of your [redacted] as [redacted] when you [redacted] something.

How to [redacted]
When [redacted] something, do the following:

1. Try to use all five of your [redacted] to find out more about the object.
2. If you know that the object is safe, pick it up, feel it, smell it and [redacted] taste it. Look at it from all sides and listen to it.
3. Look for any [redacted] in the object you are [redacted].
4. Sometimes you need to count or [redacted] the things you [redacted]

Use a ruler and a [redacted] glass to help make your [redacted] more [redacted].

Figure 4: Example of 84% text coverage

[redacted] 17: Rocket [redacted]

Starting off

People have [redacted] with rocket [redacted] for many years. They have used rockets to make [redacted] [redacted] and rocket-[redacted] vehicles. But most [redacted], people have used rocket [redacted] to travel into space. Rockets carry people, [redacted] and other [redacted] into space. People have used rockets to travel to the Moon, and rockets have even been used to get special [redacted] vehicles called [redacted] onto [redacted]. In this [redacted], you will find out more about how rocket [redacted] work and you will make a rocket [redacted] yourself.

Key words

rocket: long tall [redacted] that can be [redacted] into space

Figure 5: Example of 71% text coverage

These passages are very hard to make sense of, even for skilled readers whose L1 is English, never mind inexperienced readers who have English as their L2 or L3.

Word Families and High Frequency Words

Thus far I have compared the EFAL materials with the NST textbook, and in doing so excluded much of the vocabulary appearing in the EFAL materials. On its own, the EFAL scheme exposes learners to a total of 2683 word types, which translates into 1568 word families (see methodology for how the number of word families was determined). This amount (1568 word families) is only half of what Laufer (1992) estimates is needed for ELLs to read successfully in English.

Because learners are being exposed to a limited number of words in their EFAL materials, it becomes especially important to determine whether these word families are the most useful ones learners could be acquiring to prepare them for further study through the medium of English. One way to ascertain this is to examine whether the word families encountered in the EFAL materials are High Frequency Words.

To determine this, I compared the word families found in the EFAL list with two well-regarded international HFW lists. The first of these was the two thousand most frequent words according to the General Service List, combined with the 570 word families found in the Academic Word List (VP Classic programme). The second comparison was with the twenty-five thousand most

frequent word families found in the BNC-COCA corpus (VP Compleat programme). The results of these comparisons are summarised in the tables below. The term K-1 in the tables refers to the first thousand most frequent words according to the list, K-2 refers to the second thousand most frequent words, K-3 to the third thousand most frequent words, and so on.

EFAL List compared with the General Service List (GSL) and Academic Word List (AWL)						
	Families (EFAL)		Types (EFAL)		Tokens (EFAL)	
	Number	%	Number	%	Number	%
GSL K-1 Words (1-1000)	632	40%	1211	48%	1239	48%
GSL K-2 Words (1001-2000)	445	28%	697	28%	714	28%
AWL Words [570 fams]	39	3%	49	2%	51	2%
Off-List Words	452	29%	553	22%	561	22%
	1568	100%	2510	100%	2565	100%

Table 5: Amount of EFAL vocabulary found in the GSL and AWL

Table 5 shows that out of the 1568 word families found in the EFAL list, 632 of them (40%) can be found in the GSL's first thousand most frequent word families. A further 445 of the EFAL word families (28%) can be found in the GSL's second thousand most frequent word families. Together these add up to 1 077 word families (68%). Only 39 of the EFAL word families (3%) were found in the AWL.

EFAL list compared with the BNC-COCA K-1-25 List							
Freq. level	Families (EFAL)		Types (EFAL)		Tokens (EFAL)		
	Number	%	Number	%	Number	%	Cumul. %
BNC-COCA K-1	764	49%	1464	58%	1503	59%	58.6
BNC-COCA K-2	348	22%	504	20%	514	20%	78.6
BNC-COCA K-3	75	5%	93	4%	93	4%	82.2
BNC-COCA K-4	113	7%	139	6%	140	5%	87.7
BNC-COCA K-5	85	5%	106	4%	106	4%	91.8
BNC-COCA K-6	63	4%	72	3%	74	3%	94.7
BNC-COCA K-7	35	2%	39	2%	39	2%	96.2
BNC-COCA K-8	22	1%	25	1%	25	1%	97.2
BNC-COCA K-9	19	1%	21	1%	21	1%	98
BNC-COCA K-10	13	1%	14	1%	16	1%	98.6
BNC-COCA K-11	8	1%	9	0%	9	0%	99
BNC-COCA K-12	4	0%	4	0%	4	0%	99.2
BNC-COCA K-13	3	0%	3	0%	3	0%	99.3
BNC-COCA K-14	4	0%	4	0%	4	0%	99.5
BNC-COCA K-15	-		-		-		
BNC-COCA K-16	4	0%	4	0%	4	0%	99.7
BNC-COCA K-17	2	0%	2	0%	2	0%	99.8
BNC-COCA K-18	2	0%	2	0%	2	0%	99.9
BNC-COCA K-19	1	0%	1	0%	1	0%	
BNC-COCA K-20	2	0%	2	0%	2	0%	
BNC-COCA K-21	-		-		-		
BNC-COCA K-22	1	0%	2	0%	2	0%	100
BNC-COCA K-23	-		-		-		
BNC-COCA K-24	-		-		-		
BNC-COCA K-25	-		-		-		
Off-List	0	0%	0	0%	0	0%	
Total	1568	100	2510	100	2564	100	≈100.00

Table 6: Amount of EFAL vocabulary found in the BNC-COCA K-1-25 list

Table 6 shows that of the 1568 word families found in the EFAL list, 764 of them (49%) can be found in the BNC-COCA first thousand most frequent word families. A further 348 of them (22%) can be found in the BNC-COCA second thousand most frequent word families. Cumulatively, these add up to 1112 word families (71%). The remaining EFAL word families are predominantly concentrated in the BNC-COCA K-3 to K-10 frequency bands (a total of 425 word families, 27%) with only 31 word families (2%) found across the K-11-22 frequency bands. None of the EFAL word families appeared in the K-15, K-21 or K-23-25 frequency bands.

The implications of the above findings are as follows: approximately a third of the 1568 word families learners encounter in their EFAL materials do not fall within the two thousand most frequently used English words, according to both the GSL and the BNC-COCA list and thus cannot be considered high frequency English words. Furthermore, only 39 of the EFAL word families can be found in the Academic Word List. This means that out of a possible 2570 high frequency word families, South Africa's EFAL learners are only being exposed to approximately 1000 word families.

Given that these lists were not compiled locally, does it matter that South African learners are not learning the word families found in these lists in their Grade 1-3 EFAL materials? To answer this question, I hypothesised the following: If EFAL learners in Grade 4 did know all the word families in the different lists, how much of their NST textbook would they be able to read? If the HFW lists do not provide good coverage of a South African Grade 4 textbook, then perhaps the lists are not relevant for our context.

It was not possible to analyse the entire NST textbook at once because of software limitations (see the Methodology section), so the textbook was divided into two halves, and the results for each half are shown below.

NST Full Text (1st half): General Service List (GSL) and Academic Word List (AWL)					
Freq. Level	Families (NST)	Types (NST)	Tokens (NST)	Text Coverage %	Cumul. Coverage %
GSL K-1 (1-1000)	559	907	15944	79.4	79.4
GSL K-2 (1001-2000)	289	407	1989	9.9	89.3
AWL [570 fams]	70	110	632	3.1	92.4
Off-List		446	1512	7.53	99.93
Total	918	1871	20079	100	≈100.00

Table 7: NST full text (1st half) according to GSL and AWL

Table 7 shows that knowledge of the first two thousand most frequent words of the GSL combined with knowledge of the AWL provides a text coverage of around 92% of the first half of the NST textbook. This means that if learners know the two thousand most frequent word families from the GSL and the 570 word families in the AWL, only 8% of the running text in the first half of their NST textbook would be unknown. The vast majority of the coverage

(79.4%) comes from knowledge of the first thousand most frequent word families, while knowledge of the second thousand most frequent provides an additional 9.9% of coverage. The AWL only provides 3.1% coverage of the first half of the NST textbook.

NST Full Text (2nd half): General Service List (GSL) and Academic Word List (AWL)					
Freq. Level	Families (NST)	Types (NST)	Tokens (NST)	Text Coverage %	Cumul. Coverage %
GSL K-1 (1-1000)	576	998	17479	83.7	83.7
GSL K-2 (1001-2000)	250	351	1376	6.6	90.3
AWL [570 fams]	81	113	965	4.6	94.9
Off-List		339	1064	5.09	99.99
Total	907	1802	20885	100	≈100.00

Table 8: NST full text (2nd half) according to GSL and AWL

Table 8 shows the same information but for the second half of the textbook. Knowledge of the two thousand most frequent words of the GSL combined with knowledge of the AWL would provide a text coverage of around 95% for the second half of the NST textbook. This means that only 5% of the running text for the second half of the NST textbook would be unknown. This reaches the minimal threshold of comprehension discussed earlier. Again, the majority of this coverage (83.7%) comes from knowledge of the first thousand most frequent word families, while knowledge of the second thousand most frequent provides an additional 6.6% of coverage. The AWL provides 4.6% coverage of the second half of the NST textbook. If we average the text coverage of the two halves provided by the two thousand most frequent words in the GSL and the 570 words in the AWL, we end up with a text coverage of 93.5% as compared with the 82% coverage provided by the 1568 word families in the EFAL materials.

NST Full Text (1st half): BNC-COCA K-1-25 List					
Freq. Level	Families (NST)	Types (NST)	Tokens (NST)	Text Coverage %	Cumul. Coverage %
BNC-COCA K-1	604	957	15790	78.2	78.2
BNC-COCA K-2	277	410	2214	11.0	89.2
BNC-COCA K-3	86	122	774	3.8	93
BNC-COCA K-4	74	96	451	2.2	95.2
Coverage 95					
BNC-COCA K-5	48	65	207	1.0	96.2
BNC-COCA K-6	34	39	111	0.5	96.7
BNC-COCA K-7	18	25	141	0.7	97.4
BNC-COCA K-8	21	28	63	0.3	97.7
BNC-COCA K-9	16	24	61	0.3	98
Coverage 98					
BNC-COCA K-10	9	10	18	0.1	98.1
BNC-COCA K-11	10	12	38	0.2	98.3
BNC-COCA K-12	6	6	13	0.1	98.4
BNC-COCA K-13	3	3	8	0.0	
BNC-COCA K-14	3	3	7	0.0	
BNC-COCA K-15	5	5	7	0.0	
BNC-COCA K-16	2	4	12	0.1	98.5
BNC-COCA K-17	2	3	11	0.1	98.6
BNC-COCA K-18	1	1	2	0.0	
BNC-COCA K-19	-	-	-		
BNC-COCA K-20	4	5	10	0.0	
BNC-COCA K-21	-	-	-		
BNC-COCA K-22	1	1	9	0.0	
BNC-COCA K-23	-	-	-		
BNC-COCA K-24	-	-	-		
BNC-COCA K-25	2	2	8	0.0	
Off-List	??	82	230	1.14	99.74
Total	1226+?	1903	20185		≈100.00

Table 9: NST full text (1st half) according to BNC-COCA K-1-25

Table 9 looks at the most frequent words of the BNC-COCA list. Knowledge of the two thousand most frequent word families would provide a text coverage of 89% for the first half of the NST textbook. This means that only 11% of the running text in the first half of the NST textbook would be unknown to learners who knew the two thousand most frequent words of the BNC-COCA list. Again, the majority of this coverage (78.2%) comes from knowledge of the first thousand most frequent word families, while knowledge of the second thousand most frequent provides an additional 11% of coverage. The third thousand and fourth thousand most

frequent words would provide an additional 3.8% and 2.2% coverage respectively, which would reach the minimal threshold of comprehension, 95% coverage.

NST Full Text (2nd half): BNC-COCA K-1-25 List					
Freq. Level	Families (NST)	Types (NST)	Tokens (NST)	Text Coverage %	Cumul. Coverage %
BNC-COCA K-1	610	1018	17169	82.2	82.2
BNC-COCA K-2	249	365	2287	11.0	93.2
BNC-COCA K-3	86	112	542	2.6	95.8
Coverage 95					
BNC-COCA K-4	68	86	436	2.1	97.9
Coverage 98					
BNC-COCA K-5	30	44	117	0.6	98.5
BNC-COCA K-6	14	16	29	0.1	98.6
BNC-COCA K-7	14	14	27	0.1	98.7
BNC-COCA K-8	7	8	18	0.1	98.8
BNC-COCA K-9	3	5	7	0.0	
BNC-COCA K-10	4	5	8	0.0	
BNC-COCA K-11	6	6	19	0.1	98.9
BNC-COCA K-12	3	3	11	0.1	99
BNC-COCA K-13	3	3	4	0.0	
BNC-COCA K-14	1	1	2	0.0	
BNC-COCA K-15	2	2	4	0.0	
BNC-COCA K-16	1	1	1	0.0	
BNC-COCA K-17	2	3	5	0.0	
BNC-COCA K-18	1	1	2	0.0	
BNC-COCA K-19	-	-	-		
BNC-COCA K-20	1	1	1	0.0	
BNC-COCA K-21	-	-	-		
BNC-COCA K-22	1	1	1	0.0	
BNC-COCA K-23	-	-	-		
BNC-COCA K-24	-	-	-		
BNC-COCA K-25	3	3	4	0.0	
Off-List	??	104	191	0.91	99.91
Total	1109+?	1802	20885		≈100.00

Table 10: NST full text (2nd half) according to BNC-COCA K-1-25

Table 10 shows the same information as Table 9, but for the second half of the textbook. For this half, knowledge of the two thousand most frequent word families would provide a text coverage of 93%. This means that only 7% of the running text in the second half of the NST textbook would be unknown to learners familiar with the BNC-COCA's two thousand most

frequent words. The first thousand most frequent words provide a coverage of 82.2%. The second thousand most frequent words provide an additional 11% of coverage. The third thousand most frequent words provide 2.6% coverage, reaching the minimal threshold of 95% coverage. If we average the coverage provided by the two thousand most frequent words in the BNC-COCA list for the two halves for the NST textbook, we achieve a text coverage of 91%, as compared with the 82% text coverage given by the 1568 word families in the EFAL materials.

From the above data, we can draw a number of conclusions. Firstly, the international frequency lists do provide good coverage of a South African Grade 4 NST textbook. These lists are therefore relevant to the South African context and can be used as a guideline for determining which words should be taught to ELLs. However, if a South African English frequency list could be generated, such a list might prove even more useful for guiding the learning of South African ELLs.

A second conclusion that we can draw is that knowledge of the first thousand most frequent English word families according to the GSL provides an average coverage of 81.6% for the NST textbook, which is only marginally less than the 82% coverage provided by the 1568 word families found in the EFAL materials. This figure is only slightly lower for the BNC-COCA list, whose first thousand most frequent words provide an average coverage of 80.2% of the NST textbook. In short, using either list, the thousand most frequent English word families provide approximately the same coverage as the 1500 word families found in the EFAL materials. This implies that the EFAL materials are not exposing children to the most strategic set of vocabulary when measured against what is needed to access their NST textbook.

A third conclusion we can draw relates to the Academic Word List. Knowledge of the AWL (570 word families) provides an average coverage of 3.9% of the NST textbook, which is marginally more than the average coverage provided by the third thousand most frequent words according to the BNC-COCA list (3.2%). Even though the total coverage provided is small, given that the AWL is 570 word families, as compared with an additional thousand word families in the 3-K band of the BNC-COCA list, there seems to be merit in focusing on learning the AWL word families after having mastered both the first and the second thousand most frequent words in English. Also, since the AWL was generated using a corpus of language

predominantly drawn from university textbooks and lectures, it is almost certain the 570 word families found in this list will become increasingly valuable for learners as they progress through schooling.

Finally, based on where the comprehension thresholds fall in the HFW lists, the data appears to support Laufer's assertion that knowledge of around three thousand word families (approximately 5000 word types) is the amount of vocabulary needed to enable students to study in English, so long as those words are the most frequently used words in English. This reaffirms the fact that the CAPS vocabulary targets for EFAL (1500-2500 word types by the end of Grade 3) are completely insufficient in relation to the minimum vocabulary needed to study through the medium of English.

Chapter 5: Genre

5.1. Framework

According to Hammond and Derewianka (2001:186) who draw on a Systemic Functional Linguistics (SFL) orientation to language, "[t]he term genre is used in various educational contexts to refer to the recognisable and recurring patterns of everyday, academic and literary texts that occur within particular cultures." Texts here can be understood as referring to any meaningful stretch of language – oral or written (Derewianka, 1990). According to this view, texts are structured in different ways to achieve various purposes. For example, the purpose of *Instructions* is to tell someone how to do or make something, while the purpose of a *Recount* is to tell what happened. Each of these genres has its own schematic structure and grammatical constructions which typify it, and each makes use of specialised terms as well as other patterns of language to help it achieve its purposes. (Derewianka, 1990)

Genre is important to consider when determining the accessibility of texts (particularly those intended as learning materials) because genres give us frameworks to structure and interpret texts. This enables us to better cope with the textual demands of a given piece of writing or speech (Seligmann, 2012). In other words, the degree to which readers have mastery of the genre of a particular text informs the degree to which that text is readable for them. This is also key when we consider the need for students to internalise what they read and potentially reproduce or transform the information.

> When a reader knows a specific text structure, [s/he] is said to have acquired the appropriate textual schema for that structure. … These textual schemata allow the reader to organise the information presented to him hierarchically, and to form a global summary of the information in the text. Such a global view of the text is necessary for good comprehension. … Without knowledge of the text structure the reader will perceive all information as equally important and will only be able to produce a list of the points made. (Van Rooyen, 1990:15)

There is no one, universal list of genres since by their very nature, genres lend themselves to approximate rather than precise categories. As such, different authors subdivide them differently. Moreover, "[a]s the community's purposes grow and change, new genres arise. And with increasing complexity of purpose come increasingly complex genres – hybrid genres, genres within genres, subversive genres, and so on." (Derewianka, 2012:131) Nonetheless,

there is some consensus on broad categories of genre, especially within texts read by school children. The FP EFAL CAPS document makes provision for the following text types to be covered in the Foundation Phase: personal and factual recounts; procedures (instructions); information reports; and narratives (stories) (DBE, 2011:90).

Because the latter is a fairly limited list of genres which fails to account for the variety that exists within the language and content textbooks to be analysed, this study will instead draw on the categories of genres compiled by Derewianka (2003). She highlights eight broad categories of genre which are critical to success in educational contexts: *Description*, *Information Report*, *Procedure*, *Explanation*, *Story Genres*, *Response Genres* and *Exposition* (refer to Appendix C for more details).

Descriptions provide information about a particular person, place or thing, and can be subdivided into *objective descriptions* and *literary descriptions*. *Information Reports* serve a similar function but are distinguished by the fact that they provide information about a class of things as opposed to an individual or particular thing. *Information Reports* are abundant in expository textbooks. (Derewianka, 2003)

Procedures tell someone how to do something. Examples of sub-genres within this category include *instructions*, *experiments*, *directions* and *regulations*. *Recounts* focus on relating events that have happened and can include *personal recounts, factual recounts, historical recounts* as well as *biographical* and *autobiographical recounts*. (Derewianka, 2003)

Explanations serve to explain how or why a phenomenon takes place, and sub-genres of this category include *sequential, causal, factorial,* and *consequential explanations* as well as *explorations*. *Story Genres* refers to texts whose purpose is to explore the human condition through storying and includes the more commonly known sub-genres of *narratives, moral tales*, *fables, anecdotes*, and *plays*. (Derewianka, 2003)

Response Genres are texts whose aim is to respond to a literary text or artistic work and can take the form of *personal responses*, *reviews* and *interpretations*. *Exposition* (i.e. expository texts) are those texts which mount an argument, including *critical responses, persuasion,* and *discussions* or *debates*. (Derewianka, 2003)

In addition to the eight genres specified by Derewianka, I will include two additional categories of genre in my framework. The first of these is *Lyrical Genres*. This genre describes texts where the language is characterised by the use of rhyme, rhythm or sound devices. *Songs, rhymes* and *poetry* are all sub-genres that fall within this category. Such texts occur frequently in the EFAL textbooks and are not clearly accounted for by any of the other genres specified by Derewianka.

My second addition to the framework will be a category for *Definitions*. Although definitions could be considered a sub-category of *Descriptions*, they occur often enough in the materials to merit their own treatment. The purpose of a *Definition* is to define and/or demonstrate the meaning or use of a specific word or phrase in context. The full genre framework I will be using for my study is summarised in Table 11 below:

GENRE	PURPOSE
1. Description	To provide information about a particular person, place or thing
2. Information Report	To provide information about a class of things
3. Procedure	To tell someone how to do something
4. Recount	To tell what happened
5. Explanation	To explain how or why a phenomenon takes place
6. Story Genres	To explore the human condition through storying
7. Response Genres	To respond to a literary text or artistic work
8. Exposition	To mount an argument
9. Lyrical Genres	To entertain or instruct an audience using language characterised by rhyme, rhythm and other sound devices
10. Definitions	To define or demonstrate the meaning or use of a word or phrase

Table 11: Summary of genre framework

5.2. Methodology

For the purposes of this study, I am interested in the range and extent of exposure to particular genres that learners receive in their EFAL studies in the Foundation Phase (i.e. which genres learners encounter and how much exposure they have to each genre in the Grade 1-3 EFAL LTSMs) and how these compare with the genre demands of a Grade 4 NST textbook.

To generate this data I identified discrete reading texts within the Grade 1-3 EFAL LTSMs and the Grade 4 NST textbook and coded each of these as belonging to one or more genres, using my genre framework. The length of each of these texts (measured in word tokens) was also recorded. In those cases where texts were coded as more than one genre, the word count of the text was split equally between the relevant genres. I also sent a sample of texts along with my coding framework to two other individuals and asked them to classify the texts using the framework. I then compared their results to my own as a check for reliability.

For selected genres which occurred within both the EFAL LTSMs and the NST Learner's Book, I selected example texts from both sets of materials and made detailed annotations on the genre features of each (following the style of Derewianka, 1990). These annotations are presented and discussed side-by-side in order to compare and contrast how the same genres are treated in each set of materials. Where available, I chose example texts with similar topics from the EFAL LTSMs and the NST textbook to make the texts as comparable as possible.

5.3. Results

Below are tables summarising the distribution of genres within the EFAL materials (Table 12) and the NST textbook (Table 13).

EFAL Materials	Tokens	%
Description	1910	11.1%
Information Report	1672.5	9.7%
Procedure	2399.5	14.0%
Recount	1235.5	7.2%
Explanation	125.5	0.7%
Story Genres	6608.5	38.4%
Response Genres	0	0.0%
Exposition	51	0.3%
Lyrical Genres	1840.5	10.7%
Definitions	1352	7.9%
TOTAL	**17195**	**100.0%**

Table 12: EFAL genres

NST Textbook	Tokens	%
Description	2230	8.1%
Information Report	12425	45.4%
Procedure	6434.5	23.5%
Recount	368.5	1.3%
Explanation	2481.5	9.1%
Story Genres	189.5	0.7%
Response Genre	0	0.0%
Exposition	0	0.0%
Lyrical Genres	0	0.0%
Definitions	3268	11.9%
TOTAL	**27397**	**100.0%**

Table 13: NST genres

The first point to note is the distribution of genres in the subject EFAL as compared to NST. This is illustrated below in Figure 6 and Figure 7. While the data for both sets of materials is shown as a pie chart, readers should bear in mind that the two charts represent different amounts of text. The EFAL genre chart reflects percentages based on the 17 195 tokens encountered in the EFAL materials (see Table 12 above) while the NST genre chart reflects percentages based of the 27 397 tokens encountered in the NST textbook (see Table 13 above). Thus, while percentages in the two charts may be comparable, the absolute volumes of text they represent will not be. A comparison of the volume of exposure to particular genres will follow later.

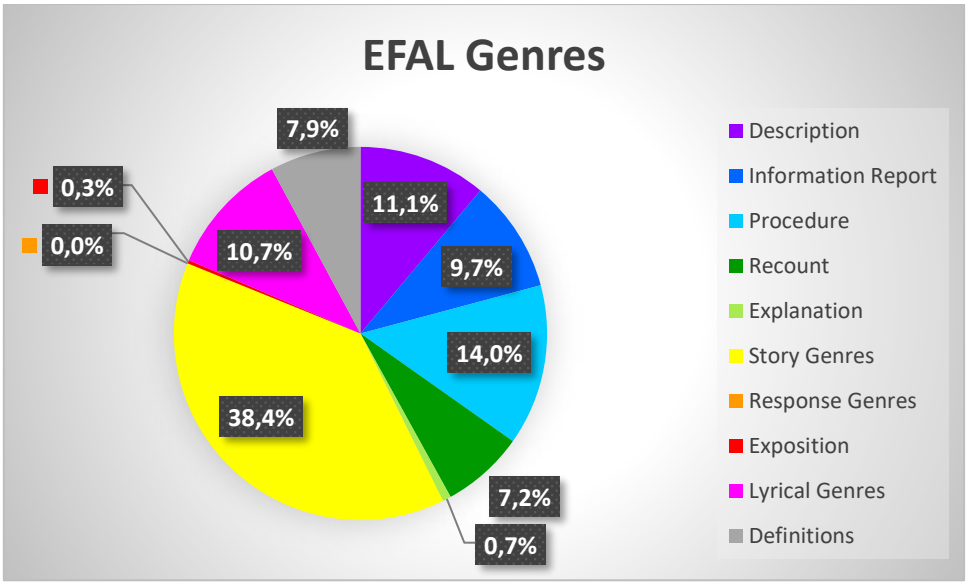


Figure 6: Genre distribution in the EFAL LTSMs

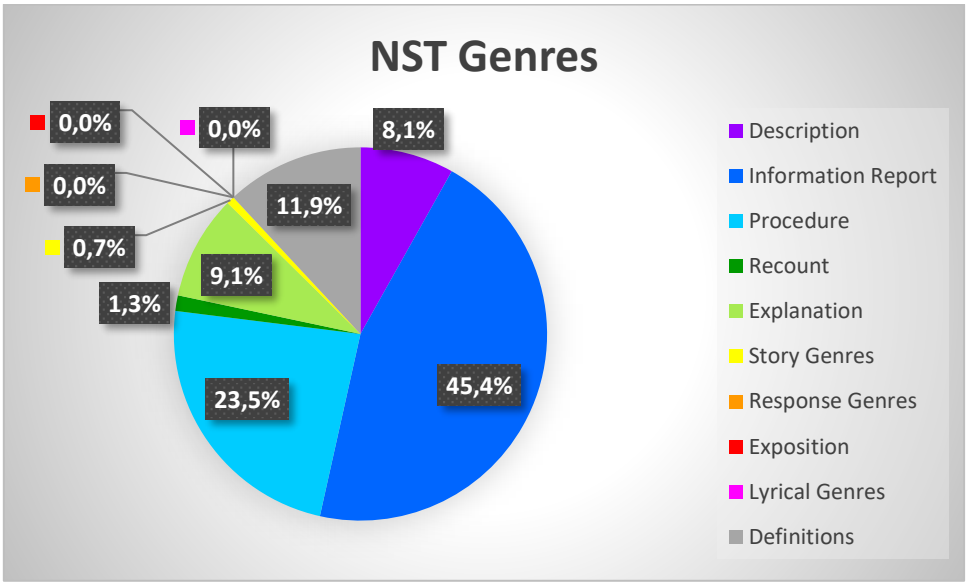


Figure 7: Genre distribution in the NST textbook

Several points merit discussion in these figures. First of all, the category *Response Genres* (whose purpose is to respond to a literary text or artistic work) is entirely absent in both sets of materials, while the category *Exposition* (whose purpose is to mount an argument) accounts for only 0.3% of the EFAL text and 0% of the NST text. Both these genres are essential for modelling critical thinking skills, and thus their absence in these materials is cause for concern.

Next, the categories *Story Genres* and *Lyrical Genres* which together account for 49.1% of the EFAL texts are almost entirely absent from the NST materials. *Story Genres* account for only 0.7% of the total NST text while *Lyrical Genres* do not feature at all. Put another way, almost

half of the texts encountered in Gr1-3 EFAL fail to prepare students for the predominant genre demands of a Grade 4 NST textbook. Of course, these genres serve other important functions in the cognitive and linguistic development of learners. However, learners are already encountering far less English text in Grades 1 to 3 as compared with one subject in Grade 4 (see discussion below). Thus, the fact that half of the text they do encounter in those years does not prepare them for the genres found within the NST textbook is problematic.

The final point worth commenting on relates to the volume of English reading students are expected to cope with as they progress through grades. The total number of word tokens encountered in the reading texts across three years of EFAL education (17 195 tokens) is substantially less than the number word tokens encountered while reading texts in a single year of NST education (27 397 tokens). Although not specifically related to genres, this issue only became apparent to me during my analysis in this area. A more rigorous comparison of the amount of text learners are expected to read is possible if we look at the total word count the EFAL materials used over the three years (including texts relating to tasks and activities which were excluded from the above tables) as compared with the total word count of the Grade 4 NST Learner's Book. This data is displayed below in Table 14 and Figure 8.

Textbook / Reader	Word Tokens	Percentage increase from previous year
Grade 1 – EFAL Learner's Book	5 257	-
Grade 2 – EFAL Learner's Book and Reader	11 729	123%
Grade 3 – EFAL Learner's Book and Reader	21 429	83%
Grade 4 – NST Learner's Book	41 842	95%

Table 14: Volume of reading in English by grade

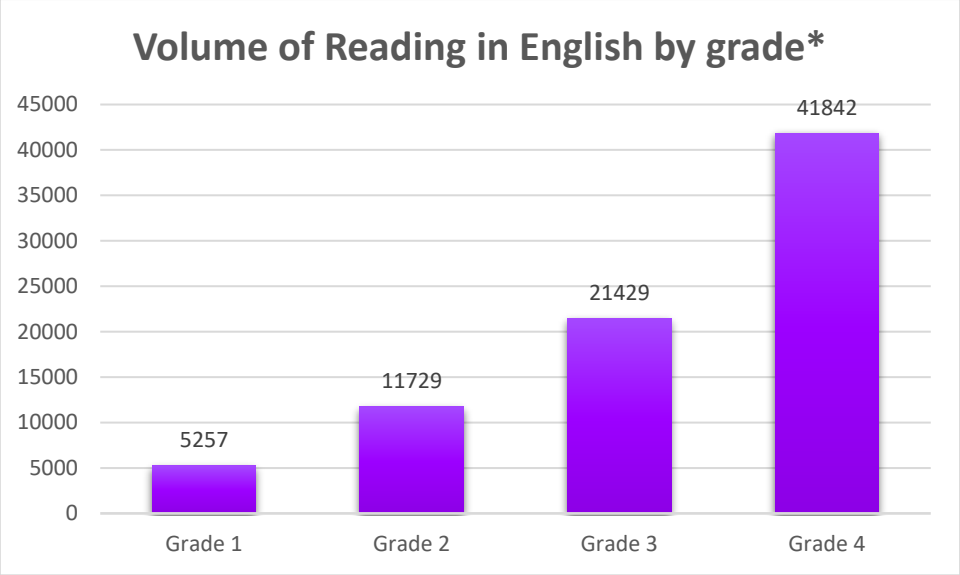


Figure 8: Volume of reading in English by grade*

**The data for Grade 4 is only taken from the subject Natural Sciences and Technology*

The data shows that the amount of English text learners are required to read more or less doubles every year, with the jump in absolute tokens from Grade 3 to Grade 4 being the largest (21 429 to 41 842 word tokens). Even before conducting a more detailed analysis of the genre characteristics of these texts, the jump learners experience from Grade 3 to 4 in terms of how much English text they are expected to read is a daunting one. If I were to factor in the amount of English reading required by other content subjects in Grade 4, this already exponential curve would become considerably steeper.

The remainder of my results will focus on a comparison of those genres which feature in both sets of materials.

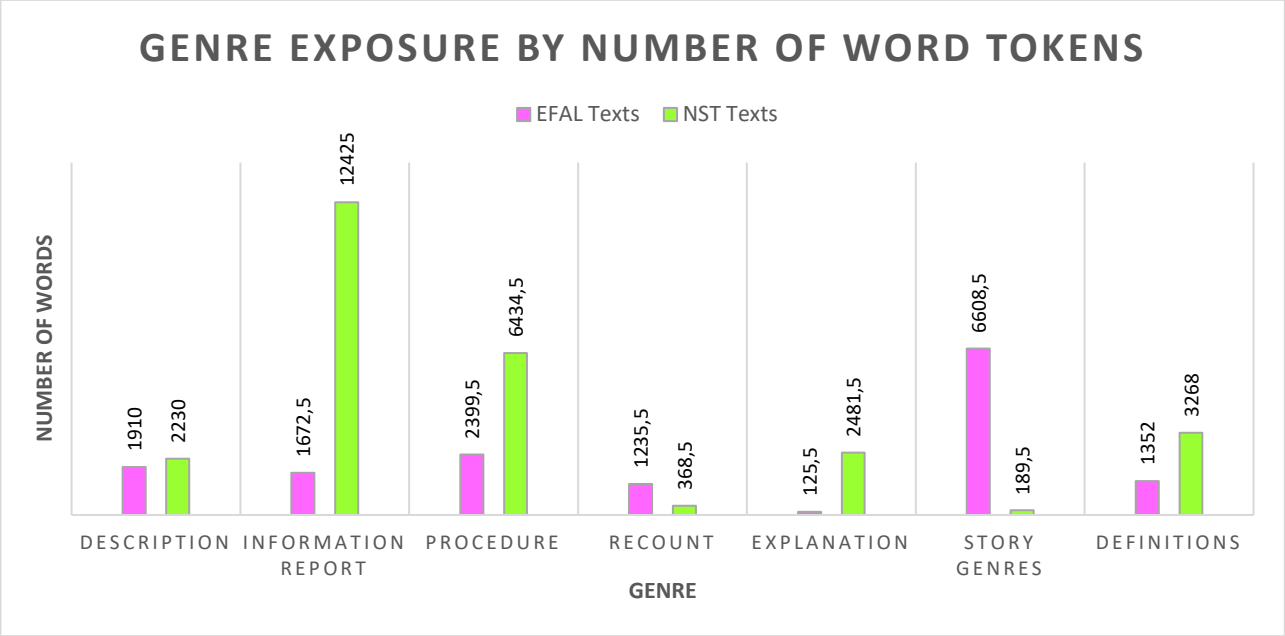


Figure 9: Genre exposure by number of word tokens

Looking at Figure 9, we can see that learners encounter an almost equivalent amount of text in both sets of materials for the *Description* genre. For the *Recount* genre (shown in the middle of the graph) and *Story Genres* (shown on the right), learners are exposed to far more text in their EFAL materials as compared with their NST textbook.

In contrast, learners receive far less exposure to the genres of *Information Report*, *Procedure*, *Explanation* and *Definitions* in their EFAL materials as compared with their NST textbook. The largest gap exists for *Information Reports* (12 425 word tokens in NST as compared with 1673 word tokens in EFAL). This is also the most common genre found in the NST textbook.

What are the implications of this in terms of preparing learners to make sense of their Grade 4 NST textbook? The more children encounter certain types of texts, the greater the degree of familiarity and eventually mastery they are able to develop with those types of texts. While new vocabulary and more complex conceptual content can make the text challenging to comprehend, familiarity with a text's genre can serve as a scaffold for the reader, greatly facilitating their ability to make predictions, inferences and process the reading materials. Familiarity with text types also play an important role in enabling learners to successfully reproduce those text forms in their own writing (Mayaba, Otterup & Webb, 2013). In this sense, the genres to which students have been frequently exposed are likely to be ones they cope well with (in this case, *Descriptions*, *Recounts* and *Story Genres*) while the genres to which they

have only limited exposure (in this case, *Information Reports*, *Procedures*, *Explanations* and *Definitions*) may prove more difficult for them to make sense of.

I shall now proceed to a qualitative comparison of some sample texts for each of the abovementioned genres from both sets of materials (with the exception of *Definitions* as vocabulary items have already been discussed at length in Chapter 4).

Information Report

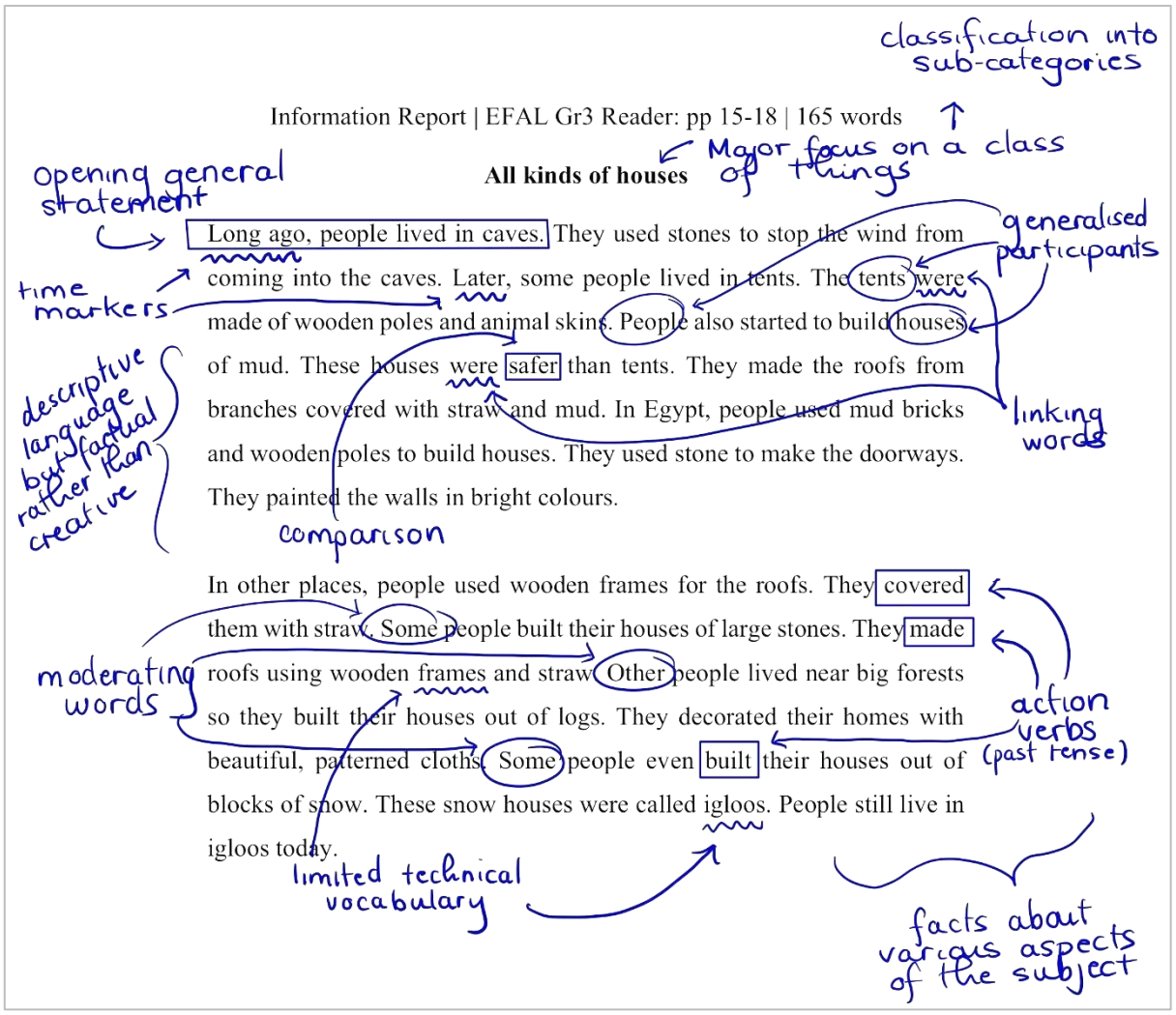


Figure 10: *Information Report* (EFAL)

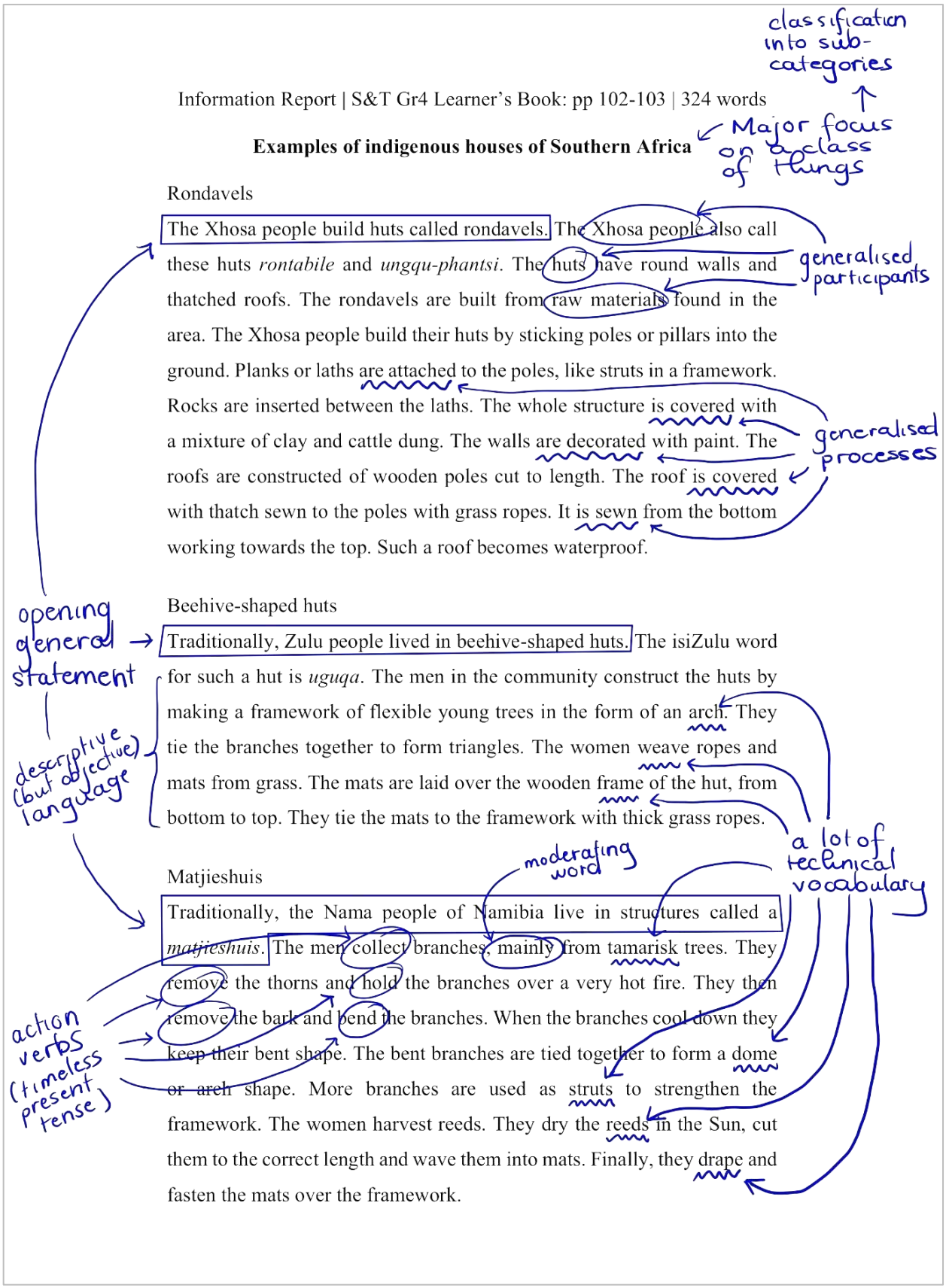


Figure 11: *Information Report* (NST)

Figure 10 displays a sample *Information Report*, taken from the Grade 3 EFAL Reader with my own annotations. Figure 11 displays a similarly annotated sample *Information Report* taken from the Grade 4 NST textbook.

When comparing these two, there are several things worth noting. Firstly, both texts display key characteristics of the *Information Report* genre such as a focus on a class of things, classification into sub-categories, generalised participants and the use of descriptive but factual language. However, the NST *Information Report* is both longer, more lexically dense (i.e. uses more specialised vocabulary) and more complex in terms of the detail and precision with which it discusses the topic.

The EFAL *Information Report* also displays some less typical characteristics for the genre such as time markers and the use of past tense, features more commonly found in *Recounts* or *Story Genres* (which abound in the EFAL materials). This might have been an intentional decision on the part of the authors to scaffold learners in processing the genre. While such a strategy has merit, the fact that there is a limited number of *Information Reports* in the EFAL materials means that if these are too simplified, it may make it even more difficult for students to cope with the genre in Grade 4. Alternatively, the Grade NST textbook will need to be simplified to align more closely with what is found in the EFAL materials.

Procedure

Figure 12 (see next page) displays a sample *Procedure* from the Gr3 EFAL Reader while Figure 13 (see page after next) shows a sample of the same genre found in the Grade 4 NST textbook.

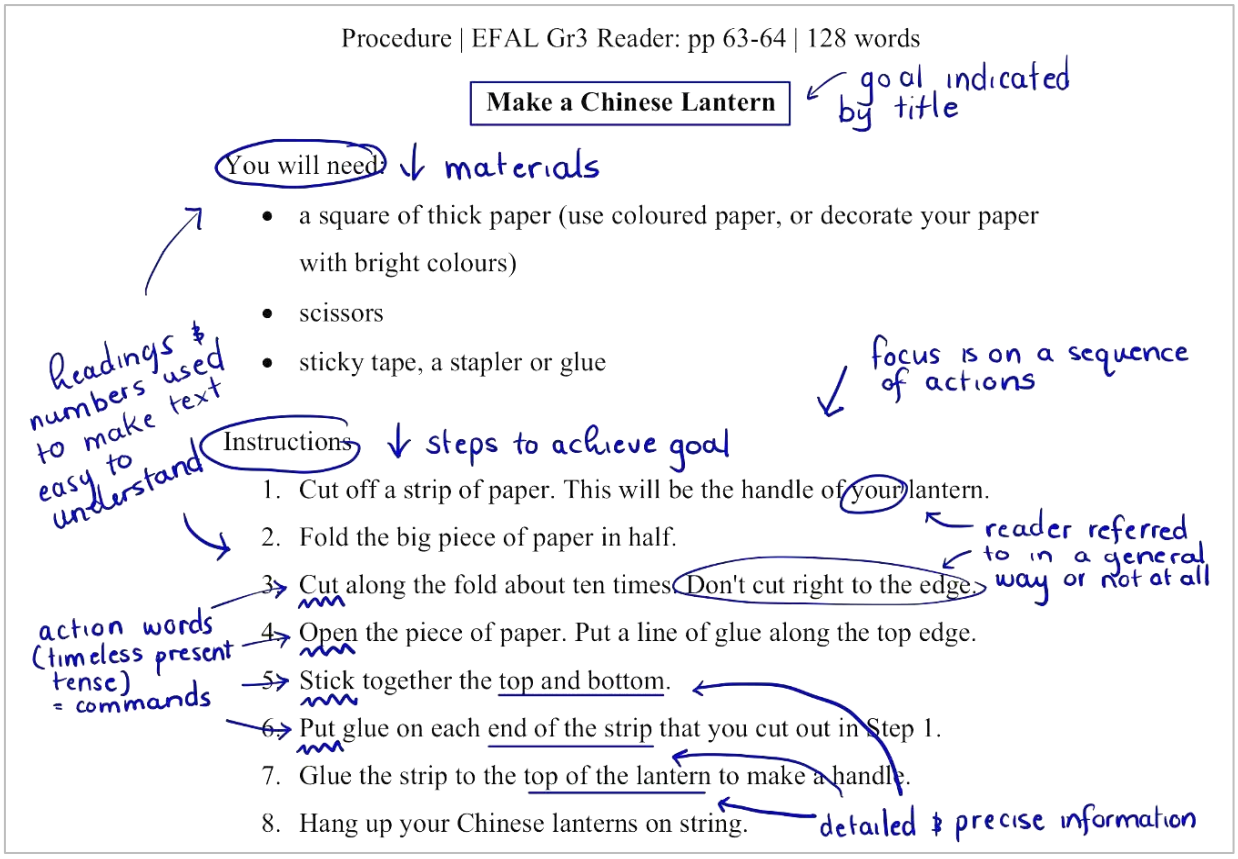


Figure 12: *Procedure* (EFAL)

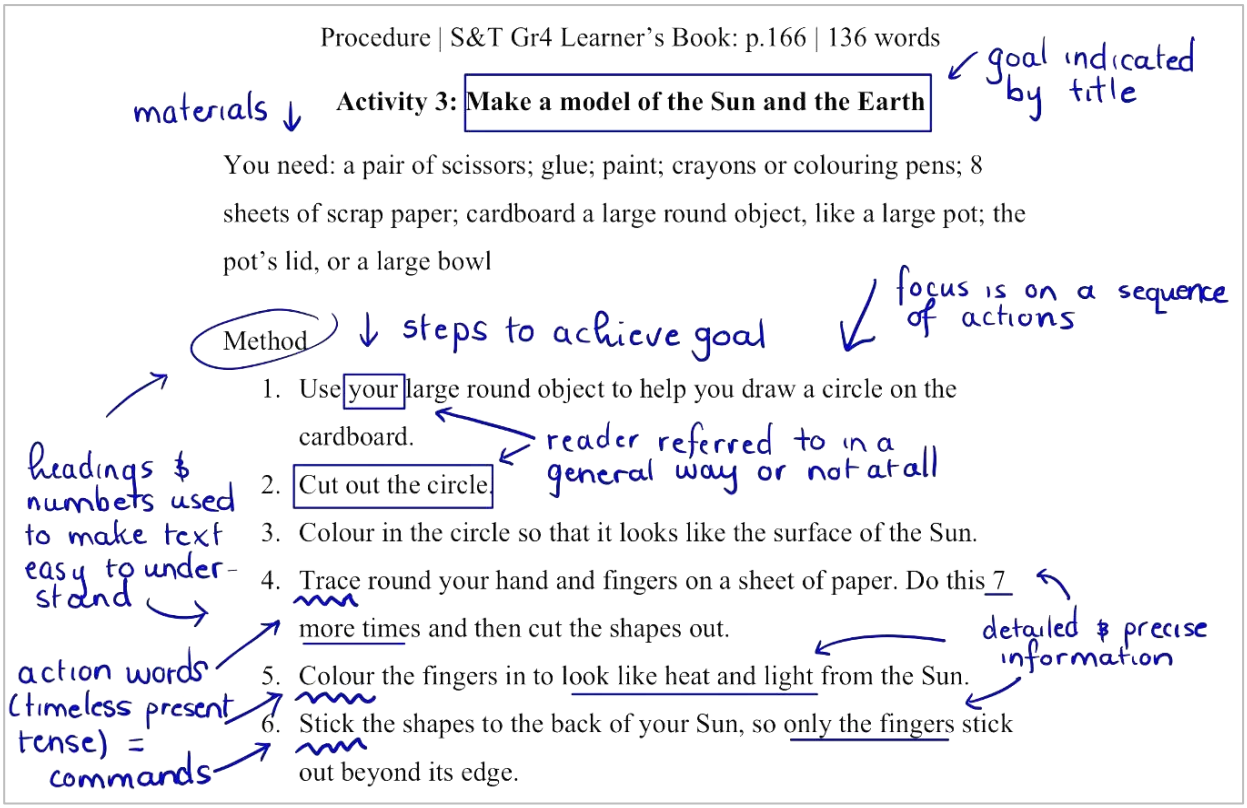


Figure 13: *Procedure* (NST)

Both texts are very similar. Both display the key features of *Procedures* such as a goal indicated by a title, a list of materials and numbered steps to achieve the stated goal. The information in both texts is detailed and precise, the verb tenses are the same and in general they could be considered of equivalent complexity. Thus, while the amount of exposure to procedural texts in EFAL is much less than will be encountered in NST, what learners do encounter closely matches what they will see in NST.

It is worth noting, however, that throughout the NST textbook *Procedures* often form part of larger activities learners are expected to complete. While the sample *Procedure* in Figure 13 is self-contained (in the sense that once the procedure is completed there is nothing further expected of learners), many other *Procedures* in the NST textbook are interspersed with questions asking learners to observe, reflect, evaluate or answer a specific question. An example of such a procedure taken from the Grade 4 NST textbook is displayed in Figure 14 (see next page).

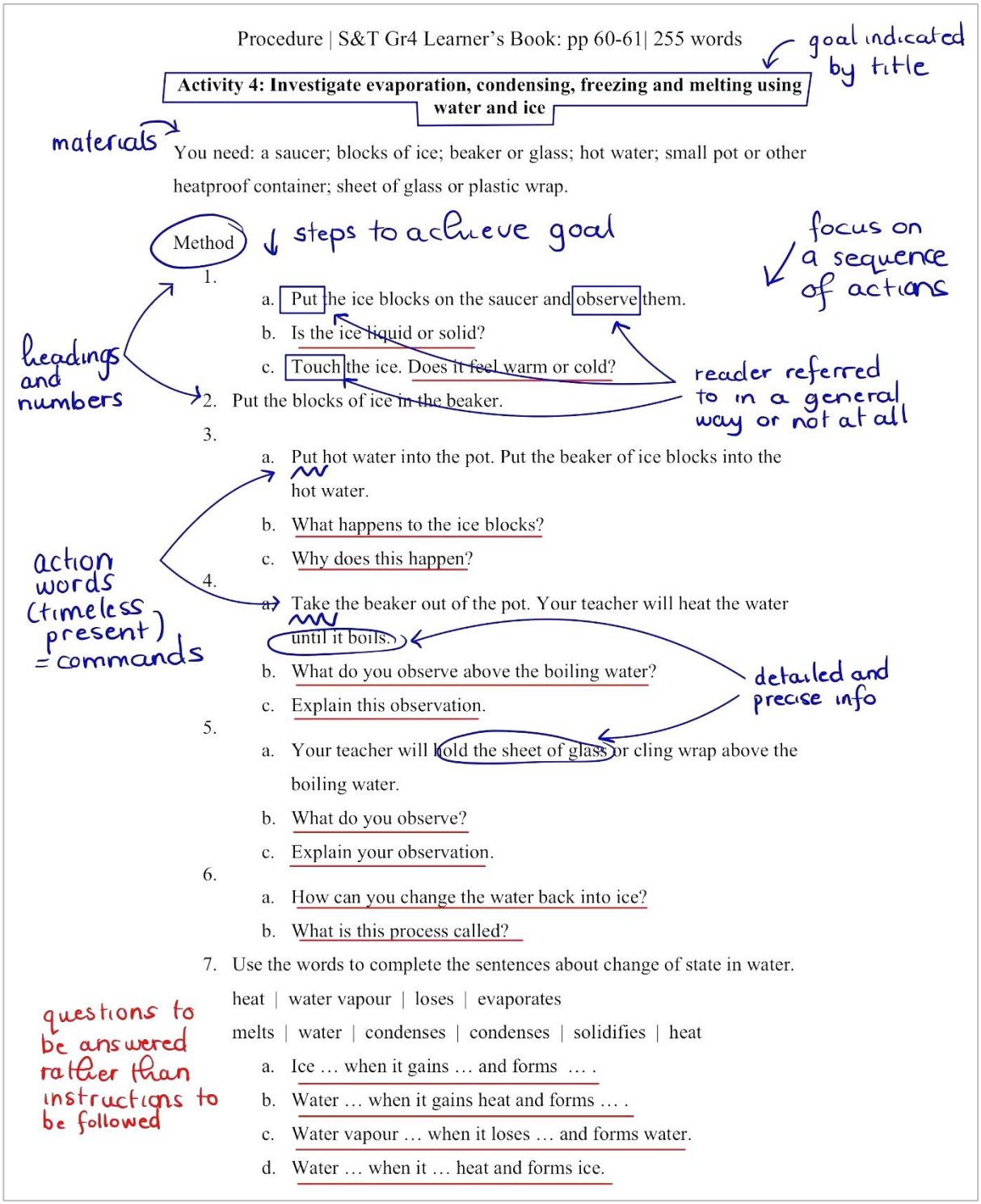


Figure 14: *Procedure* with questions (NST)

The sections of text underlined in red in Figure 14 are atypical for a procedure in that they constitute questions for learners to answer rather than instructions they must follow. Such sections of text occur commonly in the procedural texts found in the NST textbook but are not found at all in the EFAL materials. Questions for learners to answer do also appear within the

EFAL materials but they are usually found at the end of a text or even several pages after the text, never woven into the text itself. Moreover, questions rarely accompany procedural texts. They are more commonly found accompanying *Story Genres* and *Information Reports*. The extent to which this different *Procedure* format in the NST textbook may complicate or facilitate text comprehension for learners is unclear.

Explanation

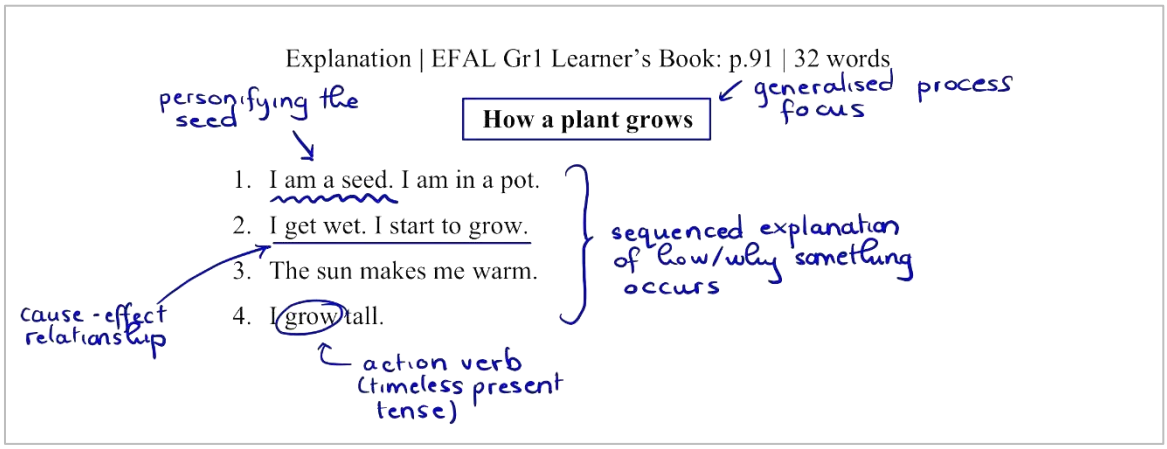


Figure 15: *Explanation* (EFAL)

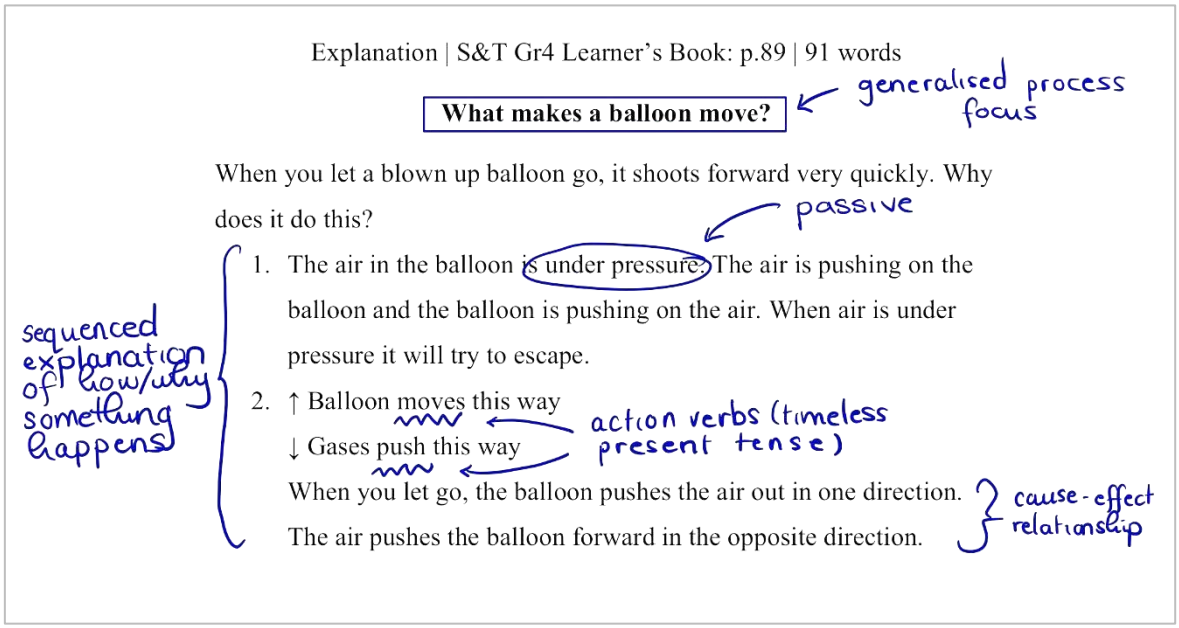


Figure 16: *Explanation* (NST)

Figure 15 displays a sample *Explanation* text from the EFAL Grade 1 Learner's Book, while

Figure 16 displays a sample *Explanation* text taken from the Grade 4 NST textbook. Before proceeding to my analysis it must be stated that there were only two texts in the entirety of the Grade 1-3 EFAL materials which displayed features of the *Explanation* genre. The first is the text seen in Figure 15. The second text, which appears in the Grade 3 Learner's book (and is therefore more appropriate for comparison with a Grade 4 text) is a hybrid text which displays features of more than one genre. It will be discussed later.

If we compare the texts in Figure 15 and Figure 16 we can see a clear difference in complexity, even just by looking at the length of the sentences. While both texts have a generalised process focus, as well as a sequenced explanation of how/why something happens and cause-effect relationships, the EFAL text is unusual in that it personifies the seed being discussed. The NST text in turn, relies on symbols such as directional arrows to support its explanation. These differences are highlighted by the illustrations accompanying the respective texts, displayed in Figure 17 (below) and Figure 18 (see next page).

Figure 17: Illustration for EFAL *Explanation*

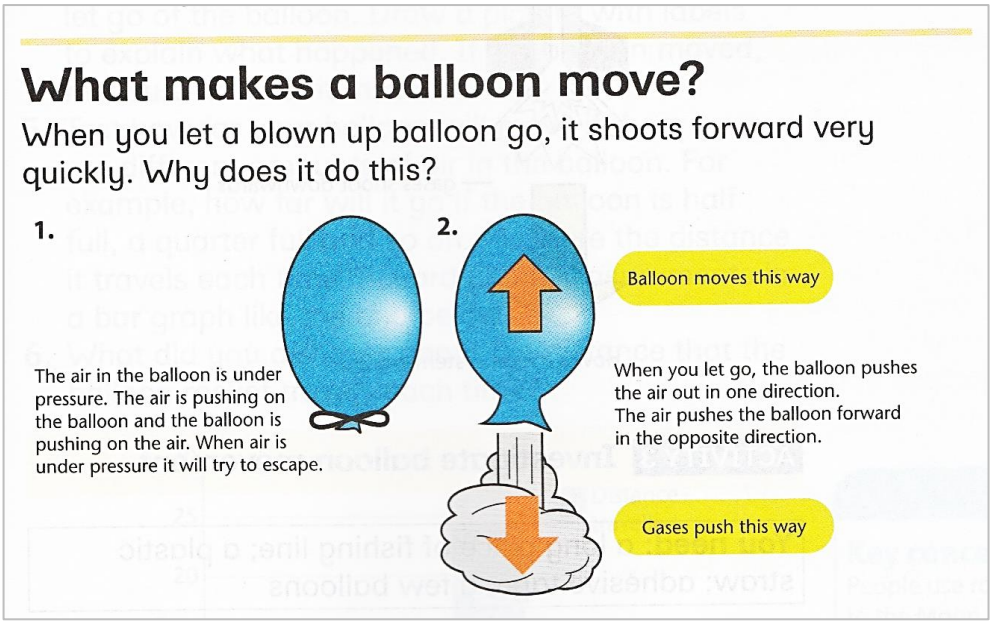


Figure 18: Illustration for NST *Explanation*

The act of personifying the seed, both in language and also in the illustration could be a strategy employed by the authors to attempt to make the content more relatable to young children, as they are typically less able to engage with and make sense of information that bears no direct relationship to them. However, this may have the opposite effect as it contradicts typical genre patterns which can guide learners, i.e. that first person pronouns such as 'I' usually refer to people (or in some narratives, animals and objects with human like personas).

In comparison, the more abstract balloon text and illustration make for a very different manifestation of the *Explanation* genre. This large gap in the style and complexity of the *Explanation* texts mean that learners are unlikely to be adequately prepared by their EFAL materials to cope with NST *Explanation* texts. The only other text found in the EFAL materials which displayed some features of the *Explanation* genre is displayed in Figure 19 (see next page).

The text in Figure 19 has a generalised process focus and the sequenced explanation of how things happen found in the other *Explanation* texts. However, it also displays features of an *Information Report* in that it focuses on various components, with features such as sub-headings and several specialised vocabulary items. It occasionally also focuses on specific

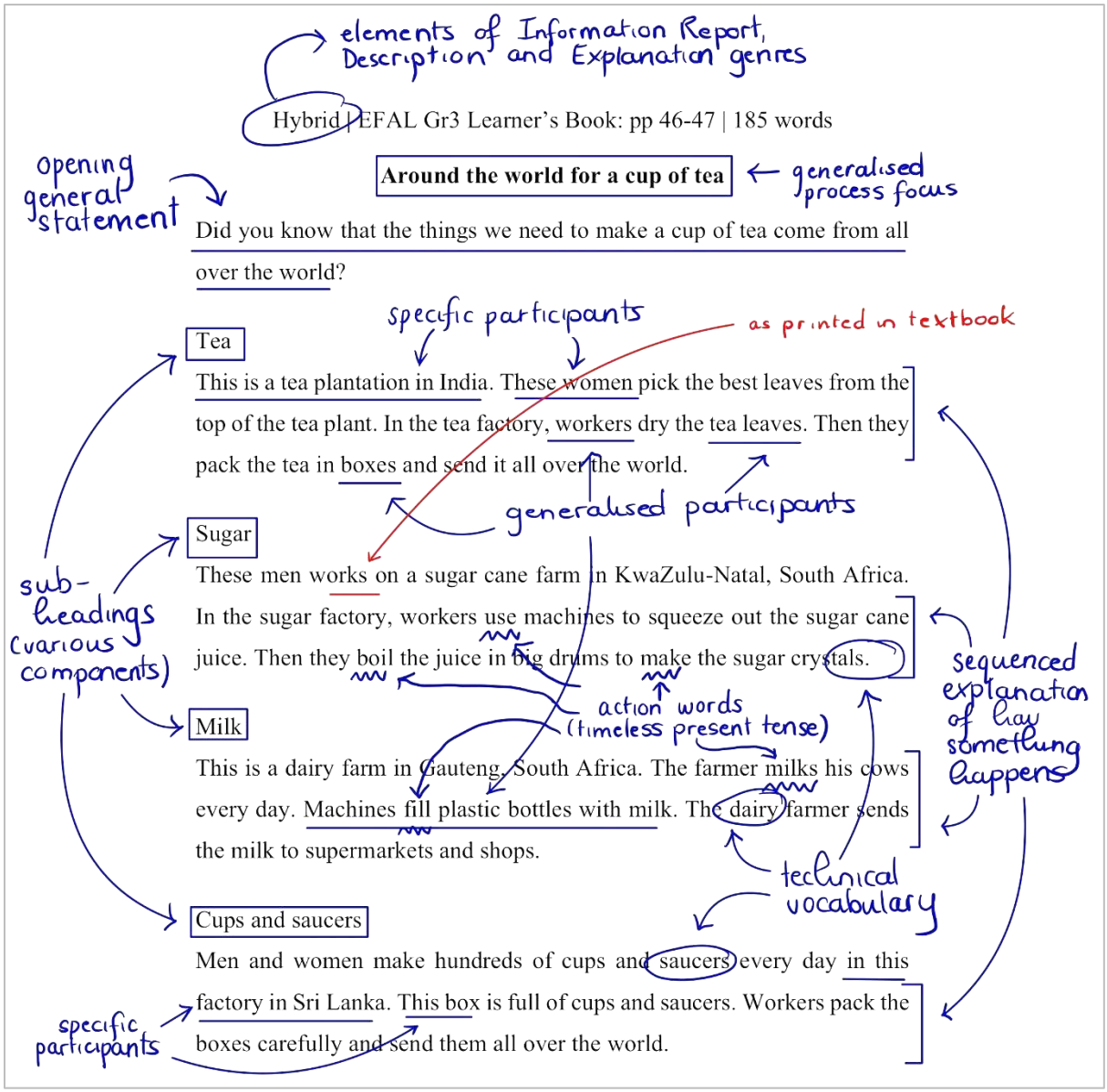


Figure 19: *Hybrid* (EFAL)

participants rather than a class of things, which is typical of the *Description* genre. This text can thus be considered an example of a hybrid genre. Because hybrid genres have an atypical structure and consist of a variety of elements, they do not assist learners in developing mastery of a particular genre.

It becomes clear from this analysis that the EFAL materials do not adequately prepare learners to cope well with the *Explanation* genre, both in terms of the volume of exposure as well the complexity of the texts encountered.

Description

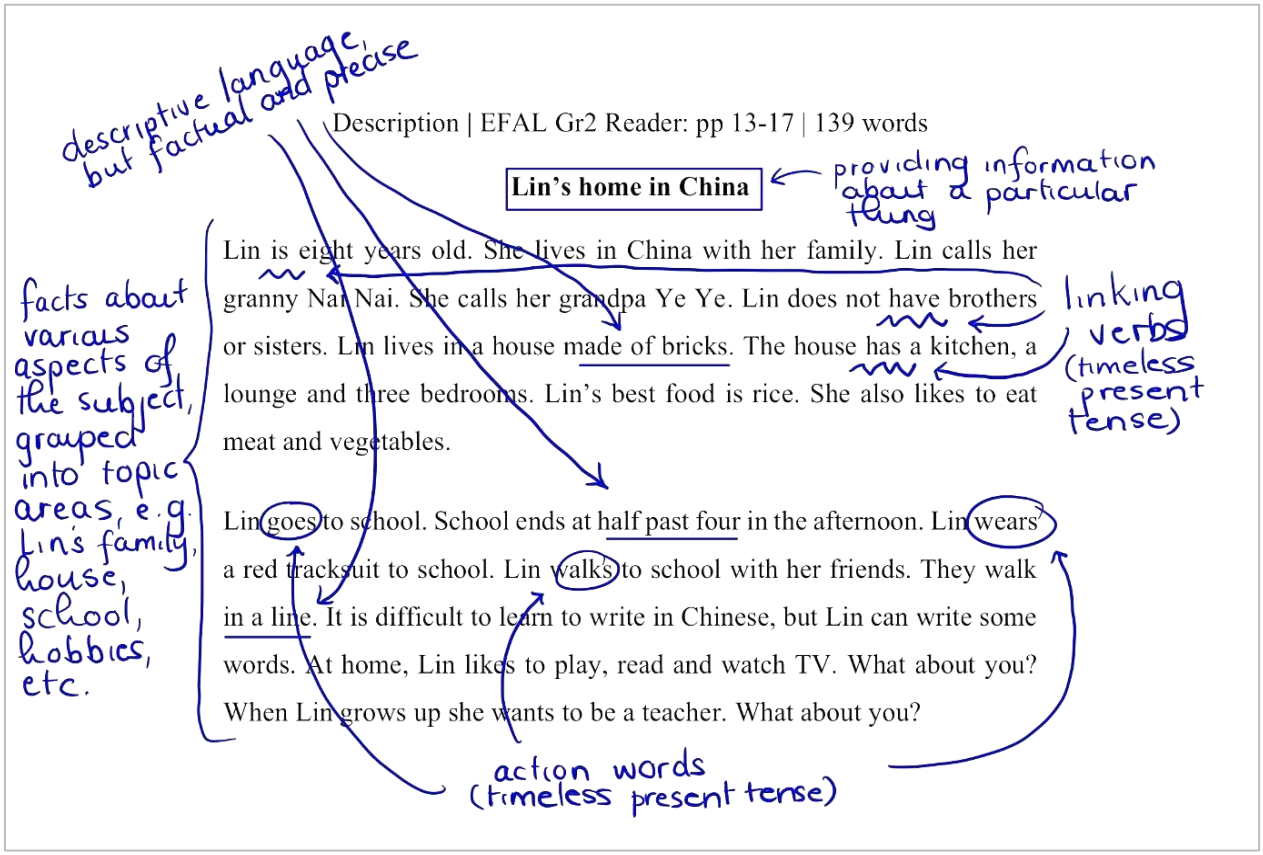


Figure 20: *Hybrid* (EFAL)

Figure 20 (above) displays a sample *Description* taken from the Grade 2 Reader. Figure 21 (see next page) displays a sample *Description* from the Gr4 NST textbook. Structurally, the two texts are very similar. Both provide information about a particular thing and this information is ordered into several topic areas. Both texts also make use of descriptive but factual and precise language. However, the NST *Description* text is more complex in that it makes use of subheadings, has far more specialised vocabulary and uses comparative language. There is thus a clear gap in complexity between the two sample texts.

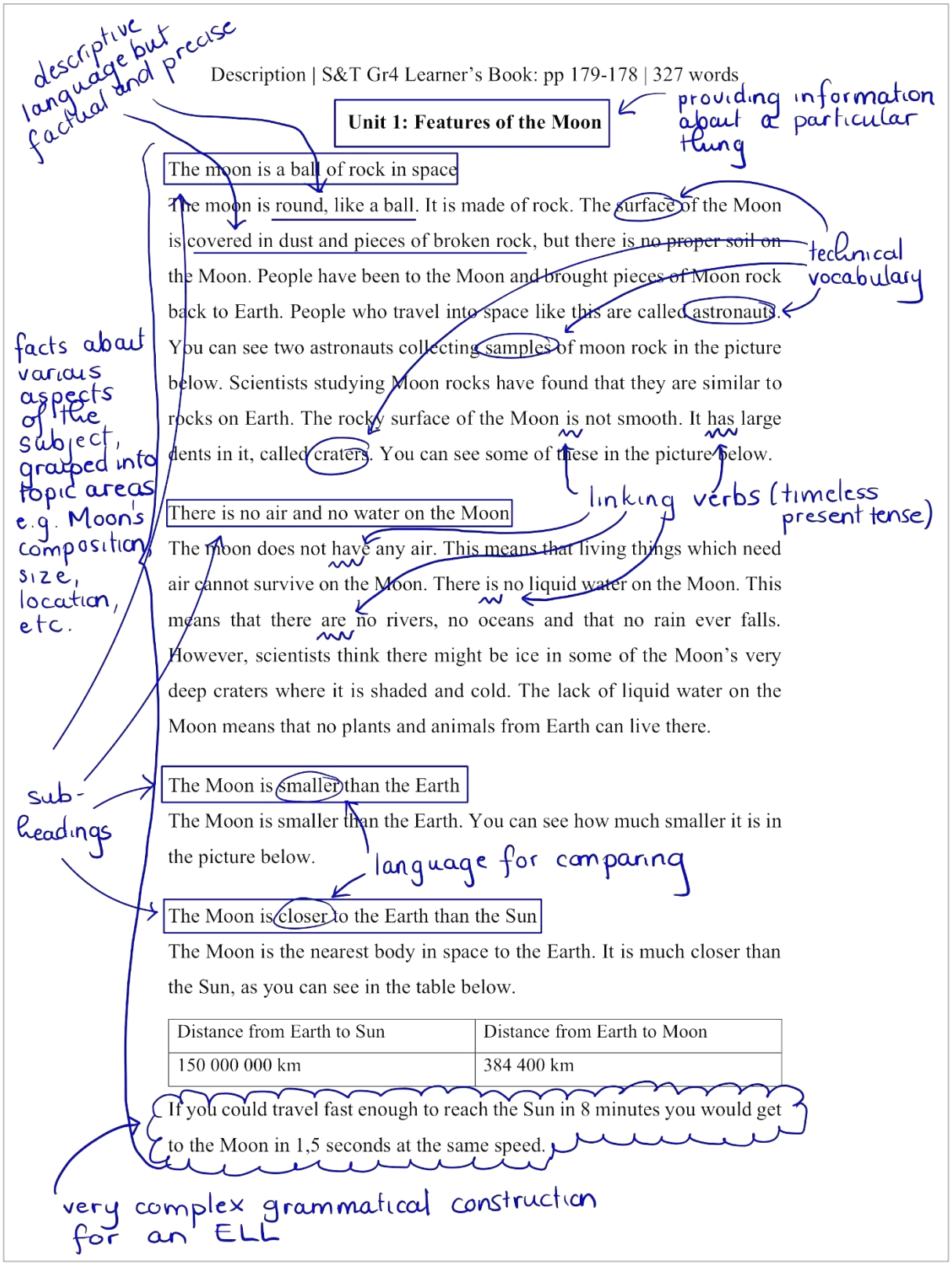


Figure 21: *Description* (NST)

Recount

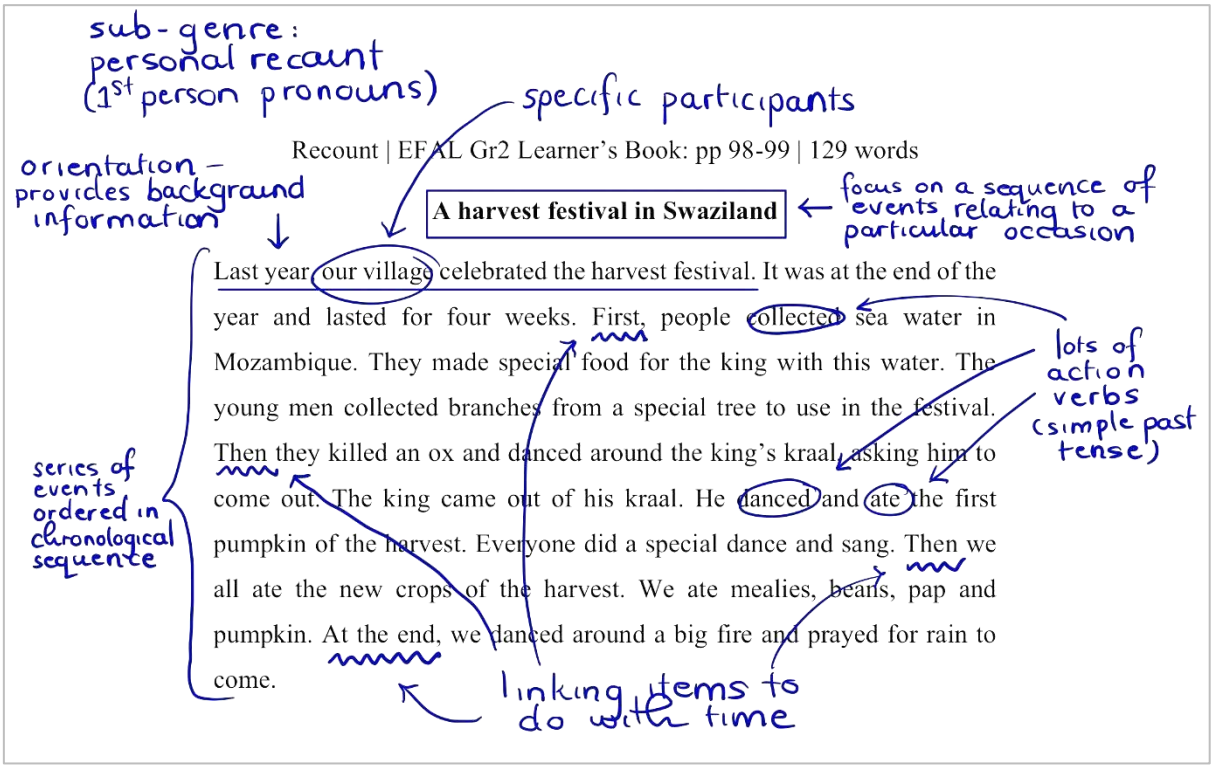


Figure 22: *Recount* (EFAL)

Figure 22 (above) displays a sample *Recount* taken from the Grade 2 EFAL Learner's Book while Figure 23 (see next page) displays a sample *Recount* taken from the Grade 4 NST textbook. Both texts focus on a sequence of events relating to a particular occasion, ordered in chronological sequence. Both start off with an orientation to provide background information for the reader, have verbs in the simple past tense and also have linking items related to time throughout the text. However, the NST *Recount* is more complex in that it also provides explanations for certain details in the text. The similarities between the texts exceed their differences. Moreover, given the high amount of exposure to *Recounts* throughout the EFAL materials compared to what is encountered in NST, it seems likely that learners would be able to cope well with this genre.

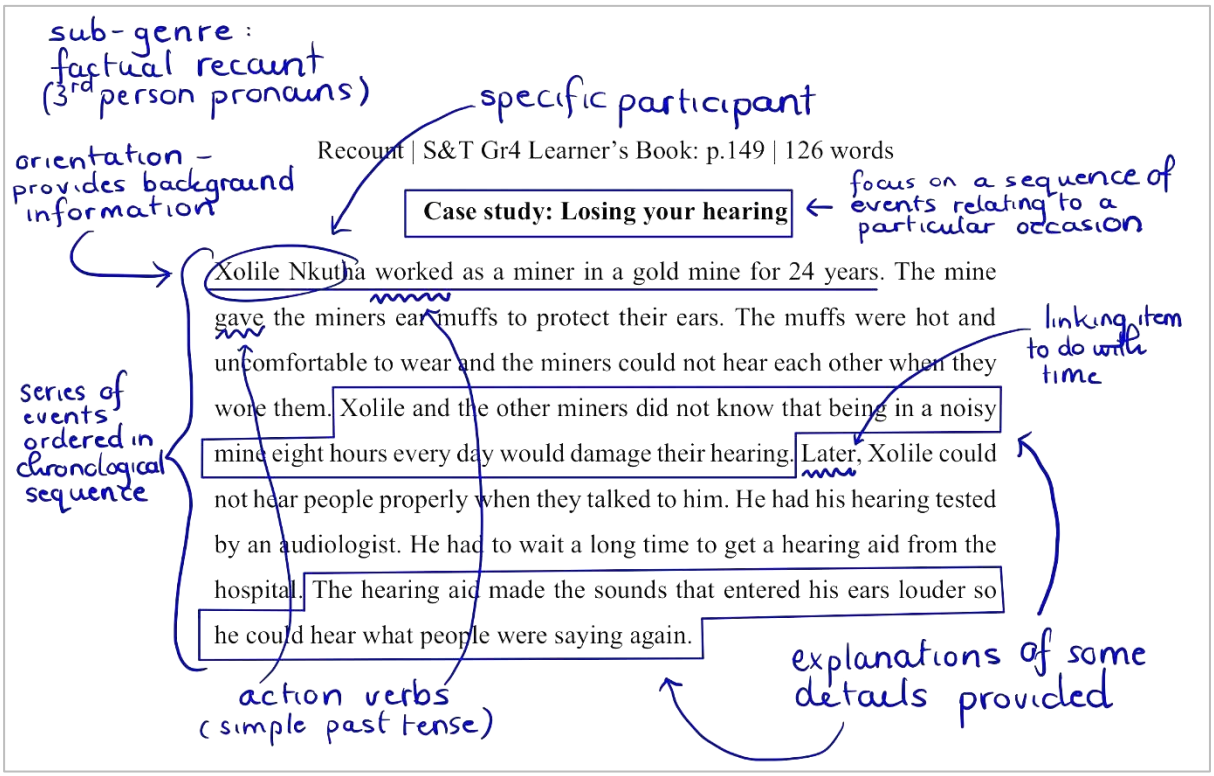


Figure 23: *Recount* (NST)

Story Genres

Figure 24 (see next page) shows a sample *Story Genre* taken from the EFAL Grade 3 Learner's Book while Figure 25 (see page after next) shows a hybrid text which shares some elements of a *Story Genre*, taken from the NST textbook. The EFAL text has all the typical conventions of a narrative (a sub-genre of *Story Genres*). It is written in the past tense, has specific individual participants with defined identities and has clear linking words related to time. It also has a complication, a resolution and dialogue between characters.

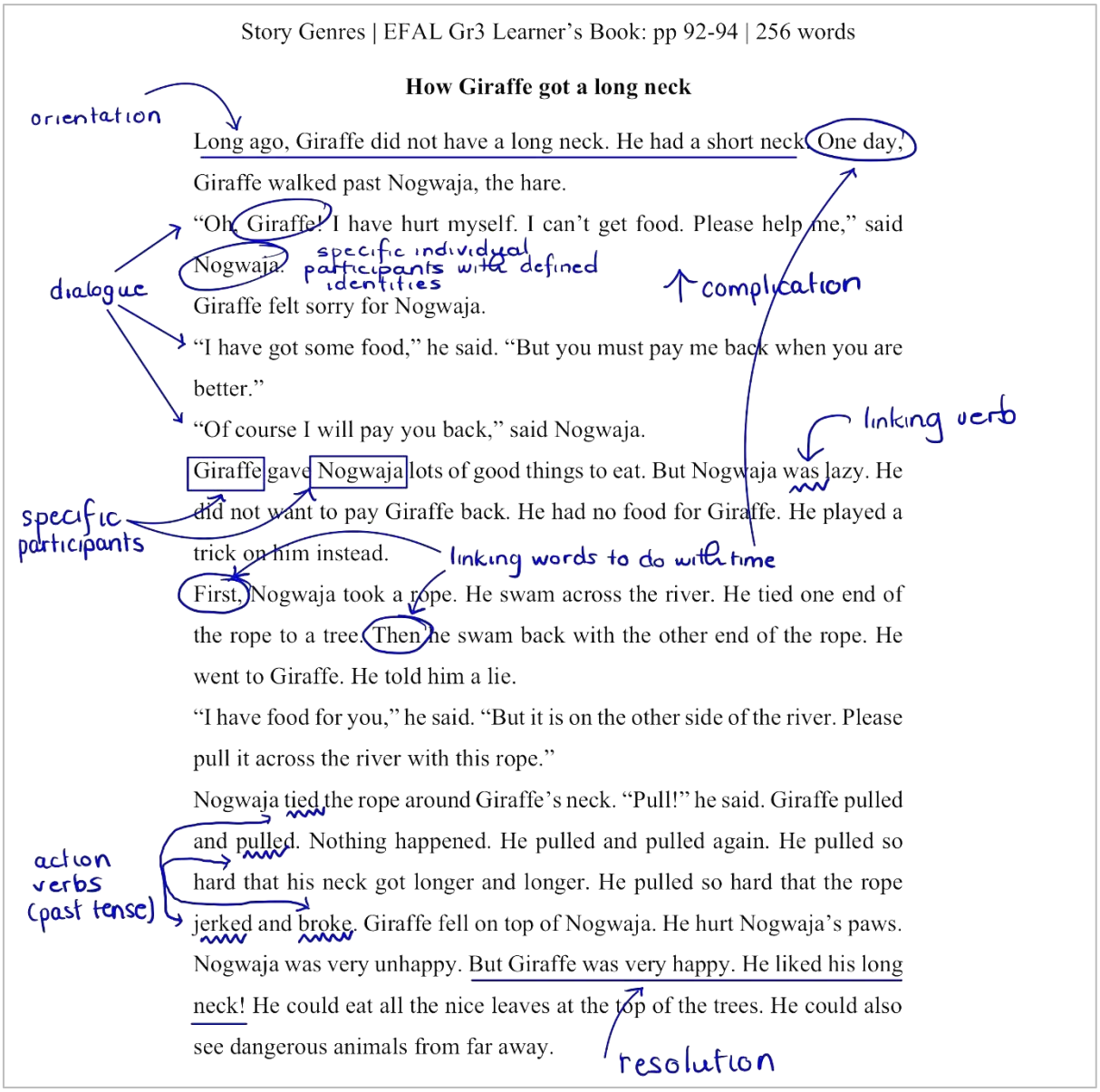


Figure 24: *Story Genre* (EFAL)

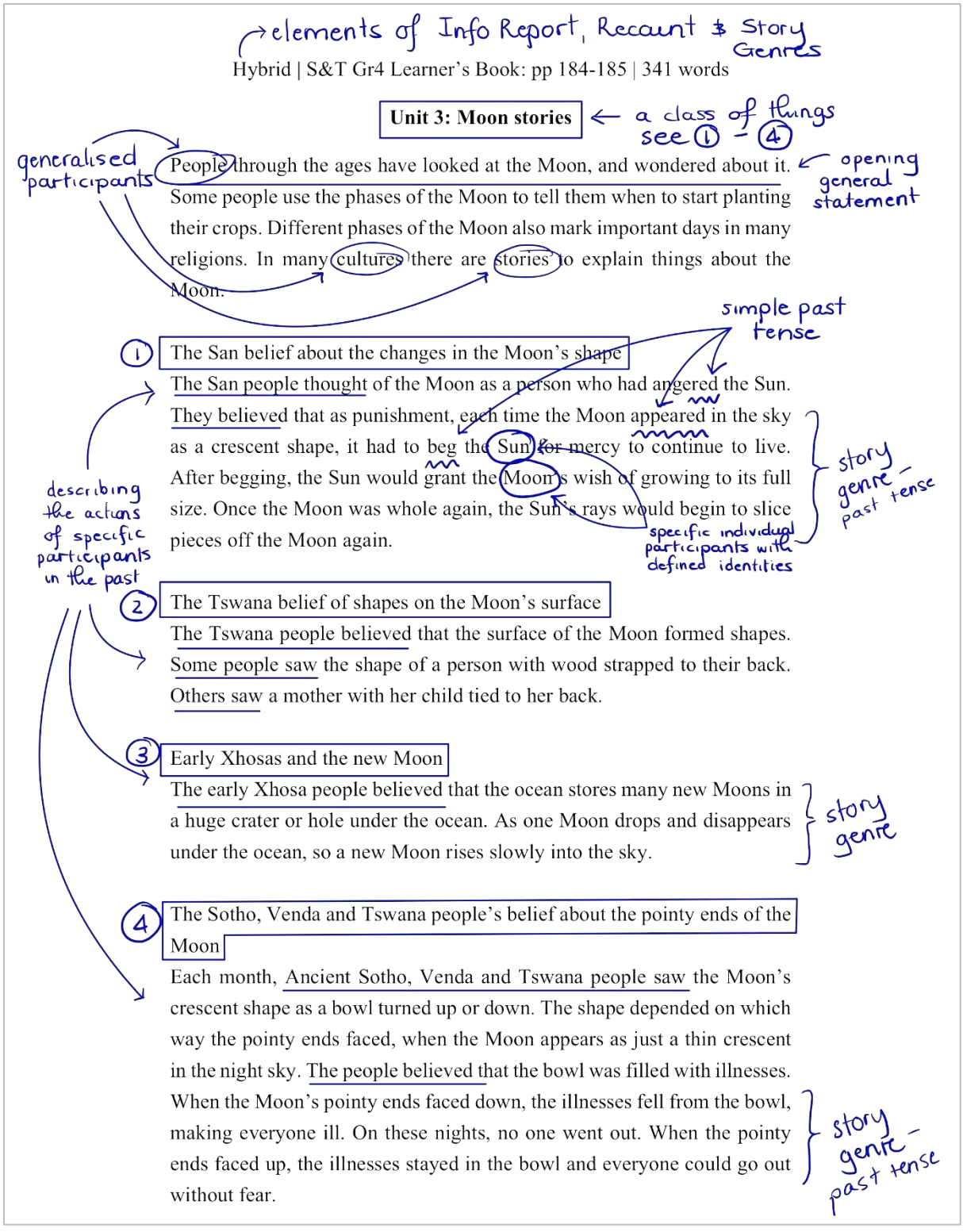


Figure 25: *Hybrid* (NST)

In comparison, the hybrid text titled 'Moon Stories' taken from the NST textbook is much harder to categorise. Insofar as it discusses a class of things (Moon Stories) and starts with an opening general statement, has generalised participants and uses the present tense, it follows

the conventions of an *Information Report*. However, as it moves into describing particular stories, it starts to display both features of a *Recount* (describing the actions of specific participants in the past) as well as *Story Genres* (specific individual participants with defined identities like the personified Sun and Moon). This text, though describing narratives from a variety of cultures, is itself not structured like a narrative and in this sense does not link well to the other examples of *Story Genres* found in the EFAL materials.

The degree to which learners are able to make sense of the 'Moon Stories' text, as well as other complex hybrid texts which have atypical structures is likely to depend more heavily on the other factors highlighted in my analytical framework (i.e. vocabulary, syntax, visuals, etc.) than the scaffolding provided by the text's genre.

In summary, for genres which appear in both sets of materials, we can say the following: the *Information Report* found in the NST textbook is both longer, more lexically dense and more complex than its EFAL counterpart. The EFAL *Information Report* also displays some less typical characteristics which may make it difficult for students to transition to the genre in Grade 4. For *Procedures*, the volume of exposure in EFAL is much less than will be encountered in NST. While some *Procedures* in the EFAL closely match what will appear in the NST, other *Procedures* in the NST textbook display additional atypical features (asking learners to observe, reflect, evaluate or answer a specific question); the extent to which these difference may impede text comprehension for learners is unclear.

For *Explanations*, the EFAL materials do not adequately prepare learners to cope well with the genre, both in terms of the volume of exposure as well the complexity of the texts encountered. For *Descriptions*, what appears in the NST is also more complex than the EFAL. For *Recounts*, while the NST text is more complex than the EFAL, the similarities between the texts exceed their differences and the high amount of exposure to *Recounts* throughout the EFAL materials makes it likely that learners would be able to cope well with this genre. *Story Genres* in the NST textbook do not bear a significant resemblance to what appears in the EFAL, but instead manifest as hybrid texts. The degree to which learners are able to make sense of hybrid texts is likely to depend more heavily on the other factors rather than the scaffolding provided by the text's genre.

In other words, for five out of the six genres compared (*Information Reports, Procedures, Explanations, Descriptions* and *Story Genres*) there are significant differences between the EFAL and NST manifestations of each, such that it cannot be assumed learners will be able to smoothly transition to making sense of these genres in their Grade 4 NST textbook. Only one genre, *Recounts,* manifests in both sets of materials in such a way that learners are likely to successfully engage with the genre in Grade 4.

Chapter 6: Syntax

6.1. Framework

Syntax forms part of the grammar of a language. Tweissi (in Nation, 2001:171) explains that while studies rank grammatical features well below vocabulary in terms of impact on readability, grammatical features can still play a significant role in how accessible a text is.

Fromkin, Rodman and Hyams (2011:13) define grammar as "the knowledge speakers have about the units and rules of their language – rules for combining sounds into words (called phonology), rules of word formation (morphology), rules for combining words into phrases and phrases into sentences (called syntax), as well as the rules for assigning meaning (called semantics)." I am particularly concerned with the impact of syntax on the accessibility of a text. Syntax concerns itself primarily with issues of sentence structure, word order and agreement between different parts of speech (nouns, verbs, adjectives, adverbs, etc.).

L1 speakers are assumed to have mastery over most of the oral syntax for the standard variety of their language by the time they engage with written text. Even so, transitioning from spoken to written language requires some adjustments. This is because "written texts employ styles and tend to use syntactic patterns that are notably different from those found in oral texts" (Goodman, 1984:89). The idiosyncrasies of written language are a result of the greater ambiguity inherent in written text as well as the editing and polishing that writing is often subjected to. The gap between conversational language and the kind of academic language learners encounter in their written textbooks and assessments has also been highlighted by Cummins (2008) in his work on the distinction between BICS and CALP, discussed briefly in Chapter 2.

Nilagupta (1977:585) states that "[t]he rules of syntax are not just the rules the writer applies to organise his statements – they are also the rules he assumes the receiver knows in order to be able to extract the meaning from statements." Once L1 readers become familiar with the unique features of written syntax, they are generally able to cope with the grammatical structures present in a written text. In contrast, L2 readers do not always have mastery over the majority of L2 syntax, whether oral or written, by the time they begin reading L2 texts (Lipka

& Siegel, 2011). Therefore, syntactic complexity becomes an especially important factor to consider when gauging the difficulty of a text intended to be read by ELLs.

To measure syntactic complexity in a valid and reliable fashion is not an easy task. Traditional readability measures account for syntactic complexity in a very blunt fashion – by using sentence length and number of total words in a passage as proxy variables. Other formulas consider the number of syllables in a sentence. The principle followed is that shorter sentences and texts are more comprehensible. However, some writers have contradicted this commonly held view, stating that "[s]hort, choppy sentences may detract more from text readability than longer sentences with examples and illustrations meant to enhance comprehensibility. Elaborated text is easier to understand" (Graves & Graves in Sibanda, 2014:157; see also Bailin & Grafstein, 2016).

Other research has delved more deeply into specific syntactic elements that have proved challenging for ELLs. Nilagupta (1977) conducted a study with Thai senior high school and tertiary ELLs in which he identified five syntactic factors which contributed to readability difficulty. The five factors were: negatives, passive voice construction, embedding of clauses, deletion of redundant words and phrases as well as nominalisation. Pitler and Nenkova (2008) found that two syntactic elements which correlated strongly with readability (beyond vocabulary and length of text), were discourse relations (e.g. as expressed by adjacent ideas or discourse connectives such as *because* and *however*) and the average number of verb phrases per sentence.

More recently, developments in computer linguistics have led to a proliferation of automated tools for text analysis. Many of these are based on more comprehensive models of text processing than traditional readability formulas. Such tools include the Coh-Metrix Text Easability Assessor (Graesser, McNamara, Louwerse and Cai, 2004; Graesser, McNamara & Kulikowich, 2011), the Reading Maturity Metric (RMM, developed by Pearson) and the TextEvaluator tool, developed by Educational Testing Service (ETS, the same entity which administers the TOEFL and TOEIC examinations of English proficiency).

Nelson, Perfetti, Liben and Liben (2012) conducted a study comparing the ability of six text difficulty metrics to predict the grade level and difficulty of a number of different texts that

had already been rated by team of expert human raters. Included among these were the RMM and TextEvaluator (then known as SourceRater) tools. The study also evaluated the Coh-Metrix Text Easability Assessor but could not compare it to other metrics as it does not provide a single, unified complexity score. Of the six metrics evaluated, SourceRater (now called TextEvaulator) provided the most accurate predictions as compared with the human ratings. For this reason, I will be using the TextEvaluator tool in my own analysis.

TextEvaluator Tool

The following description of the TextEvaluator tool is based on the work done by Sheehan, Kostin, Napolitano and Flor (2014) and Sheehan, Flor and Napolitano (2013). More details about the tool and its components can be found in these papers.

The TextEvaluator tool provides nine scores for text difficulty: one overall difficulty score which can be linked to a US grade level, as well as eight component scores which fall within four specific areas of text difficulty. The first area, which most closely aligns with my focus in this chapter, focuses on *Sentence Structure*. The component score which relates to this area is the Syntactic Complexity score. The Syntactic Complexity score is based on (among other things) average sentence length, average number of dependent clauses, average number of words before the main verb, and a measure of sentence "depth" called Average Maximum Yngve Depth (based on research conducted by Yngve, 1960, cited in Sheehan et al., 2014:194). The latter is "designed to capture variation in the memory load imposed by sentences with varying syntactic structures".

The second area of focus is *Vocabulary Difficulty*, and three component scores relate to this: Academic Vocabulary, Word Unfamiliarity and Concreteness. The Academic Vocabulary score is based on the frequency of words from the text which appear in the Academic Word List developed by Coxhead (2000) (discussed in Chapter 4) as well as the frequency of nominalisations, academic verbs, abstract nouns, topical adjectives and cognitive process nouns (the latter four categories are based on the work of Biber, 1986, 1988 and Biber et al. 2004, cited in Sheehan et al. 2014:194). The Word Unfamiliarity score is based on two word frequency indexes selected by the tool developers, while the Concreteness score is based on word concreteness ratings drawn from the Medical Research Council (MRC) psycholinguistic database. According to Alderson (cited in Sheehan, Flor & Napolitano, 2013:54) "the level of

concreteness present in a text is a useful feature to consider when evaluating passages for use on reading assessments targeted at L2 readers".

The third area the TextEvaluator tool focuses on is *Connections Across Ideas*, which consists of three component scores: Lexical Cohesion, Interactive/Conversational Style and Level of Argumentation. Lexical Cohesion focuses on aspects of the text which help readers interpret the material as a coherent message rather than a collection of unrelated clauses and sentences. The text features which produce cohesion (based on the work of Halliday and Hasan, cited in Sheehan, Flor & Napolitano, 2013) are repeated content words and explicit connectives. The Interactive/Conversational Style score reflects the frequency of conversation verbs, fiction verbs, communication verbs, first person plural pronouns, contractions and words enclosed in quotes (based on the work of Biber, cited in Sheehan, Flor & Napolitano, 2013). The Level of Argumentation component score measures the frequency of concessive and adversative conjuncts such as *although, though, alternatively, in contrast,* etc. as well as the frequency of negations such as *no, neither,* etc.

The final text area the tool focuses on is *Organisation*, which consists of only one component score: Degree of Narrativity. This score reflects the frequency of past perfect aspect verbs, past tense verbs and third person singular pronouns in the text (based on the work of Biber and Graesser et al. cited in Sheehan et al. 2014).

An important aspect to note in the design of the TextEvaluator tool is that it distinguishes between three types of genres: Informational, Literary and Mixed. The tool classifies all texts that it processes as one of these three genres and predicts different levels of difficulty based on these classifications. In other words, an Informational text and a Literary text, even if they were to have comparable component scores, would rank as different levels of difficulty because the TextEvaluator tool weights the different component scores differently depending on the genre of the text. For more information on this, see Sheehan, Flor and Napolitano (2013).

The major limitation of using this tool to process texts taken from South African textbooks is that the corpus and curriculum standards used to develop and calibrate the TextEvaluator tool were based on the US education system (i.e. the US Common Core standards and several US high stakes assessments). Therefore, the grade levels that the tool suggests need to be treated

with caution, especially for texts drawn from the EFAL materials. Even so, the information the tool provides about various aspects of text complexity, as well as the ability to evaluate texts relative to one another using the component scores make it a very valuable resource for textual analysis.

6.2. Methodology

In order to conduct my syntax analysis I selected a purposeful sample of 16 texts, 8 from the EFAL materials and 8 from the NST materials. I limited my selection from the EFAL to only those materials targeted at Grade 3 learners, on the basis that syntax is likely to grow in complexity across grade levels, and only adjacent grades can be considered comparable. For the EFAL texts, four were chosen from the Grade 3 EFAL Learner's Book and four were chosen from the Grade 3 EFAL Reader, with one text per term selected from each. All 8 texts for NST were chosen from the Grade 4 Learner's Book, with two texts per term selected.

The EFAL materials and the NST textbook are structured in different ways which influenced how texts were selected from each. For the EFAL materials, both the Learner's Book and Reader are subdivided into 16 themes, 4 per term. For the Reader, there is only a single text or in some cases two or three poems per theme followed by a few short questions. In the longer EFAL Learner's Book, each theme is divided into five sections: 1) Read and Listen; 2) Write; 3) Listen and Speak; 4) Listen, Speak and Play and 5) Sounds and Words. There is one main reading passage per theme which is clearly demarcated at the start of each theme in the 'Read and Listen' section. The remainder of the theme consists of a variety of activities as well as occasional grammar points and, in some cases, shorter reading texts (such as poems and rhymes).

The software used for the syntax analysis was not able to accurately process poetry and plays, nor any text not formatted as prose, therefore such texts were excluded from the EFAL selection. Among the remaining EFAL texts, I selected texts which represented a variety of genres (always from the main reading passage found at the start of the theme), with a bias towards texts falling in the category *Story Genres* as this comprises nearly half the EFAL materials.

In comparison, the Grade 4 NST textbook is divided into a total of 17 topics, five in Term 1 and four each in the remaining terms. Each topic consists of between one and three smaller units, beginning with a brief preview of what will be covered in the topic and concluding with a topic review. There are short 'Key Word', 'Safety', 'Did you know?' and 'Key Concepts' boxes spaced throughout the units. Beyond these, units consist primarily of prose explaining the topic interspersed with activities containing tasks and questions. In contrast to the EFAL materials which have clearly demarcated reading texts, units in the NST Learner's Book consist of far lengthier, almost continuous reading passages broken up by headings and subheadings and interspersed with the boxes and activities described earlier.

In selecting texts from the NST textbook, I excluded all previews, reviews, boxes and activities on the basis that these were either too brief or not in the prose format required by the syntax analysis software (all activities fell into the latter category). In general, the NST textbook consisted of far more running text than the EFAL materials, but units varied drastically in length, from between 300 words to more than 1000 words. For reference, Grade 3 EFAL reading texts generally range between 100 and 450 words. In order to select texts which were not much longer than those from the EFAL materials, but which still followed natural rather than artificial breaks, I selected whole units which were shorter in length, or where units were too long, sub-sections of those units. I also selected texts to represent a variety of genres with a bias towards *Information Reports* which make up nearly half the NST Learner's Book.

For all 16 texts, where pictures with captions accompanied the text these captions were retained and inserted at the end of the paragraph which addressed the picture and caption most directly. The selected texts can be found in full in Addendum D.

After selecting the texts for analysis, I formatted each text according to the specifications required by the TextEvaluator software. These specifications included ensuring clear paragraph breaks (which in some cases for the EFAL texts had to be inserted manually) as well as formatting all headings and subheadings to be the first sentence of a paragraph rather than a sentence on their own. The TextEvaluator software is freely accessible online at https://textevaluator.ets.org/TextEvaluator/. The results of my analysis can be found in the next section.

6.3. Results

Table 15 below lists the texts sampled for analysis in the chronological order that learners should encounter them, with Text 1 appearing at the start of Grade 3 and Text 16 appearing towards the end of Grade 4. The table also provides more detail on the content, genre (according to the TextEvaluator tool) and word token count of each text. Texts highlighted in pink were sourced from the EFAL materials while texts highlighted in green were sourced from the NST textbook.

TEXT #	SOURCE	TERM	TITLE	GENRE	TOKENS
Text 1	EFAL LB	Gr3 Term 1	The big match	Informational	209
Text 2	EFAL R	Gr3 Term 1	A special school day	Literary	155
Text 3	EFAL LB	Gr3 Term 2	Around the world for a cup of tea	Informational	186
Text 4	EFAL R	Gr3 Term 2	All kinds of houses	Informational	165
Text 5	EFAL LB	Gr3 Term 3	Maisie gets toothache	Literary	445
Text 6	EFAL R	Gr3 Term 3	The elephant and the rain	Literary	246
Text 7	EFAL LB	Gr3 Term 4	How Giraffe got a long neck	Literary	263
Text 8	EFAL R	Gr3 Term 4	Living things have lots of water in them	Informational	180
Text 9	NST LB	Gr4 Term 1	Living things	Informational	516
Text 10	NST LB	Gr4 Term 1	Hoot and Toot need a new home	Informational	331
Text 11	NST LB	Gr4 Term 2	Change of state	Informational	382
Text 12	NST LB	Gr4 Term 2	Examples of indigenous houses of Southern Africa	Informational	354
Text 13	NST LB	Gr4 Term 3	Energy can be stored and transferred	Informational	543
Text 14	NST LB	Gr4 Term 3	Noise pollution	Informational	326
Text 15	NST LB	Gr4 Term 4	Earth and space	Informational	361
Text 16	NST LB	Gr4 Term 4	Moon stories	Informational	379

Table 15: List of sampled texts for syntax analysis; LB = Learner's Book; R = Reader

The figure below displays the syntactic complexity ranking of the sampled texts according to the TextEvaluator tool, measured on a scale from 1-100, where 1 is easy and 100 is difficult.

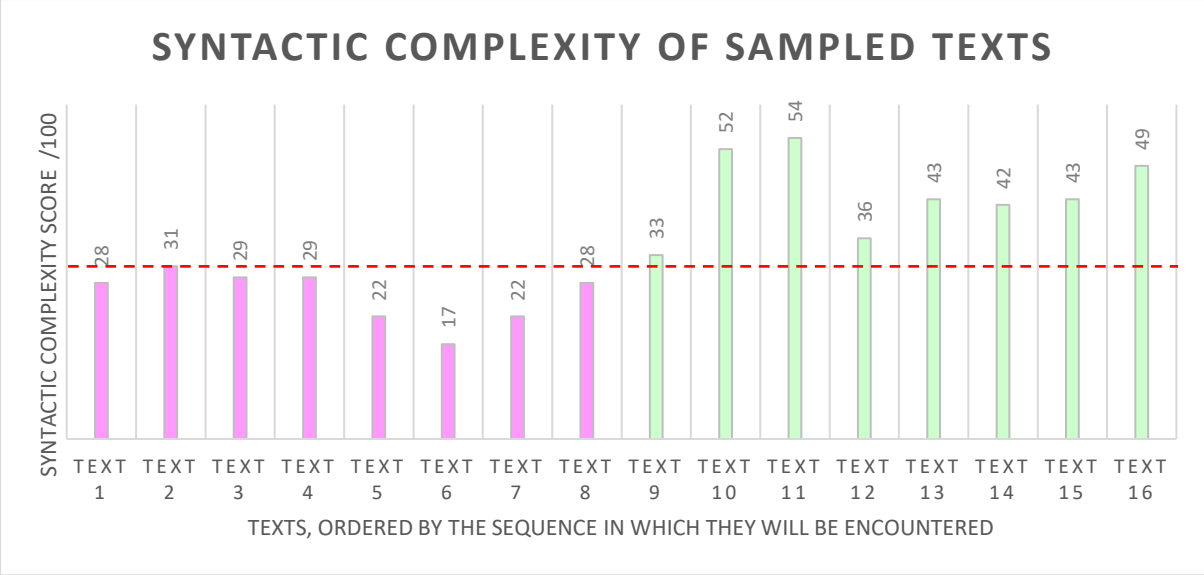


Figure 26: Syntactic complexity of sampled texts

The dotted red line in Figure 26 marks the highest complexity score found in the EFAL materials (a score of 31 for Text 2). It is clear from the figure that the NST texts (represented by the green columns) all score higher on the syntactic complexity rating than the EFAL texts (represented by the pink columns), with the EFAL complexity scores ranging from 17 to 31 while the NST complexity scores range from 33 to 54.

Note-worthy is the fact that the first text encountered in the NST textbook – Text 9 titled 'Living Things' studied at the start of Term 1 in Grade 4 – is the simplest of all the NST texts evaluated. With a complexity score of 33, it constitutes the lowest syntactic hurdle to clear based on the sample of what learners have seen in their EFAL materials. The transition from Text 2 (the most syntactically complex EFAL text with a score of 31) to Text 9 (with a score of 33) can be considered an acceptable one as these texts are close enough to one another in terms of complexity that learners should not struggle too much to adjust. The same can be said of Texts 3 and 4 (with scores of 29) and Texts 1 and 8 (with scores of 28) from the sampled EFAL materials.

Unfortunately, the complexity of the syntax found in the remaining sample of NST texts is generally much higher than that of the sampled EFAL texts and consequently more difficult to adjust to. In particular, the other sampled NST text encountered in Term 1 of Grade 4 – Text 10 titled 'Hoot and Toot need a new home' – has a very high syntactic complexity (with a score

of 52). This constitutes an increase of 21 points from Text 2 (the most syntactically complex EFAL text sampled) and as such, cannot be considered to have a comparable level of syntactic complexity to the EFAL materials. Nor can Texts 11 or 16 (with scores of 54 and 49 respectively). In other words, for at least three of the eight sampled texts encountered in Grade 4 (two of which are encountered early on), the EFAL materials have not adequately prepared learners for the complexity of syntax they will need to make sense of.

To give a sense of what these syntactic complexity differences look like in practice, I have included two excerpts from the sampled texts. Excerpt 1 is taken from the start of Text 2, 'A special school day', which has the highest syntactic complexity score for the sampled EFAL texts. Excerpt 2 is taken from Text 10, 'Case Study: Hoot and Toot need a new home', which received the second highest syntactic complexity score for the sampled NST texts.

EXCERPT 1: A special school day (EFAL Grade 3, Term 1) – SCORE: 31

> *Yesterday, I had an interesting day at school. The nurse from the clinic came. She showed us how to brush our teeth. She said, "Brush your teeth in the morning after breakfast. Brush your teeth in the evening before you go to bed." She gave us each a free toothbrush and toothpaste. 'Your teeth are precious!'*

EXCERPT 2: Case Study: Hoot and Toot need a new home (NST Grade 4, Term 1) – SCORE: 52

> *Hoot and Toot are two Spotted Eagle Owls that live at the National Zoological Gardens of South Africa in Pretoria. They go out into the zoo every day with their handlers so that visitors can see them up close. Their enclosure is in a quiet part of the zoo where visitors aren't allowed.*

Analysing a more comprehensive sample of texts from both sets of materials would have made the prevalence of gaps in complexity levels between the EFAL and the NST materials clearer. Unfortunately such an analysis was beyond the scope of this study. What is evident from the above, however, is that at least in some cases the NST textbook is written with much more complex syntax than the EFAL materials.

The TextEvaluator tool also assigns an overall complexity score to a text based on its eight areas of evaluation (of which Syntactic Complexity is only one). The overall complexity score ranges from 100 (appropriate for extremely young readers) to 2000 (appropriate for college graduates) according to the TextEvaluator guidelines. The eight areas of evaluation all have scores ranging from 1 to 100. The full set of scores for each text (ranked by overall complexity from highest to lowest) are displayed in Table 16 (see next page). Blue boxes indicate areas of analysis where a higher score is equivalent to greater complexity. Red boxes indicate areas where a lower score is equivalent to greater complexity. In other words, a high score for the category 'Academic Vocabulary' would mean that the text is considered more challenging in this regard while a high score for the category 'Concreteness' would mean that the text is considered more accessible in this regard.

			AREA A	AREA B			AREA C			AREA D
TEXT #	**TITLE**	**OVER-ALL SCORE**	**SCORE 1**	**SCORE 2**	**SCORE 3**	**SCORE 4**	**SCORE 5**	**SCORE 6**	**SCORE 7**	**SCORE 8**
Text 11	Change of state	**530**	54	38	61	61	86	23	25	22
Text 14	Noise pollution	**490**	42	50	69	59	76	57	15	56
Text 12	Examples of indigenous houses of Southern Africa	**480**	36	50	76	66	83	26	7	48
Text 10	Hoot and Toot need a new home	**430**	52	39	49	59	64	64	35	75
Text 16	Moon stories	**400**	49	32	63	67	78	46	23	70
Text 13	Energy can be stored and transferred	**370**	43	57	45	61	92	47	7	49
Text 2	A special school day	**320**	31	27	40	72	54	67	7	61
Text 9	Living things	**290**	33	38	56	62	100	62	44	29
Text 1	The big match	**290**	28	40	58	69	55	48	32	66
Text 15	Earth and space	**229**	43	27	28	70	95	16	69	63
Text 3	Around the world for a cup of tea	**50**	29	24	59	90	86	37	7	64
Text 4	All kinds of houses	**50**	29	19	36	94	68	37	7	44
Text 8	Living things have lots of water in them	**50**	28	38	28	81	78	59	34	11
Text 5	Maisie gets toothache	**50**	22	21	46	92	93	76	47	80
Text 7	How Giraffe got a long neck	**50**	22	12	30	74	69	65	69	89
Text 6	The elephant and the rain	**50**	17	14	18	82	64	78	76	75

Table 16: Full set of TextEvaluator scores for the sample texts

AREAS: A = SENTENCE STRUCTURE; B = WORD DIFFICULTY; C = CONNECTIONS ACROSS IDEAS; D = ORGANISATION

SCORES: 1 = SYNTACTIC COMPLEXITY; 2 = ACADEMIC VOCABULARY; 3 = WORD UNFAMILIARITY; 4 = CONCRETENESS; 5 = LEXICAL COHESION; 6 = INTERACTIVE/ COVERSATIONAL STYLE; 7 = LEVEL OF ARGUMENTATION; 8 = DEGREE OF NARRATIVITY

COLOURS: GREEN = S&T texts; PINK = EFAL texts; BLUE = high score equivalent to high complexity; RED = high score equivalent to low complexity

There are a few noteworthy points to be drawn from the rankings according to overall complexity. Firstly, six of the eight EFAL texts appear to be experiencing a floor effect in terms of overall complexity, with all six scoring 50 – below TextEvaluator's own stated scale of 100-2000. This is likely due to the fact that these materials have been designed for EFAL learners while the TextEvaluator tool is calibrated towards predominantly English L1 learners. Even so, it is worth noting that the gap between the score for the majority of the sampled EFAL texts (50) and the simplest NST text, Text 15, with a score of 229, is quite large. The gap becomes exceedingly large when comparing the six EFAL texts to the most complex NST text, Text 11, with a score of 530. These scores are displayed in Figure 27 (see below) with the texts ordered in the sequence with which learners are meant to encounter them.

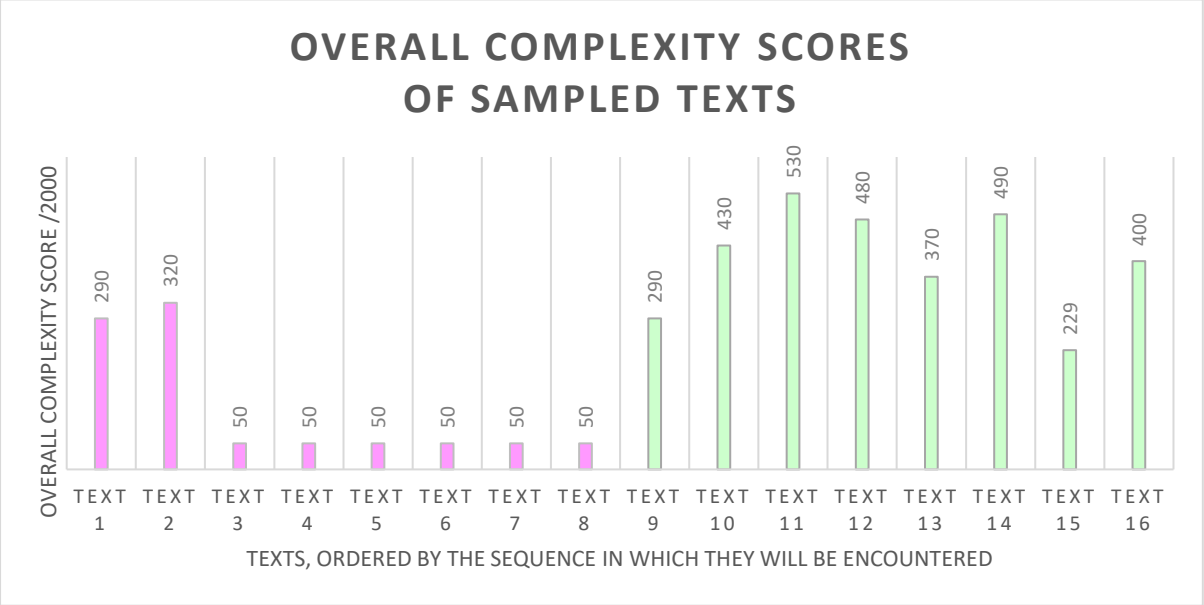


Figure 27: Overall complexity scores of sampled texts in chronological order

The overall complexity scores can also be loosely correlated with a particular US grade level by looking at the first digit of the score. In other words, a score of 530 would mean that the text can be considered appropriate for a US Grade 5 learner while a score of 480 means the text can be considered appropriate for a US Grade 4 learner and so forth. While these grade level estimates should be treated with caution given the difference in contexts and the fact that the EFAL materials are intended for ELLs, the comparison is still an interesting one to make. Using this yardstick, Texts 3, 4, 5, 6, 7, and 8 (which are targeted at South African Grade 3 EFAL learners) are considered below the level of a US Grade 1 learner. In contrast, for half of the sampled NST texts (Texts 10, 12, 14 and 16) there is a match between the intended South

African grade level (Grade 4) and the US grade level assigned by TextEvaluator, while Text 11 is above the average complexity level typically aimed at Grade 4 learners.

To give a sense of what these overall complexity differences look like in practice, I have included two additional excerpts from the sampled texts. Excerpt 3 is taken from the start of Text 6, 'The elephant and the rain', which has one of lowest complexity scores, while Excerpt 4 is taken from Text 11, 'Change of state', which received the highest complexity score.

EXCERPT 3: The elephant and the rain (EFAL Grade 3, Term 3) – OVERALL SCORE: 50

> *Elephant and Rain always argued about who was stronger. "I am the strongest." "No, I am."*
>
> *Rain said, "If I go away, you will all die." Elephant said, "I don't care. Go if you like. I will rule alone!"*
>
> *So one day, Rain went away. At first, nothing changed. There was grass to eat and the rivers were full of water.*
>
> *But after a while there was only one water hole left. The animals were thirsty.*

EXCERPT 4: Change of state (NST Grade 4, Term 2) – OVERALL SCORE: 530

> *As you learnt earlier, the three states of matter are solid, liquid and gas. When a material changes from a liquid to a solid, from a solid to a liquid, from a liquid to a gas, or from a gas to a liquid, we say that the material has a change of state.*
>
> ***Melting – changing from solid to liquid***
>
> *If solids are heated, they can change into liquids. For example, if we add heat energy to butter in a frying pan, it melts and becomes liquid. The photograph of the boy eating an ice-cream shows another example of a solid that is melting.*

To summarise, a comparative analysis of the syntax in the EFAL and NST materials reveals that all of the sampled EFAL materials are less syntactically complex than the NST materials. For the majority of the sampled EFAL materials, the gap in complexity is a substantial one, to the extent that these materials cannot be considered adequate preparation for the type of syntax learners will need to make sense of in a Gr 4 NST textbook.

Furthermore, when considering the overall text complexity score provided by the analytical tool, four of the sampled NST texts appear to be at the appropriate Grade 4 level (according to US metrics), while two others rank as Grade 2 Level, one as Grade 3 Level and one as Grade 5 level. In contrast, seven of the sampled EFAL materials intended for Grade 3 learners are ranked as being the incorrect grade level – six of the texts ranked as appropriate for a US Grade 1 learner, while one text ranked as appropriate for a Grade 2 learner. Only one of the sampled EFAL texts ranked as appropriate for a Grade 3 learner.

While I acknowledge that these EFAL materials are intended for ELLs learners and that this may account for some degree of simplification, this simplification is not nearly as prevalent in NST materials that the ELLs learners will have to transition to reading in Grade 4. There is thus a clear gap between what learners are being prepared for by the sampled Grade 3 EFAL materials and what is required of them by the sampled Grade 4 NST materials.

Chapter 7: Summary of Findings and Conclusions

7.1. Research Question and Results

In Chapter 1, I asked the following: Can children's learning of English in the subject English First Additional Language in Grades 1-3 prepare them to make sense of a Grade 4 Natural Sciences and Technology textbook? The findings of my analysis reveal that for all three key areas relating to text accessibility, namely Vocabulary, Genres and Syntax, the answer to this question is no.

Based on an analysis of materials approved by the DBE, South African learners are not being given adequate opportunity to learn (Gee, 2003) the kind of English that will enable them to make sense of a Grade 4 NST textbook. I shall briefly summarise my findings that support this conclusion below.

Regarding Vocabulary, my main findings were as follows: 1) Half of the word types found in the Grade 4 NST textbook are words learners have not encountered even once in three years of EFAL materials. 2) The percentage of running text learners can understand in their NST textbook (assuming they master all the vocabulary which appears in their Grade 1-3 EFAL scheme) is significantly below the threshold needed to read texts unassisted and dangerously close to the threshold below which no reader can make sense of text. 3) The EFAL scheme only exposes learners to around half the number of word families that research estimates ELLs need to read successfully in English; moreover, approximately a third of the word families learners encounter in their EFAL materials are not high frequency English words. 4) The findings of my study supports research which shows that around three thousand HFW word families (approximately 5000 word types) would enable students to understand the vocabulary found in their Grade 4 NST textbook. This reaffirms the fact that the CAPS vocabulary targets for EFAL (1500-2500 word types by the end of Grade 3), though they may be realistic given time allocations, are completely insufficient in relation to the minimum vocabulary needed to study through the medium of English.

Regarding Genre, the chief findings were as follows: 1) Almost half of the texts encountered in Gr1-3 EFAL fail to prepare learners for the predominant genre demands of a Grade 4 NST textbook. Learners are frequently exposed to *Descriptions*, *Recounts* and *Story* Genres in EFAL

and are likely to cope well with these in Grade 4 but have only limited exposure to *Information Reports*, *Procedures*, *Explanations* and *Definitions* which may prove more difficult for them to make sense of. 2) The amount of English text learners are required to read more or less doubles every year, with the jump from Grade 3 to Grade 4 being the largest. If one were to include the English texts that must be read for other subjects taught in Grade 4, this jump would be much greater. 3) For five out of the six genres compared qualitatively (*Information Reports, Procedures, Explanations, Descriptions* and *Story* Genres) there are significant differences between the EFAL and NST manifestations of each, such that it cannot be assumed learners will be able to smoothly transition to making sense of these genres in Grade 4. Only one genre, *Recounts,* manifests in both sets of materials in such a way that learners are likely to successfully engage with the genre in Grade 4.

Regarding Syntax, the findings were as follows: 1) In all of the texts sampled, the NST textbook is written with more complex syntax than that of the EFAL materials. In some cases the gaps in complexity between the two sets of materials are quite substantial. 2) For half of the sampled NST texts there is a match between the intended South African grade level and the US grade level assigned by the TextEvaluator tool. However, this was only the case for one of the sampled EFAL texts. Six of the EFAL texts which are targeted at South African Grade 3 EFAL learners were ranked as below the level of a US Grade 1 learner. There is thus a clear gap between what learners are being prepared for by the EFAL materials and what is required of them by the Grade 4 NST materials.

7.2. Discussion

At the start of this paper I stated that if the curriculum and learning materials approved by the DBE for English in the Foundation Phase do not adequately prepare children to study other subjects in English by the start of Grade 4, it would not be an exaggeration to say that South Africa's public education system has been designed to fail. The results of my study reveal that the sampled LTSMs approved by the DBE for the teaching of EFAL and NST in the early years of schooling are indeed fundamentally flawed. These LTSMs do not adequately accommodate the needs of linguistically diverse learners and fail to prepare them for the English language and literacy demands of classroom learning, specifically in the areas of vocabulary, genre and syntax. Given this, learners obtaining poor results in content learning undertaken through the medium of English seems a foregone conclusion.

The major design flaw underpinning this status quo is the failure of policy makers and publishers alike to recognise the diverse and socially situated nature of language and literacy resources (Blommaert & Backus, 2011; Gee, 2003). It has simply been assumed that a general introductory course of English education (completely isolated from the existing language resources of both learners and teachers in South African classrooms) will suffice to enable access to English-medium schooling after three years. There is no recognition of the need to provide learners with meaningful, situated opportunities to engage with the particular language and literacy resources required by specialised subject areas like Natural Sciences and Technology, Social Sciences, Mathematics and Life Skills to prepare them for English-medium instruction (Plüddeman, 2015).

While learners are acquiring some English resources in their three years of EFAL studies, these are centred around the disciplinary language of English-as-a-subject (Plüddeman, 2015). English-as-a-subject requires learners to be familiar with concepts and terminology related to grammar (e.g. parts of speech, tenses, punctuation, etc.), the study of literature (e.g. genres, narrative techniques, poetic devices, etc.) and the art of public speaking (e.g. fluency, voice projection, body language, etc.), to name just a few examples. Though there is some overlap with other subjects in terms of themes and topics covered in the reading materials of English-as-a-subject, in general, the language and literacy resources developed align quite poorly with what is required by content area learning in Grade 4. This has been evident in my research findings for the subject Natural Sciences and Technology.

I would like to emphasise here that the scale of the problem is considerable. While I focused on one content subject in this paper, learners are currently required to study five subjects through the medium of English in Grade 4. The demands that these additional subjects place on the English language and literacy resources of the ELLs entering Grade 4 far exceed what has been described here. The recommendations I outline below should be considered with this in mind.

7.3. Recommendations

According to the time allocated by CAPS for the study of English First Additional Language in grades 1 to 3, ELLs have between 308 and 440 hours of class time before entering Grade 4

to acquire the necessary language and literacy resources to make sense of learning five specialised content areas through the medium of English (DBE, 2011). Based on my research, it is not possible for ELLs who are not receiving additional English support outside of school to learn enough English in this amount of time to succeed in their lessons in Grade 4. There is simply not enough time to teach and learn everything that is necessary to cope with a sudden immersion into English as the LOLT in Grade 4, whether it be English vocabulary, genre knowledge, syntactic knowledge or additional language and literacy skills. Learners and their teachers must be given more time, and more support to transition from home language to English as the LOLT.

My chief recommendation is thus that learners no longer transition to English as the LOLT in Grade 4. Instead, I propose South Africa implement one of two alternative models. The first option, and my preferred alternative, is that South African schools adopt a bilingual (or where applicable, multilingual) LOLT throughout primary schooling. In this model, both the relevant home languages and English should be utilised in the study of *all learning areas* from the beginning of schooling. While the dominance of English is itself a highly politicised issue, it remains a valuable resource given the linguistic capital (Bourdieu, 1977) it affords for accessing tertiary education and employment opportunities, both within South Africa and abroad. Consequently, it remains necessary to empower South African learners to use the English language with mastery, whilst also teaching them to view its dominance in a critical light (see Janks, 2004, for a more detailed discussion on a critical approach to teaching English).

Adopting a bilingual or multilingual approach in all learning areas from the start of schooling means that teachers would need to teach every subject in every grade in a way that equips learners to both make sense of and produce language for that particular subject in their home languages as well as English. EFAL would no longer have its own slot in the timetable, but rather be woven into every lesson, with a focus on developing the specialised language and literacy resources for that particular subject. In this way, we can avoid a situation where learners develop language resources suited to English-as-a-subject but which do not equip them for the demands of content subject learning. Although children may enter Grade 1 with little existing English knowledge, they will gradually and consistently be introduced to the

specialised English vocabulary, genres and syntax utilised in all learning areas, and over time begin to develop a familiarity with the language as it manifest in the classroom.

Where explicit language instruction is required, e.g. to understand how English word order and sentence structure differs from isiXhosa (or Sesotho, or any home language), this should be done in a single, consolidated language arts period, where all the languages of the classroom are utilised and studied. This will have multiple benefits. Firstly, students will be able to analyse and compare the different language structures and patterns with the help and guidance of a teacher. Secondly, instead of duplicating the teaching of specialised language skills such as reading, writing, public speaking and the study of literature in both home language and EFAL periods, these skills can become the focus of a single period. Teachers will be able to, and indeed will have to draw on the multiple languages of the classroom when choosing literature and setting tasks and assignments.

This recommendation has major implications for two areas: the design of LTSMs and the training of teachers. For the former, in order for bilingual instruction to be meaningful, the LTSMs used in all subjects would need to reflect this bilingual approach. This would require that new textbooks and other LTSMs be developed for all content subjects across the primary grades which would simultaneously utilise both the home languages of learners as well as English. In practice, this would mean things like bilingual glossaries and instructions, explanations in a combination of languages and also tasks and activities as well as formal assessments that require learners to draw on both their home language and English language and literacy resources. Examples of the latter might include speaking in one language and then writing in another, speaking and writing in both languages in alternation, learners and teachers translating texts, etc. (Garcia & Li, 2014; Nation, 2001; Hall & Cook, 2012). See ‘Home Language’ section in Appendix A for a more detailed discussion on the value of the above-mentioned strategies.

Developing such materials well will be no small undertaking. It will require the collective efforts of experts from multiple fields, including individuals knowledgeable in the structure and teaching of all the South African home languages, experts in the field of bilingual education, as well as bilingual teams of textbook writers. Fortunately, there is a wealth of research related to best practice in bilingual education (Clegg & Afitska, 2011; Hall & Cook,

2012; Barrett & Bainton, 2016; Milligan et al., 2016; Clegg & Simpson, 2016; Lin, 2019) which can be drawn on to develop such materials and guide classroom practice.

It is important to note here that no one size will fit all. The exact nature of bilingual instruction in a given classroom will depend on the particular language resources of both the learners and the teachers in that classroom (Lin, 2019), and LTSMs will need to make allowances for this. Exactly because of this rich diversity, and in order to utilise the bilingual materials which are developed appropriately, existing teacher training programmes will need to be revised to allow for more specialised training and guidance in bilingual teaching practices. Moreover, additional time and resources will need to be allocated within curricula to enable pre-service teachers to master additional languages of teaching and learning beyond the ones they may already have facility with. While practically difficult to execute, these are essential investments that must be made in order for all other teaching to be successful.

A second alternative model which could be implemented, should the above proposal be considered too drastic a shift, would be to delay the transition to English as a LOLT from Grade 4 to either Grade 6 or 7, in order to provide two to three additional years in which to learn more English. Schools could also gradually phase in the English LOLT, by implementing a bilingual approach at the initial point of transition and then slowly phasing out home language use. Should this approach be preferred, the EFAL curriculum will require considerable revision in order to more closely reflect the language and literacy demands of the content subjects to be taught in English in later grades.

Beyond the above (and regardless of which approach is taken), I have three further recommendations. Firstly, because LTSMs play such a key role in shaping classroom practice, I believe that it would be valuable for the DBE to mandate a comparative evaluation of LTSMs using a framework similar to the one introduced in Chapter 3 of this study, approximately every 5 years in order to scrutinise what publishers are producing and incentivise them to improve the quality of their LTSMs. Such an evaluation will also be vital should the EFAL course of study need to be revised to reflect the specific language demands of the various content subjects.

Secondly, in terms of English vocabulary, the findings of my study indicate that international high frequency word lists such as the GSL, AWL and BNC-COCA lists are promising resources to guide English textbook developers in their development of materials. However, there would also be merit in generating a high frequency English word list for South African learners, and this should be put on the research agenda of the DBE.

Thirdly and finally, in relation to the syntax and overall text complexity of school LTSMs, there needs to be a closer alignment and gradual progression in terms of what learners encounter in their reading materials when transitioning from one grade to the next. Too many of the texts found in the Grade 3 EFAL materials were oversimplified as compared with the texts found in the Grade 4 NST textbook. Publishers need to ensure that the syntactic complexity of texts in adjacent grades are more comparable to allow for a smooth reading transition for learners.

7.4. Conclusion

In many ways, this entire research paper merely serves to reiterate (with more recent theory and data as evidence) what has long been acknowledged by the majority of teachers working in South African classrooms – that our abrupt transition to English medium instruction in primary school is failing learners. We cannot continue to implement, year-after-year, a curriculum and materials that offer no possibility of success for the majority of learners whilst ignoring the poor educational outcomes that necessarily follow from this. It is time to act on the knowledge and expertise already widely in circulation regarding language learning and education. Only by doing so can we finally offer South African children the real possibilities for future success that they have too long been denied.

References

Babbie, E. 2014. *The Basics of Social Research*. 6th edition. Wadsworth: Cengage Learning.

Bailey, B. 2007. Heteroglossia and Boundaries. In Heller, M. (ed). *Bilingualism: A social approach.* Hampshire and New York: Palgrave Macmillan.

Bailin, A. & Grafstein, A. 2016. *Readability: Text and Context.* London: Palgrave Macmillan. DOI: https://doi.org/10.1057/9781137388773

Barrett, A. M. & Bainton, D. 2016. Re-interpreting relevant learning: an evaluative framework for secondary education in a global language. *Comparative Education*, 52(3):392-407.

Barrett, A. M., Mtana, N., Osaki, K., & Rubagumya, C. 2014. *Language Supportive Teaching and Textbooks: Baseline Study Report.* Bristol: Language Supportive Teaching and Textbooks.

Biemiller, A. 2006. Vocabulary Development and Instruction: A Prerequisite for School Learning. In Neuman, S.B. & Dickinson, D.K. (eds). 2006. *Handbook of Early Literacy Research, Volume 2.* New York, London: The Guilford Press, 41-51.

Blommaert, J. & Backus, A. 2011. Repertoires revisited*:* 'Knowing language' in superdiversity. *Working Papers in Urban Language and Literacies,* 67:1–26.

Bourdieu, P. 1977. The economics of linguistic exchanges. *Social Science Information*, 16(6):645-668. DOI:10.1177/053901847701600601

Cameron, D. 2001. *Working with spoken discourse.* London: Sage Publications.

Canagarajah, A. S. 1999. *Resisting linguistic imperialism in English teaching.* Oxford: Oxford University Press.

Carney, R. N. & Levin, J. R. 2002. Pictorial illustrations still improve students' learning from text. *Educational Psychology Review*, 14(1):5-26, March.

Carter, R. 2001. Vocabulary. In Carter, R. & Nunan, D. (eds). *The Cambridge Guide to Teaching English to Speakers of Other Languages.* Cambridge: Cambridge University Press, 42-47.

Chu, Y. 2017. Twenty years of Social Studies Textbook Content Analysis: Still "Decidedly Disappointing?" *The Social Studies*, 108(6):229-241.

Clegg, J. & Simpson, J. 2016 Improving the effectiveness of English as a medium of instruction in sub-Saharan Africa. *Comparative Education*, 52(3):359-374.

Clegg, J. & Afitska, O. 2011. Teaching and Learning in Two Languages in African Classrooms. *Comparative Education*, 47(1):61–77.

Cummins, J. 2008. BICS AND CALP: Empirical and Theoretical Status of the Distinction. In Street, B. & Hornberger, N. H. (eds). *Encyclopedia of Language and Education, 2nd Edition, Volume 2: Literacy*. New York: Springer Science and Business Media LLC: 71-83.

Dale, E. & Chall, J.S. 1949. The concept of readability. *Elementary English*. 26(1):19-26, January.

Department of Education. 1997. Language in Education policy, 14 July 1997. Pretoria: Department of Education.

Department of Basic Education. 2011. *Curriculum and Assessment Policy Statement Grades 1-3 English First Additional Language.* Available at: https://www.education.gov.za/Curriculum/CurriculumAssessmentPolicyStatements(CAPS)/CAPSFoundation.aspx [1 October 2019].

Department of Basic Education. 2019. *Report on the 2019 National Senior Certificate Diagnostic Report Part 2*. Pretoria: DBE. Available at: https://www.education.gov.za/Resources/Reports.aspx [28 February 2021]

Derewianka, B. 1990. *Exploring How Texts Work.* Sydney: Primary English Teaching Association.

Derewianka, B. 2003. Trends and Issues in Genre-Based Approaches. *RELC Journal,* 34(2):133-154.

Derewianka, B. 2012. Knowledge about Language in the Australian Curriculum: English, *Australian Journal of Language and Literacy,* 35(1):127-146.

Desai, Z. 2016. Learning through the medium of English in multilingual South Africa: enabling or disabling learners from low-income contexts? *Comparative Education*, 52(3):343-358.

Fabrício, B. F. & Santos, D. 2006. The (re-)framing process as a collaborative locus for change. In Edge, J. (ed.), *(Re)locating TESOL in an age of empire.* Basingstoke: Palgrave Macmillan, 65–83.

Fromkin, V., Rodman, R. & Hyams, M. 2011. *An Introduction to Language.* 9th ed. Boston, MA: Wadsworth, Cengage Learning.

Garcia, O. & Wei, L. 2014. *Translanguaging: Language, Bilingualism and Education.* Basingstoke: Palgrave Macmillan.

Gee, J. P. 2003. Opportunity to Learn: a language-based perspective on assessment. *Assessment in Education,* 10(1):27-46.

Gee, J. P. 2015. Discourse, small-d, Big D. In Tracy, K., Ilie, C. & Sandel, T. (eds). *International Encyclopedia of Language and Social Interaction.* Wiley-Blackwell & ICA.

Goodman, K. 1984. Unity in Reading. In Purves, A. C. & Niles, O. eds. *Becoming Readers in a complex society.* Eighty-third yearbook of the national society for the study of education. Chicago: University of Chicago Press.

Graesser, A. C., McNamara, D. & Kulikowich, J. M. 2011. Coh-Metrix: Providing Multilevel Analyses of Text Characteristics. *Educational Researcher*, 40(5):223-234.

Graesser, A. C., McNamara, D. Louwerse, M. M. & Cai, Z. 2004. Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers,* 36:193-202.

Gunning, T. G. 2003. The Role of Readability in Today's Classrooms. *Top Language Disorders*, 23(3):175-89.

Hall, G. & Cook, G. 2012. Own-language use in language teaching and learning. *Language Teaching*, 45(3):271-308.

Hammond, J. & Derewianka, B. 2001. Chapter 27: Genre. In Carter, R. & Nunan, D. (eds). In *The Cambridge Guide to Reaching English to Speakers of Other Languages.* New York: Cambridge University Press.

Heath, S. B. 1982. What no bedtime story means: Narrative Skills at Home and School. *Language in Society*, 11(1):49-76.

Heugh, K. 2013. Multilingual Education Policy in South Africa Constrained by Theoretical and Historical Disconnections. *Annual Review of Applied Linguistics*. 33: 215-237.

Horst, M., Cobb, T. & Meara, P. 1998. 'Beyond a Clockwork Orange: acquiring second language vocabulary through reading'. *Reading in a Foreign Language*, 11(2):207-223.

Hu, M. & Nation, P. 2000. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language.* 13(1):403-430.

Hymes, D. H. 2005. On Communicative Competence. In Pride, J. B. & Holmes, J. (eds). *Sociolinguistics: Selected Readings*. Harmondsworth: Penguin:269-293.

Janks, H. 2004. The Access Paradox. *English in Australia*, 12(139):33-42.

Joe, A. 1995. 'Text-based tasks and incidental vocabulary learning'. *Second Language Research*, 11(2):149-158.

Joe, A. 2010. 'The quality and frequency of encounters with vocabulary in an English for Academic Purposes programme'. *Reading in a Foreign Language*, 22(1):117-138.

Katunich, J. 2006. Equity and English in South African higher education: Ambiguity and colonial language legacy. In Edge, J. (ed.), *(Re)locating TESOL in an age of empire.* Basingstoke: Palgrave Macmillan, 139–157.

Kim, K. & Clariana, R. B. 2017. Text signals influence second language expository text comprehension: knowledge structure analysis. *Educational technology research and Development*, 65:909-930.

Laufer, B. 1992. How Much Lexis is Necessary for Reading Comprehension? In Arnaud, P.J.L. & Béjoint, H. (eds.), *Vocabulary and Applied Linguistics*. London: Macmillan, 126-132.

Laufer, B. & Hulstijn, J. 2001. Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement. *Applied Linguistics*, 22(1):1-26.

Laufer, B. & Ravenhorst-Kalovski, G. C. 2010. Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1):15-30.

Laufer, B. & Rozovski-Roitblat, B. 2015. Retention of new words: Quantity of encounters, quality of task and degree of knowledge. *Language Teaching Research*, 19(6):687-711.

Laufer, B. & Sim, D. D. 1985. Measuring and Explaining the Reading Threshold Needed for English for Academic Purposes Texts. *Foreign Language Annals,* 18(5):405-411.

Lemarié, J., Lorch, R. F. & Péry-Woodley, M. 2012. Understanding how headings influence text processing. *Discours,* 10:1-22, July. DOI: 10.4000/discours.8600

Liddell, C. 1997. Every picture tells a story – or does it? Young South African children interpreting pictures. *Journal of Cross-Cultural Psychology*, 28(3): 266-283, May.

Lin, A. M. Y. 2019. Theories of trans/languaging and trans-semiotizing: implications for content-based education classrooms. International Journal of Bilingual Education and Bilingualism, 22(1):5-16. DOI: 10.1080/13670050.2018.1515175

Lipka, O. & Siegel, L.S. 2012. The development of reading comprehension skills in children learning English as a second language. *Reading and Writing*, 25(8):1873-1898.

Littlejohn, A. 2011. The analysis of language teaching materials: inside the Trojan Horse. In Tomlinson, B. (ed). *Materials Development in Language Teaching*. 2nd ed. Cambridge: Cambridge University Press: 179-211.

Littlewood, W. & Yu, B.-H. 2011. First language and target language in the foreign language classroom. *Language Teaching,* 44:64–77.

Lorch, R. F. 1989. Text-Signalling Devices and Their Effects on Reading and Memory Processes. *Educational Psychology Review*, 1(3):209-234.

Macaro, E. 2006. Codeswitching in the L2 classroom: A communication and learning strategy. In E. Llurda (ed.), *Non-native language teachers: Perceptions, challenges and contributions to the profession*. Amsterdam: Springer, 63–84.

Macdonald, C. & Burroughs, E. 1991. *Eager to Talk, Learn and Think: Bilingual Primary Education in South Africa*. Cape Town: Maskew Miller Longman.

Macdonald, C.A. 1990. *Crossing the threshold into Standard Three in Black Education: The Main Report of the Threshold Project*. Pretoria: HSRC.

Malmkjær, K. (ed.). 1998. *Translation & language teaching: Language teaching & translation.* Manchester, UK: St. Jerome Publishing.

Mayaba, N., Otterup, T. & Webb, P. 2013. Writing in Science Classrooms: Some Case Studies in South African and Swedish Second-language Classrooms. *African Journal of Research in Mathematics, Science and Technology Education,* 17(1-2)74-82. DOI: 10.1080/10288457.2013.826972.

McKinney, C. 2017. *Language and Power in Post-Colonial Schooling: Ideologies in Practice.* New York and London: Routledge.

McKinney, C. & Tyler, R. 2019. Disinventing and reconstituting language for learning Science. *Language and Education*, 33(2):141-158. DOI: 10.1080/09500782.2018.1516779

Mbude, N. N. 2020. 'IsiXhosa as the language of teaching and learning mathematics in Grade Six: investigating the mother tongue based bilingual education mathematics pilot in the Eastern Cape Province, South Africa'. PhD thesis. Rhodes University, Makhanda.

Milligan, L. O., Clegg, J. & Tikly, L. 2016. Exploring the potential for language learning in English medium instruction: a Rwandan case study. *Comparative Education*, 52(3):328-342.

Mini, B. M. and Botha, C. R. 2020. Eastern Cape's Language-in-Education policy implementation initiative for bilingual education: A historical and onion metaphor conceptualization. *South African Journal of African Languages*, 40(2):139-147.

Nation, I. S. P. 2001. *Learning Vocabulary in Another Language.* Cambridge: Cambridge University Press.

Nation, P. 2006. Vocabulary: Second Language. In Brown, K. 2006. *Encyclopedia of Language & Linguistics.* 2nd edition. Boston: Elsevier, 448-454.

Nation, I. S. P. 2012. The BNC/COCA word family lists. Available at: https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/vocabulary-lists [1 June 2021]

Nelson, J., Perfetti, C., Liben, D. & Liben, M. 2012. *Measures of Text Difficulty: Testing their Predictive Value for Grade Levels and Student Performance*. Technical Report. Available at: https://achievethecore.org/page/1196/measures-of-text-difficulty-testing-their-predictive-value-for-grade-levels-and-student-performance [16 July 2021]

Nilagupta, S. 1977. The relationship of syntax to readability for ESL students in Thailand. *Journal of Reading.* 20(7):585-594, April.

Ouane, A. & Glanz, C. (eds). 2011. *Optimising Learning, Education and Publishing in Africa: The Language Factor.* Hamburg and Tunis: UIL-UNESCO & ADEA.

Pitler, E. & Nenkova, A. 2008. Revisiting Readability: A Unified Framework for Predicting Text Quality. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Honolulu, Hawaii, USA. 25-27 October 2008.

Plüddeman, P. 2015. Unlocking the grid: language in-education policy realisation in post-apartheid South Africa. *Language and Education,* 29(3):186-199. DOI: 10.1080/09500782.2014.994523.

Polio, C. & Duff, P. 1994. Teachers' language use in university foreign language classrooms: A qualitative analysis of English and target language alternation. *The Modern Language Journal*, 78(3):313–326.

Probyn, M. 2001. Teachers' voices: Teachers' reflections on learning and teaching through the medium of English as a second language. *International Journal of Bilingual Education and Bilingualism,* 4(4):249–66.

Probyn, M. 2009. 'Smuggling the vernacular into the classroom': conflicts and tensions in classroom codeswitching in township/rural schools in South Africa. *International Journal of Bilingual Education and Bilingualism*, 12(2):123-136. DOI: 10.1080/13670050802153137

Probyn, M. 2015. Pedagogical translanguaging: bridging discourses in South African science classrooms, *Language and Education*, 29(3):218-234. DOI: 10.1080/09500782.2014.994525

Rott, S. 1999. The effect of exposure frequency on intermediate language learners' incidental vocabulary acquisition and retention through reading. *Studies in Second Language Reading,* 21(4):589-619.

Saragi, T., Nation, I. S. P. & Meister, G. F. 1978. Vocabulary learning and reading. *System,* 6:72-78.

Schmidt, R. 2001. Attention. In P. Robinson (ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139524780.003

Schmitt, N., Cobb, T., Horst, M. & Schmitt, D. 2017. 'How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2006)'. *Language Teaching*, 50(2):212-226.

Seligmann, J. 2012. *Academic Literacy for Education Students.* Revised ed. Cape Town: Oxford University Press.

Sheehan, K. M., Flor, M. & Napolitano, D. 2013. 'A Two-Stage Approach for Generating Unbiased Estimates of Text Complexity'. *Proceedings of the 2nd Workshop of Natural Language Processing for Improving Textual Accessibility (NLP4ITA)*. Atlanta, Georgia, USA. 14 June 2013.

Sheehan, K. M., Kostin, I. Napolitano, D. & Flor, M. 2014. The TextEvaluator Tool: Helping Teachers and Test Developers Select Texts for Use in Instruction and Assessment. *The Elementary School Journal*, 115(2):184-209.

Sibanda, L. 2014. The readability of two Grade 4 natural sciences textbooks for South African schools. *South African Journal of Childhood Education,* 4(2):154-175.

Statistics South Africa. 2018. *General Household Survey 2018*. Pretoria: Statistics South Africa. Available at: https://www.statssa.gov.za/publications/P0318/P03182018.pdf [2 November 2021].

Stahl, S.A. 2003. 'Vocabulary and Readability: How Knowing Word Meanings Affects Comprehension'. *Topics in Language Disorders,* 23(3):241-247.

Storch, S.A. & Whitehurst, G.J. 2002. 'Oral Language and Code-Related Precursors to Reading: Evidence From a Longitudinal Structural Model'. *Developmental Psychology*, 38(6):934-947.

Street, B. 1992. Introduction: The New Literacy Studies. In: Street, B. ed. 1992. *Cross Cultural Approaches to Literacy*. Cambridge: Cambridge University Press.

Van Broekhuizen, H., Van der Berg, S. & Hofmeyr, H. 2016. Higher Education Access and Outcomes for the 2008 National Matric Cohort. *Economic Working Paper 16/16*. Bureau for Economic Research: University of Stellenbosch. Available at: https://resep.sun.ac.za/wp-content/uploads/2016/10/Van-Broekhuizen-et-al.pdf [29 October 2019].

Van Rooyen, H. 1990. *The disparity between English as a subject and English as the medium of learning: A final report of the Threshold Project.* Pretoria: HSRC.

Vekiri, I. 2002. What is the value of graphical displays in learning? *Educational Psychology Review.* 14(3):261-312, September.

Waring, R & Takaki, M. 2003. 'At what rate do learners learn and retain new vocabulary from reading a graded reader?' *Reading in a Foreign Language*, 15(2):130-163.

Webb, S. 2007. 'The Effects of Repetition on Vocabulary Knowledge'. *Applied Linguistics*, 28(1):46-65. DOI: 10.1093/applin/aml048

West, M. 1953. *A general service list of English words*. London: Longman, Green & Co.

Witte, A., Harden, T. & Ramos de Oliveira Harden, A. 2009. *Translation in second language learning and teaching*. Bern: Peter Lang.

Zahar, R. Cobb, T. & Spada, N. 'Acquiring Vocabulary through Reading: Effects of Frequency and Contextual Richness'. *The Canadian Modern Language Review*, 57:541-572.

Appendices

Appendix A – Support Features of Text

Visuals

Visuals refers to images (photographs and illustrations) as well as graphics (diagrams, graphs, maps and charts) which accompany texts. According to Carney and Levin (2002) well-selected or well-constructed images reliably improve the reading-to-learn process.

Not all images perform the same functions or add equal value when combined with texts. Drawing on earlier work by Levin, Carney and Levin (2002) provide the following five categories of images: decorational, representational, organisational, interpretational and transformational. Decorational images serve only to decorate the page and are not significantly related to the text content. Representational images reflect the meaning of at least part of the text (these are the most common type of illustration). Organisational images generate a structural framework to organise text content. Interpretational images serve to clarify the meaning of complex text. Finally, transformational images (which occur very rarely) include memory enhancing features which can assist readers to remember the information contained in the text.

How images are used also depends on learner characteristics such as the age of the learner, the learner's reading ability and the learner's visual literacy (ability to read images) (Peeck in Carney & Levin, 2002). With regards to visual literacy, Liddell (1997) argues that children must be taught to interpret images and understand their conventions, highlighting that these skills will not simply develop as a function of maturation.

Liddell (1997), whose research examined the image interpretation skills of South African children in rural Limpopo, highlighted that rural children have less exposure to picture books than their urban counterparts and sometimes misinterpret certain picture elements which relate to perspective, depth and shading. She reported that school becomes one of the main sites where visual literacy skills develop for these children.

To be able to evaluate the quality of visuals accompanying text, Vekiri (2002) identified three characteristics which mediate the impact of images on student learning. Vekiri argues that, in

order for images to have a positive impact: they need to address the goal of the task; they should be provided along with explanations and guidance; and they need to be spatially and timeously coordinated with the text. Regarding the first point, this links to the different functions of images identified by Carney and Levin (2002) summarised earlier. In short, not all images are good for learning. Only images whose function align with the goal of the task are fit for purpose.

Regarding the second point, the key finding of research has been that students need to be explicitly instructed (by the text itself or by a teacher) to make use of the images in a text. Rieber (in Vekiri, 2002:275) reported that "students often do not know what information they need to observe in a display, and they are likely to draw wrong conclusions from what they see". Regarding the third point of coordinating images with text, Vekiri (2002) states that images must be spatially close or presented simultaneously with verbal informal information in order to be effective. Carney and Levin (2002:21) echo this principle when they highlight the need for images to be "proximally adjunct" to text.

I believe these three principles of Vekiri (2002) combined with the classification system provided by Carney and Levin (2002) for categorising images could provide a useful framework for evaluating the impact of visuals in supporting text accessibility.

Home Language

The use of a learner's home language (L1) in classrooms where the subject or medium of instruction is an L2 has long been a common-sense approach of teachers but one which has historically been frowned upon (Probyn, 2001). Proponents of monolingual teaching (where the use of a learner's L1 is forbidden) argue that an L2 immersion environment is the most effective means of teaching a language. There are also pragmatic benefits to this approach, such as instructors not needing to be proficient in the languages of students (especially useful in classes where students are from diverse language backgrounds) and publishers not needing to adapt language learning materials for different language groups.

In recent years, however, the monolingual bias in language and content teaching has been exposed (McKinney, 2017; Garcia & Wei, 2014) and has begun giving way to a wider acceptance of bilingual approaches. There is now an increasing recognition of the value and

benefit of making strategic use of learners' own languages in teaching. For a fuller discussion on the transition from monolingual to bilingual approaches as well as the theoretical perspectives that underpin each, see Hall and Cook (2012).

There are many reasons to make judicious use of a child's home language in teaching a new language. Exactly how much of the home language should be used, in what way and in which situations remains a topic of much debate in the literature. Nonetheless, there are clear benefits for the use of L1 in at least some cases and to some extent. These benefits can broadly be divided into three categories: pedagogical benefits, social and affective benefits and political issues relating to the status and power of languages.

Regarding pedagogical benefits, using the L1 can assist teachers in classroom management, giving complex instructions to learners efficiently and clearly as well as providing quick and easy to understand translations of grammatical points and terms that are already familiar concepts in the learner's L1. The benefits of translating vocabulary in particular are well documented, including the benefits of bilingual dictionaries and glossaries (Nation, 2001; 2006; Hall & Cook, 2012). The learner's home language can also be a valuable resource for negotiating understanding in small groups (peer-to-peer learning) and can be a valuable tool for teachers to check understanding.

Macaro (2006) explains that "code-switching and immediate translation during breakdowns in classroom discourse can also lighten the cognitive load on learners, facilitating the processing of other input." Using the L1 can also enable teachers to engage with learners in a more cognitively demanding fashion (e.g. to question and critique what is being taught), than their current L2 skills allow for (Hall & Cook, 2012).

Furthermore, Witte et al. (2009) state that engaging in translation activities is "a means through which learner's language awareness, intercultural competence and understanding of conceptual metaphors and literary texts may be developed (in addition to the benefits provided to learners' fluency and accuracy)". Indeed, Malmkjær (1998) describes translation as a distinct and valuable 'fifth skill' for language learners to develop which draws on the other four skills of listening, speaking, reading and writing.

Moving on to the social and affective benefits of using the L1, learners can easily experience high levels of fear and anxiety in L2 immersion environments which inhibit their ability participate in lessons and absorb what is being taught. Littlewood and Yu (2011) describe how monolingual teaching can 'alienate' learners. Teachers can use the L1 to reassure and develop rapport with learners to counteract such affective barriers. Polio and Duff (1994) describe how teachers make use of the L1 to demonstrate empathy or show solidarity while Canagarajah (1999) details how her use of the learners own-language puts them at ease and creates a less-threatening environment.

There are also significant political issues relating to the status and power of languages that come into play when utilising monolingual or bilingual approaches in teaching. Fabrício and Santos (2006) explain that the use of children's home languages can create an awareness of how the use of one language over another has been influenced by the history and power relationships between different languages. This is an especially important issue in the South African context because of the country's linguistic diversity and history of oppression (e.g. the 1976 Soweto uprisings). Katunich (2006) explains how "English only or English-mainly teaching in post-apartheid South Africa is a form of colonialism which denies black [...] language learners the possibility of additive multilingualism and limits their identities and educational trajectories."

Hall and Cook (2012:276) also highlight how the monolingual approach is underpinned by an assumption that ELLs wish to develop native-like proficiency in a language. They argue that such goals are not only unrealistic but out of touch with the global reality of English today, where the language functions as a *lingua franca* among non-native speakers wishing to preserve their own cultural and linguistic identity. Bilingual approaches to teaching are much more in line with this reality.

Given the benefits described above, it seems clear that any course of English language instruction and transition to English-medium instruction in South Africa needs to make allowance for the use of L1 in the classroom. Learning materials could even support teachers by guiding the extent and nature of L1 use in the classroom and by including activities designed in line with best practices in bilingual teaching.

I acknowledge that the number of official languages in South Africa means that publishers would need to make significant efforts to adapt materials for the country's variety of language groups, as well as for linguistically homogenous classrooms (such as those found in parts of the rural Eastern Cape and KwaZulu-Natal) and heterogenous classrooms (such as those found in urban areas). However, I believe this is an inescapable necessity for a multilingual society that wishes to support the academic development of its learners and protect the status of different languages. While more research is needed to develop a rigorous framework for evaluating the inclusion of L1 in L2 teaching materials, the presence of home language support in these materials needs to be seen as non-negotiable.

Sign-posting

Sign-posting (also known as signalling) refers to writing devices which emphasise aspects of a text's organisation and/or content. They include outlines, overviews, summaries, all systems of headings, enumeration devices (e.g. numbering points in an argument) as well as words and phrases which serve to draw attention to points and orient readers (e.g. 'it should be noted that', 'thus', 'in summary'). Sign-posting also includes typographical cues such as boldface, italics, underlining, the use of colour and the alignment of text (indentation, centering). (Lorch, 1989)

Sign-posting can play a very powerful role in supporting the comprehension and recall of especially expository text. It does so by assisting readers to develop well-organised mental representations of texts. Spyridakis and Standal (in Van Rooyen 1990:16) explain that sign-posting serves "to emphasise and preview content relationships and help a reader identify superordinate content and make inference from it. In theory, signals should aid a reader in instantiating the appropriate schema, in forming a hierarchical framework in which to store textual information, in deciding what information is important and in checking the correctness of his or her integration and storage of information in memory".

A number of studies (summarised in Lemarié, Lorch & Péry-Woodley, 2012 as well as Lorch, 1989) have verified that sign-posting can aid in the comprehension and recall of texts. While much of the research done in this area has related to L1 readers, a recent study by Kim and Clariana (2017) has confirmed that the use of coherent text signals can improve comprehension of science expository text for both low proficiency and high proficiency ELLs. Kim and

Clariana (2017) emphasise that sign-posting adds the most value when the goal of a text is to convey information to readers who are unfamiliar with the content..

Van Rooyen (1990) recommends that the following questions be asked to determine the usefulness of sign-posting in increasing text readability: 1) Do headings predict the following content? 2) Is the content of a section organized in a logical fashion, i.e. in such a way that propositions follow from the previous ones? 3) Is an overview of the chapter or section of it provided in comprehensible language? 4) Do paragraphs have explicitly stated, salient, main idea formulations? Such questions provide a useful point of departure to begin evaluating the presence and usefulness of sign-posting in LTSMs developed for ELLs.

Appendix B – Texts with unknown words blacked-out (originals)

Topic 1: Living and non-living things

Skills focus: Observe

All living things grow. We can watch the changes in living things as they get bigger. To observe means to use as many of your senses as possible when you study something.

How to observe
When observing something, do the following:

1. Try to use all five of your senses to find out more about the object.
2. If you know that the object is safe, pick it up, feel it, smell it and perhaps taste it. Look at it from all sides and listen to it.
3. Look for any changes in the object you are observing.
4. Sometimes you need to count or measure the things you observe.

Use a ruler and a magnifying glass to help make your observations more accurate.

Topic 17: Rocket systems

Starting off

People have experimented with rocket systems for many years. They have used rockets to make deadly weapons and rocket-driven vehicles. But most importantly, people have used rocket systems to travel into space. Rockets carry people, equipment and other spacecraft into space. People have used rockets to travel to the Moon, and rockets have even been used to get special robotic vehicles called rovers onto Mars. In this topic, you will find out more about how rocket systems work and you will make a rocket system yourself.

Key words

rocket: long tall structure that can be propelled into space

Appendix C – Text organisation and language features of selected genres

GENRE	PURPOSE	TYPES	TEXT ORGANISATION	LANGUAGE FEATURES
Recounts	To tell what happened.	Personal recount, factual recount, imaginative recount.	Focus is on a sequence of events, all of which relate to a particular occasion. Generally begins with an orientation (giving background information – who, where, when), followed by a series of events ordered in chronological sequence. At various stages there may be some personal comment on the incident.	• Specific participants • Use of simple past tense • Use of action verbs [material processes] • Use of linking items to do with time
Instructions [Meta-Genre: Procedure]	To tell someone how to do or make something.	Recipes, science experiment, craft instructions, game rules, appliance manuals, how-to-do-it kits, directions to reach a destination	Focus is on a sequence of actions. Structure generally consists of **Goal** (often indicated in the main heading and/or diagram), **Materials** (listed in order of use) and **Method** (steps oriented towards achieving the goal). The text may also include comments on the usefulness , significance, danger, fun, etc. of the activity. Headings, subheadings, numbers, diagrams, photos, etc. are often utilised to make Instructions as clear and easy to understand as possible.	• Generalised participants referring to a whole class of things • The person following the instructions is referred to in a general way (one/you) or sometimes not at all • Linking words to do with time • Mainly action verbs [material processes] • Tense is timeless, e.g. simple present tense • Detailed factual description of participants • Detailed information on ***how, where*** and ***when***
Narratives [Meta-Genre: Story Genres]	To entertain, i.e. to gain and hold the reader's interest in a story. May also teach, inform, embody writer's reflections on experience or nourish and extend reader's imagination.	Fairy stories, mysteries, science fiction, choose-your-own-adventures, romances, horror stories, "heroes and villains" (e.g. TV cartoons), adventure stories, parables, fables and moral tales, myths and legends, historical narratives.	Usually begins with an orientation, introduces main characters and indicates where and when action is taking place. Often there is foreshadowing of the action to follow, drawing readers into the story. The story is pushed along by a series of events, during which some sort of complication arises temporarily thwarting the main character(s) from reaching their goal. In a "satisfying" Narrative, a resolution of this complication is brought about, though it may be resolved for better or for worse.	• Specific, often individual participants with defined identities. • Mainly action verbs [material processes], but also many verbs which refer to what the human participants said, or felt, or thought [verbal and mental processes] • Normally past tense • Many linking words to do with time • Dialogue often included (may be in present or future tense) • Descriptive language chosen to create images in the reader's mind. • Can be written in the 1st or 3rd person

Information Reports	To document, organise and store factual information on a topic. Information Reports classify and describe phenomena of our world. We use them when we talk about a whole class of things whereas a Description only talks about one specific thing.	The term "report" is used in everyday language to refer to many different types of factual texts. Here, however, the term Information Report is being used to refer *only to texts used to store information about a class of things.* Information reports can contain information of various kinds, e.g. **classification** into different types (subclasses), an examination of **components**, or a look at various **aspects**.	Major focus is on a "thing" (a class of things) rather than a sequence. Topic is usually introduced by an **opening general statement/general classification** locating what is being talked about in the universe of things. Often takes the form of a classification or definition. The rest of the report will consist of **facts about various aspects of the subject**. Each aspect might be elaborated by referring to distinctive characteristics of the subject, or by giving examples , or by comparing and contrasting, or by describing components and their function. Reports don't usually contain an "ending". Diagrams, photos and illustrations are often used to lend clarity to the text and may be accompanied by labels or captions.	• Generalised participants: a whole class of things rather than specific participants • Some action verbs [material processes] • Many "linking verbs" [relational processes] (is, are, has, have, belongs ton) • Usually in "timeless" present tense • Descriptive language, but factual and precise rather than imaginative or "lively" • Likely to contain technical vocabulary • Writing style is relatively formal and objective; the writer's opinions are not generally appropriate in this type of writing
Explanations	To give an account of how something works or reasons for some phenomenon.	A) explaining ***how***, e.g. *mechanical* explanation (How does a pump work?), *technological* explanation (How does a computer work), *system* explanation (How does a company work?), *natural* explanation (How are mountains formed?) B) explaining ***why***, e.g. Why do some things float and others sink? Why is the ozone layer thinning? etc.	Explanations have a "process" focus rather than a "thing" focus. They are often concerned with a logical sequences. To position the reader, there is usually some **statement about the phenomenon in question** (often in the form of a heading or question), followed by a **(sequenced) explanation of how/why something occurs.**	• Generalised non-human participants • Time relationships (first, then, following, finally), especially in how-explanations • Cause-and-effect relationships (if/then, so, as a consequence, since), especially in why-explanations • Mainly action verbs [material processes] • Some passives • Timeless present tense
Arguments [Meta-Genre: Exposition]	To take a position on some issue and justify it.	**Argument** texts belong to a genre group called "Exposition", concerned with the analysis, interpretation and evaluation of the world around us. There are several different expository genres, depending on whether your major aim is to, interpret, or evaluate. Some texts contain a mix of expository genres. In contrast to Argument texts, essays which require the writer to discuss both sides of a case and then form some opinion based on the preceding arguments are referred to as **Discussion** texts.	Major focus is on an issue and a logical sequence of argument related to this issue. The beginning of an argument usually consists of a **statement of position [thesis statement]** often accompanied by some background information about the issue in question. There may also be some broad foreshadowing [preview] of the line of argument to follow. To justify the position, the writer must now present the **argument**. Usually there is more than one point put forward in the argument, and each one should be supported by evidence, and possibly by examples. All points should relate directly back to the statement of position. Finally, there is an	• Generalised participants – sometimes human but often abstract (issues, ideas, opinions, etc.) • Possibility of technical terms relating to the issue. • Variety of verb [process] types – action [material], linking [relational], saying [verbal], and mental. • Mainly timeless present tense but may change when providing background or making predictions • Frequent use of passives • Actions often changed into "things" (nominalised) to make the argument sound more objective

			attempt at **summing up the position** in the light of the argument presented.	• Connectives associated with reasoning (therefore, so, because of, the first reason, etc.) • Use of emotive words more appropriate to spoken than written arguments

(Derewinaka, 1990)

Appendix D – Sampled texts used in syntax analysis

TEXT 1

EFAL LB | Term 1 | pp. 2-4 | Literary | 209 tokens

The big match. Last Saturday, our school team played a soccer match against Greenfields Primary School. We were all very nervous and excited as we put on our soccer clothes.

FINAL SOCCER MATCH: Greenfields vs Parkside. Venue: Parkside Primary School Sportsground. Date: 15 September. Time: 2 pm.

As the teams ran onto the field, everyone shouted and cheered. Then the match began. Our team scored a goal in the first five minutes. The crowd jumped, cheered and clapped.

Just before the end of the game, Greenfields got a penalty kick. The ball flew through the air towards the goal. But Thomas, our goalkeeper, jumped very high and saved the ball. The whistle blew and it was the end of the match. We shook hands with the Greenfields players.

The principal of our school gave us a trophy. The captain of our team made a speech. All our teachers and families were really proud of us. Someone came to take a photograph of us.

There was a picture in the newspaper the next day, with a short story about the match. It said that our school had the better team. ‘Thabsile Lehoko wins the game for Parkside Primary School.’ The school celebrated our win with songs at the school assembly.

TEXT 2

EFAL Reader | Term 1 | pp. 5-7 | Literary | 154 tokens

A special school day. Yesterday, I had an interesting day at school. The nurse from the clinic came. She showed us how to brush our teeth. She said, “Brush your teeth in the morning after breakfast. Brush your teeth in the evening before you go to bed.” She gave us each a free toothbrush and toothpaste. ‘Your teeth are precious!’

It was fun when the whole class pretended to brush their teeth. Then the mayor came to open our new school library. All our parents had helped to raise money for library books. First, the mayor gave a speech. Then our principal spoke.

After the speeches, we went to the library. We were very excited when we saw the beautiful covers of all the books. There were five computers in one corner of the library. There was a television and a DVD player in another corner. I can't wait to use our wonderful library!

TEXT 3

EFAL LB | Term 2 | pp.46-47 | Informational | 186 tokens

Around the world for a cup of tea. Did you know that the things we need to make a cup of tea come from all over the world?

Tea. This is a tea plantation in India. These women pick the best leaves from the top of the tea plant. In the tea factory, workers dry the tea leaves. Then they pack the tea in boxes and send it all over the world.

Sugar. These men work on a sugar cane farm in KwaZulu-Natal, South Africa. In the sugar factory, workers use machines to squeeze out the sugar cane juice. Then they boil the juice in big drums to make the sugar crystals.

Milk. This is a dairy farm in Gauteng, South Africa. The farmer milks his cows every day. Machines fill plastic bottles with milk. The dairy farmer sends the milk to supermarkets and shops.

Cups and saucers. Men and women make hundreds of cups and saucers every day in this factory in Sri Lanka. FRAGILE. This box is full of cups and saucers. Workers pack the boxes carefully and send them all over the world.

TEXT 4

EFAL Reader | Term 2 | pp. 15-18 | Informational | 165 tokens

All kinds of houses. Long ago, people lived in caves. They used stones to stop the wind from coming into the cave.

Later, some people lived in tents. The tents were made of wooden poles and animal skins.

People also started to build houses of mud. These houses were safer than tents. They made the roofs from branches covered with straw and mud.

In Egypt, people used mud bricks and wooden poles to build houses. They used stone to make the doorways. They painted the walls in bright colours.

In other places, people used wooden frames for the roofs. They covered them with straw.

Some people built their houses of large stones. They made roofs using wooden frames and straw.

Other people lived near big forests so they built their houses out of logs. They decorated their homes with beautiful, patterned cloths.

Some people even built their houses out of blocks of snow. These snow houses were called igloos. People still live in igloos today.

TEXT 5

EFAL LB | Term 3 | pp. 62-64 | Literary | 443 tokens

Maisie gets toothache. Maisie loved to eat sweet things. She loved chocolates. She loved ice-cream. She loved biscuits. She loved cake. She also loved to drink sweet things. She loved fizzy drinks. She loved sweet juices. She loved chocolate milk. When she drank tea, she put in six great big spoons full of sugar. “I only like sweet stuff,” said Maisie.

Her mother gave her vegetables, but Maisie gave them to the dog. Her mother gave her beans and fish, but Maisie gave them to the cat. Her mother gave her apples, but Maisie didn't eat them.

One day Maisie woke up with a very sore tooth. She could not eat anything. “Ow! Ow!” she cried. “My tooth is sore. Ow! Ow! Help me.” She cried and cried. “You have got toothache,” said her mother.

Maisie’s mother took her to the dentist. They sat in the waiting room. Some children were waiting with sore teeth. They were crying. Some children were waiting for a check-up. They were reading magazines and playing with blocks.

There were lots of posters on the walls. BRUSH YOUR TEETH TWICE A DAY. Brush up, brush down, brush in front and behind. WATCH OUT! FIZZY DRINKS, SUGAR, SWEETS – TOO MANY OF THESE GIVE YOU TOOTHACHE. HEALTHY MEALS MEAN HEALTHY TEETH.

At last it was Maisie’s turn to see the dentist. He looked inside her mouth. “This tooth must come out,” said the dentist. “You are lucky it is a baby tooth. You will grow another one in its place.” “Ow! Ow!” cried Maisie. “It will be sore when you take it out.”

The dentist washed his hands. “Sweet things hurt you,” said the dentist. “They leave dirt on your teeth. Your teeth get holes and go bad. You get toothache. I will give you an injection so that it doesn’t feel too sore when I pull out your tooth.”

The dentist put a toy rabbit in front of Maisie. “Look at the rabbit. Think about its big strong teeth,” said the dentist. While Maisie was looking at the rabbit, the dentist gave her an injection. Then he took out the bad tooth.

“There,” he said, “now you will feel better. But you must look after your teeth now. Eat healthy food. And brush your teeth very well. Stay away from sweets. Do you want a present?” “Yes, please,” said Maisie. “Here is a toothbrush,” said the dentist. “Do you want a blue one or a

yellow one?" "A yellow one, please," said Maisie. "Thank you. I will use it every day." "Good," said the dentist. "I will see you again in six months. I want to see healthy teeth."

TEXT 6

EFAL Reader | Term 3 | pp. 48-51 | Literary | 245 tokens

The elephant and the rain. Elephant and Rain always argued about who was stronger. "I am the strongest." "No, I am."

Rain said, "If I go away, you will all die." Elephant said, "I don't care. Go if you like. I will rule alone!"

So one day, Rain went away. At first, nothing changed. There was grass to eat and the rivers were full of water.

But after a while there was only one water hole left. The animals were thirsty.

Elephant asked Vulture to help. But Vulture did not help. "No!" Elephant asked Crow to make rain. But Crow could only make a few drops.

There was still no water. Elephant went to look for food. He told Tortoise to look after the water hole till he came back. "Go away! This is Elephant's water."

Then lion came. He did not listen to Tortoise. "I am the boss," he said. "Get out of my way!" "Get out of my way!"

Then all the other animals came too. They came down the dry hills and over the dry grass. They drank up all the water.

When Elephant came back he was angry. He picked Tortoise up with his trunk and swallowed him.

Tortoise did not want to be inside Elephant. He pulled and pushed. He started to bite his way out. "Ow! Ow!"

Rain felt sorry for the animals, so Rain came back. Water fell from the sky. The earth was green and wet again.

TEXT 7

EFAL LB | Term 4 | pp. 92-94 | Literary | 262 tokens

How Giraffe got a long neck. Long ago, Giraffe did not have a long neck. He had a short neck.

One day, Giraffe walked past Nogwaja, the hare. “Oh, Giraffe! I have hurt myself. I can’t get food. Please help me,” said Nogwaja.

Giraffe felt sorry for Nogwaja. “I have got some food,” he said. “But you must pay me back when you are better.” “Of course I will pay you back,” said Nogwaja.

Giraffe gave Nogwaja lots of good things to eat. But Nogwaja was lazy. He did not want to pay Giraffe back. He had no food for Giraffe. He played a trick on him instead.

First, Nogwaja took a rope. He swam across the river. He tied one end of the rope to a tree. Then he swam back with the other end of the rope. He went to Giraffe. He told him a lie. “I have food for you,” he said. “But it is on the other side of the river. Please pull it across the river with this rope.”

Nogwaja tied the rope around Giraffe’s neck. “Pull!” he said. Giraffe pulled and pulled. Nothing happened. He pulled and pulled again. He pulled so hard that his neck got longer and longer. He pulled so hard that the rope jerked and broke. Giraffe fell on top of Nogwaja. He hurt Nogwaja’s paws.

Nogwaja was very unhappy. But Giraffe was very happy. He liked his long neck! He could eat all the nice leaves at the top of the trees. He could also see dangerous animals from far away.

TEXT 8

EFAL Reader | Term 4 | pp. 60-61 | Informational | 180 tokens

Living things have lots of water in them. This glass is two-thirds full.

Did you know? A tree is two-thirds water. Your brain is two-thirds water. We need to drink enough clean water.

Did you know? You can live for weeks without food, but only a few days without water. Pipes bring water to taps and toilets.

Did you know? Fewer people in the world have toilets than cell phones. Clean water helps us to stay healthy.

Did you know? If we all washed our hands properly, millions of the world's sick people would be healthy.

We do experiments to find out how things work, or to see what happens when we mix things together. We also do experiments to test that something is true. This experiment helps us to see how water moves through plants.

Do an experiment. How does water move through plants? You will need: a glass of water, food colouring, a piece of celery.

Mix the food colouring with water in a glass. Put the celery in the glass overnight. What has happened to the celery?

TEXT 9

NST | Term 1 (1st) | pp. 2-6 | Informational | 515 tokens

Unit 1: Living things. There are many different kinds of living things. Some living things are big, such as an elephant. Other living things are so small that we cannot see them with the naked eye such as germs. Look at the pictures of some living things.

Some living things: elephant, fish, fern, sunflower, mealie plant, bacteria/germs, bread mould, locust.

Life processes of living things. All living things carry out life processes. Life processes are the things plants and animals do that show us they are alive. There are seven life processes.

Growing. They grow up and become adults. Plants grow. People grow. Reproducing. They make babies. Feeding. They need food. Moving. They can move. Breathing. They need to take in air. Excreting. They get rid of waste products. Sensing. They can see, hear, feel, smell or taste their environment.

Things that seem not living can be alive. Sometimes things seem to be not living or dead, but they are really living. A seed seems to be not living, but when we plant it and give it water and warmth, it starts to grow. When a seed starts to grow we say it germinates. The germinating seed grows into a plant. A seed germinates and grows into a plant.

Yeast is a very small organism. It is used to make bread. Yeast looks like it is not living but when we give yeast food and warmth it grows. We use yeast to make bread.

An egg seems to be not living, but there can be a young chick growing inside the egg. When the chick is big enough, it breaks the shell with its beak. We say the chick hatches. A young chick grows inside an egg.

What makes things that seem not living come alive? When you give the seeds water, they start to grow. You also put the seeds in a warm place. Seeds need warmth and water to start growing and come alive. The yeast starts to grow when you give it sugar for food. It also needs warmth and water in order to grow. The chick needs warmth, food and air to grow big enough to hatch from the egg. The mother bird sits on the egg to keep the chick warm. The young chick gets food from the egg yolk. It gets air through the egg shell.

Skills focus: Observe. All living things grow. We can watch the changes in living things as they get bigger. To observe means to use as many of your senses as possible when you study something.

How to observe. When observing something, do the following: Try to use all five of your senses to find out more about the object. If you know that the object is safe, pick it up, feel it, smell it and perhaps taste it. Look at it from all sides and listen to it. Look for any changes in the object you are observing. Sometimes you need to count or measure the things you observe. Use a ruler and a magnifying glass to help make your observations more accurate.

TEXT 10

NST | Term 1 (2nd) | pp. 49-50 | Informational | 330 tokens

Case Study: Hoot and Toot need a new home. Hoot and Toot are two Spotted Eagle Owls that live at the National Zoological Gardens of South Africa in Pretoria. They go out into the zoo every day with their handlers so that visitors can see them up close. Their enclosure is in a quiet part of the zoo where visitors aren't allowed.

Robynn Ingle trains Hoot and Toot. She knows a lot about owls' habits and how they live in the wild. She knew that the owls needed a bigger enclosure with some space to fly. They needed choices about where to sit. The enclosure needed to protect them and give them privacy, as well as allowing them to look out. They needed a place to hide away and rest.

Robynn knows that all animals kept by humans need: fresh, clean water and proper food; a comfortable place to live; to be safe and healthy; to feel safe and happy; to be able to do the kinds of things they would do in the wild.

Robynn helped the Friends of the zoo to design and build a new enclosure for Hoot and Toot to meet all their needs. Hoot in her new enclosure. Robynn Ingle with Hoot.

Hoot and Toot also needed new nesting boxes to hide in and rest. Some students from the University of Pretoria designed and made the nesting boxes. First they did research on owls' habits.

They discovered that Spotted Eagle Owls like to roost on the ground. The students thought about the size and shape of the boxes and what materials they could use. They learnt that owls must be able to turn around in their nesting boxes and their heads must not touch the top. The

nests needed to protect the birds from sun, wind and cold. The nests could not cost too much money to make but they needed to be safe, comfortable and easy to clean. Hoot in her new nesting box.

TEXT 11

NST | Term 2 (1st) | pp. 58-60 | Informational | 382 tokens

Unit 2: Change of state. As you learnt earlier, the three states of matter are solid, liquid and gas. When a material changes from a liquid to a solid, from a solid to a liquid, from a liquid to a gas, or from a gas to a liquid, we say that the material has a change of state.

Melting – changing from solid to liquid. If solids are heated, they can change into liquids. For example, if we add heat energy to butter in a frying pan, it melts and becomes liquid. The photograph of the boy eating an ice-cream shows another example of a solid that is melting.

Evaporating – changing from liquid to gas. When liquids are heated enough they can change into gases. When this happens we say that the liquid evaporates. Liquids evaporate when they gain heat. For example, when you heat liquid water it changes to a gas called water vapour.

Condensing – changing from gas to liquid. When a gas cools, it loses heat and becomes a liquid again. When a gas changes back into a liquid we say that it condenses. Gases condense when they lose heat. For example, when you breathe out air on a cold day the water vapour in your breath cools down and condenses to form liquid water.

Ice-cream melts and becomes liquid quickly on a hot day. The water in the wet washing evaporates into the air on warm days. When liquid water is heated enough it changes to a gas. When water vapour loses heat it condenses to form drops of water.

Solidifying – changing from liquid to solid. Liquids lose heat when they are cooled and they can then change into solids. For example, when you put water into the freezer, it becomes a solid. When the water has been in the freezer for a few hours it becomes ice. The water changes from a liquid into a solid when we cool it. We say that the water solidifies. However, not all substances have to be to be put in the fridge or freezer to solidify. For example, butter and

margarine are solids at room temperature, as you can see in the photograph on the left. Some materials are solids at room temperature. They melt and become liquid when they are heated.

TEXT 12

NST | Term 2 (2nd) | pp. 99-103 | Informational | 354 tokens

Examples of indigenous houses of Southern Africa. Rondavels: The Xhosa people build huts called rondavels. The Xhosa people also call these huts rontabile and ungqu-phantsi. The huts have round walls and thatched roofs. The rondavels are built from raw materials found in the area. The Xhosa people build their huts by sticking poles or pillars into the ground. Planks or laths are attached to the poles, like struts in a framework. Rocks are inserted between the laths. The whole structure is covered with a mixture of clay and cattle dung. The walls are decorated with paint.

The roofs are constructed of wooden poles cut to length. The roof is covered with thatch sewn to the poles with grass ropes. It is sewn from the bottom working towards the top. Such a roof becomes waterproof. Examples of a Xhosa rondavel. Structure of roof from inside. Look at the triangles made by struts.

Beehive-shaped huts: Traditionally, Zulu people lived in beehive-shaped huts. The isiZulu word for such a hut is uguqa. The men in the community construct the huts by making a framework of flexible young trees in the form of an arch. They tie the branches together to form triangles. The women weave ropes and mats from grass. The mats are laid over the wooden frame of the hut, from bottom to top. They tie the mats to the framework with thick grass ropes. Example of an uguqa.

Matjieshuis: Traditionally, the Nama people of Namibia live in structures called a matjieshuis. The men collect branches, mainly from tamarisk trees. They remove the thorns and hold the branches over a very hot fire. They then remove the bark and bend the branches. When the branches cool down they keep their bent shape. The bent branches are tied together to form a dome or arch shape. More branches are used as struts to strengthen the framework. The women harvest reeds. They dry the reeds in the Sun, cut them to the correct length and weave them into mats. Finally, they drape and fasten the mats over the framework. Framework of a matjieshuis. Completed matjieshuis. A Swazi hut.

TEXT 13

NST | Term 3 (1st) | pp. 118-121 | Informational | 543 tokens

Unit 1: Energy can be stored and transferred. Energy around us. Energy is all around us. For example, during the day the Sun brings us energy in the form of light and heat. We can see things happening and changing.

Even at night, when it is dark, we are aware of the energy around us. We hear people talking and making music. Sound is a type of energy. We feel the wind on our faces. Wind is moving air; it has movement energy.

Some of the energy around us comes from objects that we use. For example, lamps, stoves and radios.

Sources of stored energy. In Topic 10, you learnt that a source of energy is something that we get energy from. Food is a source of energy for our bodies. The energy that is stored in food comes from plants. Plants use energy from the Sun to make food.

People also use energy that was stored in the bodies of plants or animals long ago. Sources of this energy include wood, coal, crude oil and natural gas. Coal was made from plants that lived millions of years ago. Coal is a source of energy for most of the mains electricity in South Africa.

Crude oil is found deep underground. It was made from the bodies of very small plants and animals that lived in the sea millions of years ago. Petrol, diesel, paraffin, jet fuel and candle wax are all products made from crude oil. We use petrol, paraffin and liquid petroleum gas (LPG) as sources of energy. They are made from a type of oil called crude oil.

Energy can be transferred. Energy can go from a source of energy to where it is needed. This is called energy transfer. During this process one type of energy may change to another type.

For example, when we burn wood to make fire the stored energy in the wood is changed to light and heat, as show in the picture below:

Burning wood gives light and heat. Light energy from the burning wood helps us to see when it is dark. Heat energy from the fire goes into the air to make us warm. Heat energy from the burning wood goes to the pot. The heat cooks the food in the pot.

Skills focus: Draw a flow diagram for a process. A flow diagram helps us to understand the order in which something happens. A flow diagram shows you the steps in a process in the correct order.

The energy chains that you drew in Topic 10 are examples of flow diagrams. A flow diagram can also be used to show how energy is transferred from a source to where it is needed.

How to draw a flow diagram. Write the steps that show the order in which something happens. You can write them down the page or across the page. Draw a shape around each step; it can be a circle, a square or a rectangle. This makes each step easy to see. Draw arrows between each step. Look at the flow diagram below. It shows how an electric light works. Turn the light switch on. Electricity goes along wires to the light bulb. The light bulb changes the electricity to light.

TEXT 14

NST | Term 3 (2nd) | pp. 148-149 | Informational | 326 tokens

Unit 3: Noise pollution. Noise is any unwanted sound. If loud, unpleasant or harmful sounds carry on for a long time, we call it noise pollution. Most noise pollution comes from machines, tools and vehicles made by humans.

Volume (how loud or soft a sound is) is measured in decibels (dB). If you are near sounds that are louder than 85 dB a lot of the time, your hearing will slowly get worse. If you hear a noise that is louder than 130 dB your hearing can be damaged right away. A hearing loss caused by noise will not get better. We say that someone who has a hearing loss is hearing-impaired.

Some machines, tools and vehicles make loud noises: a jackhammer, aeroplane, truck, lawn mower.

Case study. Losing your hearing. Xolile Nkutha worked as a miner in a gold mine for 24 years. The mine gave the miners ear muffs to protect their ears. The muffs were hot and uncomfortable to wear and the miners could not hear each other when they wore them. Xolile and the other miners did not know that being in a noisy mine eight hours every day would damage their hearing. Later, Xolile could not hear people properly when they talked to him. He had his hearing tested by an audiologist. He had to wait a long time to get a hearing aid from the hospital. The hearing aid made the sounds that entered his ears louder so that he could hear what people were saying again. Miners use equipment that makes very loud noises.

Description: dB levels – noise source. Damages hearing immediately: 140 – shotgun blast; 130 – plane taking off; 120 – Vuvuzela. Extremely loud: 110 – car hooter; 100 – underground in a mine; 90 – big truck. Damages hearing over time: 80 – factory noise. Loud: 70 – classroom noise; 60 – people talking in the street. Comfortable: 50 – people talking at home; 40 – fridge humming. Soft: 30 – whisper; 20 – watch ticking; 10 – normal breathing.

TEXT 15

NST | Term 4 (1st) | pp. 160-161 | Informational | 360 tokens

Unit 2: Earth and space. Earth is a planet in space. Earth is a planet. A planet is a large object that moves around a star. Our Sun is a star. Earth is a planet of the Sun. So Earth always moves around the Sun. The Sun has eight planets that move around it and Earth is one of them. Earth is one of eight planets of the Sun: Sun, Earth.

The Sun and its eight planets are in space. Space is a never-ending darkness that is filled with stars and other planets. In space there are many other objects, such as gas and dust but no air. Astronauts are people who travel to space in space shuttles. A space shuttle is launched with a rocket.

From Earth we can see the Sun, Moon and stars. The Sun is a bright star that is near enough to Earth for us to see it. We can only see the Sun in the daytime when our half of the Earth is facing towards it. We can use a slowly turning ball as a model to represent the turning Earth. We can use the light from a torch to represent the light of the Sun. If we shine the torch onto the ball, only half of the ball will be lit up. As the ball turns slowly, a different part of it will

face the Sun. Only half of Earth can see the Sun at any one time: light from torch; side facing the light; side turned away from the light.

We see the Moon because it is the nearest object to Earth in space. The Moon is always close to us. The Moon is cold and dark but gets its light from the Sun. So when we see a bright shiny Moon, we are seeing the light of the Sun reflected from the Moon.

There are millions of stars in space. We can only see some of them as small lights because they are so far away. Stars shine all the time but we cannot see them during the day because the Sun is too bright. The moon is close to Earth.

TEXT16

NST | Term 4 (2nd) | pp. 184-185 | Informational | 379 tokens

Unit 3: Moon stories. People through the ages have looked at the Moon, and wondered about it. Some people use the phases of the Moon to tell them when to start planting their crops. Different phases of the Moon also mark important days in many religions. In many cultures there are stories to explain things about the Moon.

The San belief about the changes in the Moon's shape. The San people thought of the Moon as a person who had angered the Sun. They believed that as punishment, each time the Moon appeared in the sky as a crescent shape, it had to beg the Sun for mercy to continue to live. After begging, the Sun would grant the Moon's wish of growing to its full size. Once the Moon was whole again, the Sun's rays would begin to slice pieces off the Moon again. A Moon sliver begs the Sun to continue to live.

The Tswana belief of shapes on the Moon's surface. The Tswana people believed that the surface of the Moon formed shapes. Some people saw the shape of a person with wood strapped to their back. Others saw a mother with her child tied to her back. Shapes formed on the Moon.

Early Xhosas and the new Moon. The early Xhosa people believed that the ocean stores many new Moons in a huge crater or hole under the ocean. As one Moon drops and disappears under the ocean, so a new Moon rises slowly into the sky. One Moon drops into the ocean and a new Moon rises.

The Sotho, Venda and Tswana people's belief about the pointy ends of the Moon. Each month, Ancient Sotho, Venda and Tswana people saw the Moon's crescent shape as a bowl turned up or down. The shape depended on which way the pointy ends faced, when the Moon appears as just a thin crescent in the night sky. The people believed that the bowl was filled with illnesses. When the Moon's pointy ends faced down, the illnesses fell from the bowl, making everyone ill. On these nights, no one went out. When the pointy ends faced up, the illnesses stayed in the bowl and everyone could go out without fear. The crescent shaped Moon's pointy ends face downwards.