BENCHMARKS REPORT

Sesotho-Setswana Early Grade Reading

OCTOBER 2022

basic education
Department:
Basic Education
REPUBLIC OF SOUTH AFRICA

ZENEX FOUNDATION

Funda Wande
Reading for Meaning

Room to Read

SALDRU
Southern Africa Labour and Development Research Unit

University of Cape Town · iYunivesithi yaseKapa · Universiteit van Kaapstad

RESEP
Research on Socio-Economic Policy

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

UNIVERSITY OF LIMPOPO

UNISA
university of south africa

UFS·UV
UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT

BENCHMARKS REPORT

Sesotho-Setswana Early Grade Reading

October 2022

Nompumelelo Mohohlwane (Department of Basic Education)

Prof Cally Ardington (SALDRU, University of Cape Town)

Dr Gabrielle Wills (Research on Socio-Economic Policy, Stellenbosch University)

Lesang Sebaeng (Department of Basic Education)

Zamangwe Zwane (Department of Basic Education)

Dr Eileen Pooe (North West University & Setswana National Language Body - PanSALB)

Dr Refilwe Ramagoshi (University of Pretoria)

Dr Connie Makgabo (University of Pretoria)

Dr Ablonia Maledu (University of Limpopo)

Dr Mamosebatho Julia Ramabenyane (University of Free State)

Prof Elizabeth Pretorius (UNISA)

Christine Beggs (Room to Read)

Pinaki Jodar (Room to Read)

Acknowledgements

This report was funded by the Zenex Foundation and Funda Wande. Additionally, contributions by Nompumelelo Mohohlwane and Lesang Sebaeng and have been funded by the Department of Basic Education. Contributions by Zamangwe Zwane were funded by the World Bank.

This report synthesizes the findings from two distinct reports, the Setswana language benchmarks report and the Sepedi language benchmarks report, and includes new analyses of Sesotho language data.

Contributors to the Setswana Reading Benchmarks report include the authors Gabrielle Wills, Cally Ardington, Elizabeth Pretorius, Eileen Pooe, Refilwe Ramagoshi and Lesang Sebaeng. Funding for the Early Grade Reading Study that the Setswana benchmarks are based on were made by the Department of Basic Education, the Department of Planning, Monitoring and Evaluation, the North-West Department of Education, the International Initiative for Impact Evaluation, Zenex Foundation, the United Nations Children's Fund (UNICEF), the United States Agency for International Development (USAID), and the Chairman's Fund of Anglo American.

Contributors to the Sepedi Reading Benchmarks report include the authors Cally Ardington, Nompumelelo Mohohlwane, Lesang Sebaeng, Ablonia Maledu, Connie Makgabo, Christine Beggs, Pinaki Joddar and Zamangwe Zwane. Funda Wande and Zenex Foundation funded an extension to the planned data collection for the impact evaluation of the Funda Wande programme in Limpopo to allow for the establishment of Sepedi benchmarks. In addition, existing data from Room to Read monitoring and evaluation studies were incorporated into the Sepedi benchmarking exercise. Contributions by Pinaki Joddar and Christine Beggs were funded by Room to Read. Contributions by Nompumelelo Mohohlwane and Lesang Sebaeng and have been funded by the Department of Basic Education. Contributions by Zamangwe Zwane were funded by the World Bank.

Data for the Sesotho language is drawn from Save the Children's Literacy Boost study in the Free State.

DG Forward

The mastery of reading in the early grades continues to be a sector priority underpinning the foundational skills required for a successful learning trajectory for South African young learners and South Africa at large. These early skills are the bedrock upon which all later skills are developed.

The Sesotho-Setswana benchmarks in this report bring together African linguistic relationships across the three languages within this language group: Sesotho, Setswana, and Sepedi. It also merges theory and leadership from African language scholars and academics, education experts, as well as quantitative researchers.

These efforts are intended to empower and equip teachers, parents, universities, and the sector more broadly with educationally meaningful and scientifically valuable approaches to support the teaching of reading in Grade 1 through to Grade 3. It is our hope that along with all the other complementary efforts, these benchmarks will contribute to the improvement of reading in the country.

While this work was led by the Department of Basic Education, it was only possible through broad stakeholder collaboration. The data used for the Sesotho-Setswana benchmarks was primarily based on the Early Grade Reading Study North West studies funded by the Department of Basic Education, the Department of Planning, Monitoring and Evaluation, the North West Provincial Education Department, the Department of Planning, Monitoring and Evaluation, the Initiative for Impact Evaluation, the Zenex Foundation, the United Nations Children's Fund (UNICEF), the United States Agency for International Development (USAID), and the Anglo American Chairman's Fund. In addition, data was contributed by Room to Read, Funda Wande and the Allan Gray Orbis Foundation Endowment and Save the Children.

While these benchmarks are valuable in their own right, their true value will be found as they are used in classrooms, homes, and universities. I encourage all stakeholders to leverage this public good.

MR HM MWELI

DIRECTOR-GENERAL

DATE: 12/10/2020

Contents

Abbreviations and Acronyms

CAPS	Curriculum and Assessment Policy Statement
clspm	correct letter sounds per minute
CCV	Consonant Consonant Vowel
CV	Consonant Vowel
CVV	Consonant Vowel Vowel
cwpm	correct words per minute
DBE	Department of Basic Education
EGRA	Early Grade Reading Assessment
EGRS	Early Grade Reading Study
LOLT	language of learning and teaching
ORF	oral reading fluency
PanSALB	Pan South African Language Board
PIRLS	Progress in International Reading and Literacy Study
RSP	Reading Support Project
SAECMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
TVET	Technical and Vocational Education and Training
USAID	United States Agency for International Development

Executive summary

Background

> "Early reading is the basic foundation that determines a child's educational progress through school, through higher education and into the workplace. All other interventions – from the work being done to improve the quality of basic education to the provision of free higher education for the poor, from our investment in TVET colleges to the expansion of workplace learning – will not produce the results we need unless we first ensure that children can read" (Ramaphosa, 2019)

In recognition of low learning outcomes in reading, the Department of Basic Education (DBE) has increasingly invested in supporting early grade reading through research such as the Early Grade Reading Studies; supplementary curriculum guidance including the Framework for Reading in African Languages and more recently, the creation of early grade reading benchmarks.

Under the leadership of the DBE, collaborative efforts involving African language linguists, donors, literacy organisations, and quantitative researchers have made considerable progress towards establishing reading benchmarks in all South African languages. Firstly, a consultative process resulted in a design report which detailed the technical and linguistic approach, along with the grades and skills to benchmark. The report determined that the focus should be on letters and passage reading fluency benchmarks for Grades 1 to 3.

From the onset, there was consensus that each African language needed to be benchmarked individually taking into account differences in phonology (system of speech sounds), orthography (writing) and morphology (words and their constituent parts) across languages. In other words, the adoption of existing international English benchmarks was not appropriate and benchmarks for one language, such as isiZulu, would not automatically be adopted for another, such as Sepedi. However, once data analysis was completed for each language within a language family (i.e. Nguni languages and Sesotho-Setswana languages), there was agreement to adopt one benchmark for the language family if the individual language benchmarks proved to be similar.

So far, efforts have resulted in the development of early grade reading benchmarks for the Nguni language group (isiZulu, isiXhosa, Siswati and IsiNdebele) in 2020 based on the availability of large-scale data from early grade reading studies across the country. As a new development, this report provides the benchmarks for the Sesotho-Setswana language group. It is based on the stand-alone Setswana and Sepedi reading benchmarks reports. This report synthesises the analyses in these reports along with the incorporation of available Sesotho data to provide a consolidated report for the Sesotho-Setswana language family.

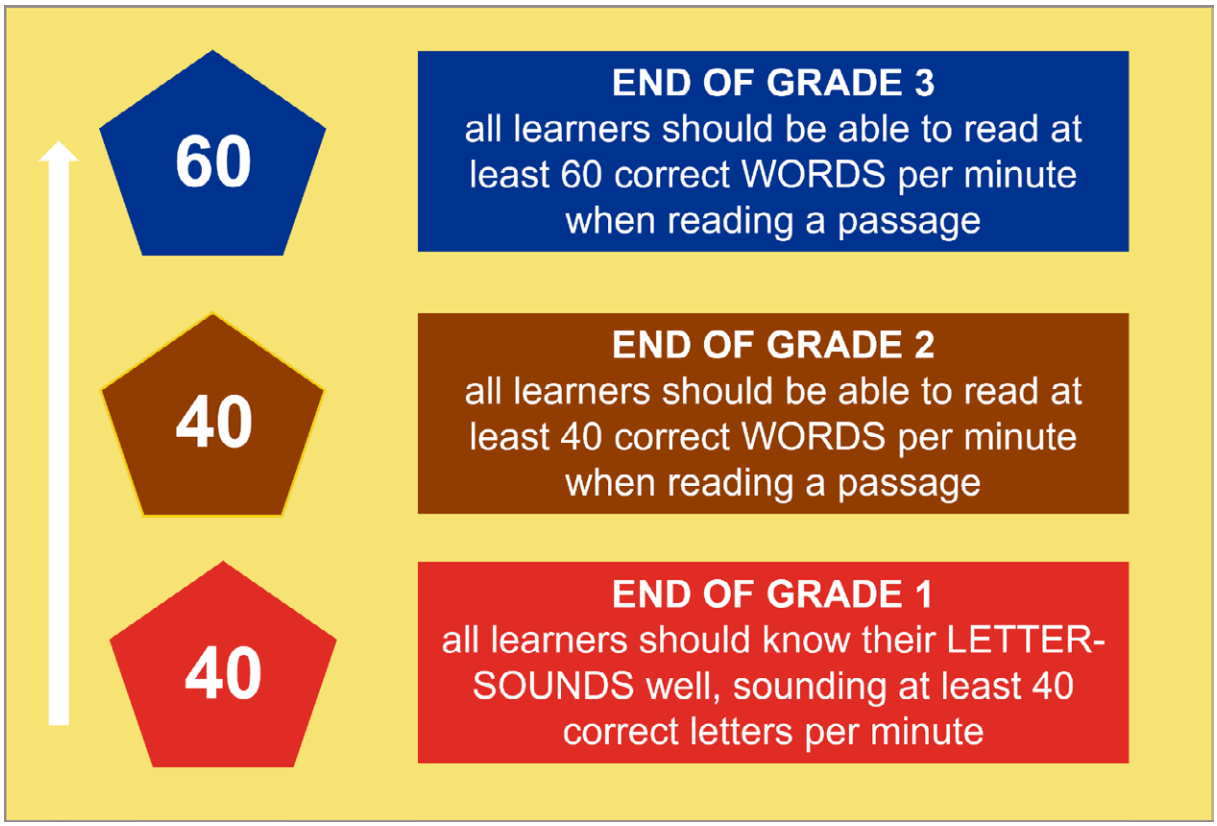
Why Do We Need Benchmarks?

International assessments such as the Progress in International Reading Literacy Study (PIRLS) show that the majority, 78 percent, of learners in South Africa cannot read for meaning by Grade 4. While this implies that mastery of *early* reading skills is not taking place in the Foundation Phase, the PIRLS data provide no guidance on where these foundational gaps lie. Reading benchmarks in the early grades afford the sector a standard by which to measure its learners and to monitor progress towards targets such as having all ten-year-olds reading for meaning by 2030 (South African Government, 2019). Closer to the ground, benchmarks inform the teacher about which learners are on track to become proficient readers and which have gaps in foundational skills. Teachers can therefore implement informed and adequate intervention strategies *early* in the reading journey.

How Do We Establish Benchmarks?

From the onset, the approach used to establish benchmarks was multidisciplinary. Three integrated features informed the benchmarks set: reading development theory, linguistic expertise of each of the Sesotho-Setswana languages (Sesotho, Setswana and Sepedi) and quantitative analysis of large-scale data. These three features were balanced together against the demands of the Home Language Foundation Phase curriculum. For this study, we used existing early grade reading assessment data and undertook new data collection activities. The existing data early grade reading assessment data was drawn from the first four rounds of DBE's first Early Grade Reading Study (EGRS I) in North West province and monitoring and assessment data collected by Room to Read in Limpopo, and Save the Children in the Free State. In addition, we use data from the fifth round of EGRS I and the Funda Wande evaluation in Limpopo that were purposively designed to incorporate benchmark-specific requirements.

What Are The Sesotho-Setswana Early Grade Benchmarks?



As illustrated in the figure above, the benchmarks are as follows:

- By the end of Grade 1, *all* learners should be able to correctly sound 40 letters per minute.
    - This is the same benchmark as for the Nguni languages. While pronunciation may be different, the letters across languages are the same supporting the use of one benchmark.
    - Letters are a good early predictor of oral reading fluency (ORF) levels acquired by the end of the Foundation Phase. Improvements in letter-sound speed stagnate at 40 letters.
    - Once learners have achieved this level of letter-sound knowledge, phonics instruction should focus on blending of sounds and complex consonants while decoding instruction should focus on helping learners apply word attack strategies.

- By the end of Grade 2, *all* learners should be able to correctly read at least 40 words per minute.
    - Below this threshold, accuracy is poor and we find little evidence that learners can understand what they have read. Quite simply, they are making too many mistakes and reading too slowly to comprehend what they are reading. For learners not meeting the Grade 2 benchmark, instruction should focus on improving decoding skills.
    - Once learners have reached this level, they would benefit from instruction that focuses on developing fluency and exposes them to a wider range of texts.

- By the end of Grade 3, *all* learners should be able to correctly read at least 40 words per minute.
    - At this level of fluency, reading comprehension becomes increasingly possible when learners read on their own. Once learners reach this level of fluency, it appears that poor comprehension skills become the limiting factor to further literacy development.
    - Once learners have reached this fluency level, instruction should shift to strengthening comprehension skills through continued development of vocabulary, language skills and encouraging learners to engage critically with text.

Reaching these benchmarks in the Foundation Phase is within reach of learners, including those attending less resourced schools. By example, across non-representative samples of learners in no-fee schools, we find from pre-pandemic assessments that:

- By the end of Grade 1, 24 percent of a Setswana learner sample and 32 percent of a Sepedi learner sample attained the letter-sound benchmark of 40 correct letters per minute.
- By the end of Grade 2, 42 percent of Setswana learner samples, 32 percent of Sepedi learner samples and 51 percent of Sesotho learner samples were reaching the Grade 2 fluency benchmark.
- By the end of Grade 3, 24 percent of Setswana learners were reaching the Grade 3 fluency benchmark and 51 percent were reaching the Grade 2 fluency benchmark.

These examples confirm that the benchmarks are attainable, yet learners are acquiring decoding skills (such as letter-sound knowledge) and fluency far too slowly. Considerable progress will need to be made for all learners to reach these benchmarks in the Foundation Phase.

As benchmarks are increasingly used to assess and track reading in the Foundation Phase through systematic measurement, we will gain more understanding of how well children are able to keep pace with these African language benchmarks to support improvements in reading for meaning.

1. Introduction

Although more children in low- and middle-income countries are in school and staying in school longer than ever before, a large proportion are not acquiring the fundamental skills required to build the necessary human capital to support sustainable growth and poverty reduction. This 'learning crisis' has led to international organisations and governments prioritising foundational literacy. In the 2019 State of the Nation Address, the President of South Africa stated that "all 10-year-olds should be reading for meaning" as a priority (South African Government, 2019).

There is a growing consensus that *early* measurement matters locally and internationally. South Africa is one of the few low- and middle- income countries that participates in international assessments such as the Progress in International Literacy Reading Study (PIRLS). Despite significant gains in reading proficiency over the past decade, seen in both PIRLS and the regional Southern and Eastern Africa Consortium for Monitoring Educational Quality (SEACMEQ), reading outcomes in South Africa remain weak. The 2016 PIRLS results revealed that, according to their five-tier benchmark categories, 78 percent of South African Grade 4 learners were unable to reach the lowest benchmark[1], in stark contrast to the international average of 4 percent (Howie et al. 2017).

The PIRLS results are valuable in highlighting low learner competency at the Grade 4 level in the ultimate reading skill, written comprehension. However, PIRLS provides no information on where learners are falling behind in their foundational reading skills that underpin being able to read with comprehension. With research on learning to read in African languages still in a nascent phase, there has been limited available evidence to uncover the causes of poor reading comprehension performance in Grade 4.

In response, the Department of Basic Education (DBE) has led a collaborative effort with various stakeholders including South African linguists, academics and reading practitioners, funders, and international benchmarking specialists to establish early grade reading benchmarks in all South African languages. Efforts to date include a design report outlining methodological approaches and grades and skills to benchmark; and the establishment of benchmarks for Nguni languages through the quantitative analysis of existing reading assessment data drawn from several early grade reading studies involving 16,400 unique learners in more than 660 schools (Ardington et al. 2020).

Based on the DBE's ongoing first Early Grade Reading Study (EGRS I), other existing data and planned data collection, the Sesotho-Setswana languages were identified as the next priority. The EGRS I has been collecting longitudinal reading assessment data on a large sample of Setswana learners since 2015. The fifth EGRS I wave of data collection in 2021, provided an opportunity to collect data specifically for benchmarking purposes. The DBE identified that Room to Read and Save the Children had existing Grade 2 Sepedi and Sesotho data respectively and approached Funda Wande to expand the planned data collection for their impact evaluation in Limpopo. All three organisations agreed to collaborate to enable the development of Sepedi benchmarks and the anticipated creation of Sesotho-Setswana benchmarks.

Aims

The purpose of this report is to provide a consolidated set of early grade reading benchmarks for Sesotho-Setswana languages. The report is intended to provide a joint analysis from the primary reports and data for the Sepedi and Setswana reading benchmark reports, as well as incorporate Sesotho data. The individual language reports provide a rich and extensive discussion and analysis, serving as technical reference reports. This report serves as a summary report bridging across all three languages.

Report Structure

The next section provides a brief overview of the theoretical framework for reading development that motivates benchmarking specific foundational skills and discusses the need to develop benchmarks for African language groups. The following section focuses on the three Sesotho-Setswana languages, synthesizing information on cross-language differences and similarities in phonology, morphology and orthography. Section 4 outlines the methodological approach, while Section 5 describes the data and provides summary results of the benchmarking analyses. The established benchmarks are then presented in Section 6. The report concludes with a summary and recommendations for future use and analysis.

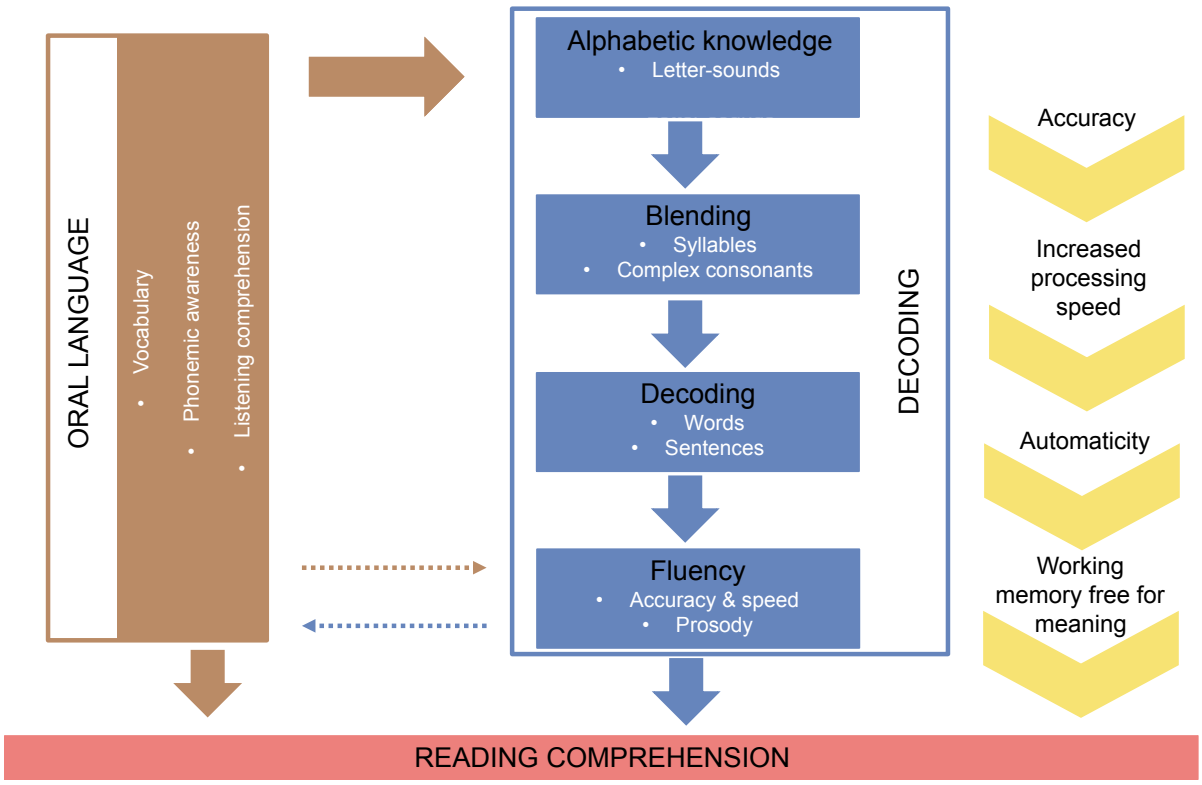
2. Background

2.1. Reading acquisition and foundational skills

While the goal of reading is to construct meaning from text, reading comprehension is a complex and hierarchical process requiring the development and coordination of multiple foundational skills and processes. Figure 1 illustrates how these skills progress and build on one another, culminating in reading comprehension. Oral language skills (vocabulary, listening comprehension, phonemic awareness), acquired through listening and speaking, reflect a child's understanding of the language in which he/she will learn to read. The initial connection between the language a child understands and the written code of that language is realized through phonics where alphabetic knowledge – knowing how sounds (phonemes) are represented by letters (graphemes). This is the first level of decoding. The next level on this hierarchy is the blending together of these sounds represented by letters to form syllables and words. The subsequent level of reading acquisition is fluency, the ability to read with accuracy, speed, and proper expression (prosody).

Within each skill, accuracy tends to develop first, followed by increased speed as decoding becomes more automatic, rapid, and effortless freeing up working memory and attention for meaning construction. Betts (1946) classified learners as reading at either the independent, instructional or frustration level based on a combination of their word reading accuracy and comprehension. In terms of accuracy, learners reading at the independent level read with at least 99 percent accuracy, those at the instructional level read with at least 95 percent accuracy and readers at the frustration level are reading with less than 90 percent accuracy[2]. A review of recent evidence supports the continued use of these levels (Allington, McCuiston & Billen 2015).

Figure 1. Reading acquisition processes



Fluency is necessary, albeit not sufficient, for learners to fully comprehend what they are reading. Although the skills depicted in Figure 1 are hierarchical, they are inter-connected and do not develop in a strictly linear fashion. For example, knowledge of the language (e.g. vocabulary) is essential for comprehension and there is a feedback from gains in fluency towards improved vocabulary through greater exposure to the language.

While fluency builds a bridge between decoding and reading comprehension (Chard, Pikulski & McDonagh 2006), there may be non-linearities in the relationship between fluency and comprehension. The decoding threshold hypothesis put forward by Wang et al. (2019) suggests that reading comprehension is unlikely to develop until decoding exceeds a lower bound threshold level. They also suggest that there may be an upper threshold, beyond which there are no additional gains (in comprehension) for increasing decoding skills.

2.2. Why do we need benchmarks?

At the root of poor reading comprehension outcomes at the Grade 4 level are gaps in fundamental skills that are essential for learning to read in the Foundation Phase. Benchmarks foster increased awareness of early milestones in reading development and enable teachers and schools to monitor the educational needs and progress of children. They provide a standard against which teachers can measure learners' reading subskills and identify *early on* learners who are at risk of not learning to read for meaning by age 10. This, in turn, supports remediation at an earlier age. Additionally, as specific learners reach different benchmarks, this can help teachers adapt their instructional focus to meet the learners' needs at their reading level. Beyond the classroom, benchmarks facilitate the monitoring of reading outcomes and the measurement of progress towards the goal of having all learners on track for a successful reading trajectory.

2.3. What skills should we benchmark?

The value of benchmarks clearly lies in their use. In deciding which skills to benchmark - a process that should be aligned to curriculum demands - consideration should also be given to ease of measurement, interpretation and clarity of communication. On the basis of these determining factors, a decision was made to produce one grade-specific benchmark for each year of the Foundation Phase. These benchmarks articulate the level that all learners should reach at the end of the grade to be on track to read for meaning by age 10. While comprehension is the ultimate goal of reading, there are serious challenges in defining, measuring and standardizing comprehension as a construct. For Grades 2 and 3, we therefore focus on Oral Reading Fluency (ORF) as it is easily measured and understood. ORF is a necessary skill to acquire to be able to read for meaning and is highly predictive of reading comprehension outcomes. As such, tracking fluency provides a useful screening measure of progress towards reading for meaning. For Grade 1 learners, letter-sound knowledge was selected as the appropriate skill to benchmark, providing an early indication of learners' decoding development.

Box 1. Measuring alphabetic knowledge and fluency

Alphabetic knowledge is the understanding of how sounds (phonemes) are represented by written letters (graphemes) in a language. In this report we measure alphabetic or letter-sound knowledge as the number of letters correctly sounded within one minute.

Oral Reading Fluency (ORF) is the ability to read aloud with accuracy, speed, and prosody. Accuracy is measured as the percentage of words read correctly while speed is measured as the number of words read within a time period, typically one minute. Prosody refers to the reading of words in a natural way that conforms to the speech rhythms with intonation patterns reflecting punctuation in the language. The assessment of prosody is subjective and difficult to measure in field studies. In this report we use the term fluency to describe the combination of speed and accuracy. Specifically, ORF is measured as the number of words read correctly within one minute.

2.4. Why do we need specific African language benchmarks?

The distinguishing structural features of a language affect the process of reading development (Malda, Nel & van de Vijver 2014) and make simple comparisons in fluency across languages uninformative. For too long, not enough was known about the processes involved in learning to read in African languages (De Vos, Van der Merwe& Van der Mescht 2014), inhibiting efficient reading instruction and monitoring for more than 70 percent of South African learners. Recently, several evaluations of early grade reading programmes and other studies, including the benchmarking efforts, have made considerable contributions to advancing knowledge of these processes. Box 2 highlights key structural features of the Sesotho-Setswana language family.

Box 2. Key structural characteristics of Sesotho-Setswana languages

Morphology refers to the internal structure of words and how they are put together.

Sesotho-Setswana languages are *agglutinating*. This means that words are made up of a sequence of morphemes (the smallest meaningful unit in a language) with each component of meaning represented by its own morpheme.

Phonology refers to the system of speech sounds in a language.

Sesotho-Setswana languages have a *tonal* phonology and include click consonants.

Orthography refers to how spoken language is represented in written form.

Sesotho-Setswana languages have a *transparent* orthography which means that there is mostly a one-to-one mapping between letters (graphemes) and sounds (phonemes).

Sesotho-Setswana languages have a *disjunctive* orthography which means that a spoken word may correspond to a number of written words. For example, ke a ba rata(Setswana) - I love them (English).

Unlike English which has an opaque orthography, languages that fall under the Sesotho-Setswana and Nguni groups have a transparent orthography which means that a letter (grapheme), in most instances, matches one-for-one to a sound (phoneme) (Machobane, Matlosa & Mokitimi 2003; Wilsenach 2019). Accuracy tends to develop more rapidly in languages with a transparent orthography than in languages with an opaque orthography (Malda, Nel & van de Vijver 2014; Wilsenach 2019; Wills, Ardington & Sebaeng 2022). However, this advantage is partly offset by the prevalence of complex consonant sequences especially in Nguni languages and to a lesser extent in Sesotho-Setswana languages, even in early grade texts (Malda, Nel & van de Vijver 2014). In Setswana, examples include digraphs *ng, ts,* trigraphs such as *tsh,* and blends such as *ngw, tshw.* The complexity of these consonant sequences may either be phonological or orthographic. Regardless of the source of their complexity, knowledge of complex consonant sequences is vital to learn how to read.

Differences in language structure that affect reading acquisition can also occur within broader language family groups. For example, whilst the Sesotho-Setswana language group and Nguni group both fall under the Southern Bantu[3] language family they differ in orthography. Sesotho-Setswana languages are disjunctive, consisting of shorter words whereas Nguni languages have conjunctive orthographies. The reading speed and accuracy scores identified across orthographies confirm that foundational skills like phonological decoding develop in response to the nature of the language-specific orthography (Share 2021). It is therefore imperative to move away from a 'one size fits all' approach to establishing benchmarks for each language group.

The specific linguistic and orthographic features of the languages in the Sesotho-Setswana family, including differences between them, are explored in more detail in the following section.

3. Sesotho-Setswana languages

The Sesotho-Setswana language group is made up of three languages: Sesotho, Setswana and Sepedi. All three languages are part of the 11 official languages recognised in the South African Constitution (1996) reflecting the cultural diversity of the Republic of South Africa. The demographics of the speakers of these languages are summarised in Table 1.

Table 1. Sesotho-Setswana language speaker demographics

	Setswana	Sesotho	Sepedi
Population speaking the language as a First Language	4.1 million	4.6 million	4.6 million
% of the SA Population as first-language speakers	8%	8%	9%
Province(s) where language is spoken	North-West and Gauteng	Free State, Eastern Cape and Gauteng	Limpopo, Mpumalanga and Gauteng
Non-SA locations where language is spoken	Botswana, Namibia and Zimbabwe	Lesotho, Namibia and Zambia	Botswana

Source: Statistics South Africa, 2012

Sesotho-Setswana languages are classified under the Southern Bantu language group (Messerschmidt et al. 2008). Each of the languages have several dialects. However, even with the presence of dialectal differences within each of the languages, these three languages share many common features such as nouns, pronouns, demonstratives, qualificatives, verbs, copulatives, adverbs, ideophones, interjections, conjunctions and interrogatives that allow speakers to communicate easily and clearly with one another. The differences lie in phonological segmental, tonal, morphological, and syntactic differences, which qualify each language as autonomous in its own right (Poulos and Msimang 1998).

3.1. Tone features in Sesotho-Setswana languages

Sesotho, Sepedi and Setswana are all tonal languages, spoken using mainly two contrasting tones: low and high. The most important property of tonal languages, which distinguishes them from languages that merely use the pitch as part of intonation, is the existence of numerous tonal minimal pairs. Often, a few words may be composed of exactly the same syllables/phonemes yet have different characteristic tones. Speakers of these languages vary their voice and pitch (either high or low, or level, rising, or falling) when articulating certain words in order to distinguish the meaning of one word from the other (Demuth 1993). For example nòkà (waistline); nókà (spicing food) and nóká (river)[4]. It is difficult for learners at the foundation phase level to differentiate homonyms according to their tonal differences. Tonal differences may make it difficult for learners to pronounce and comprehend the words effectively.

3.2. Vowels

Sesotho, Setswana and Sepedi languages have structures which are based on three types of sounds, namely; vowels, semi-vowels, and consonants. These languages have seven vowels. Table 2 shows the standard vowels with examples words.

Table 2: Standard vowels in Sesotho, Sepedi and Sepedi

Vowels	Example words					
	Sesotho	English	Sepedi	English	Setswana	English
a	hapa	conquer	bala	read	rata	like, want, love
e	sebetsa	cork	sepela	walk	lema	plough
ê	-	-	bêka	to marry	rêma	chop
i	bina	sing	rita	brew	dira	do
o	motho	human being/person	motho	human being/person	motho	human being/ person
ô	-	-	bôla	rot	tôrô	dream
u	hula	pull	bula	open	khudu	tortoise

The three languages differ in the use circumflex diacritic mark (ˆ) for (ê) and (ô), for Setswana and Sepedi, the circumflex sign/diacritic mark should be used to differentiate the ê from e and ô from o while Sesotho does not. (Department of Education and Training, 1988:6; PanSALB, 2019). However, in Setswana and Sepedi classrooms the diacritic markings are not typically taught or used. The difference in sounds is generally derived from context.

3.3. Semi-Vowels

The Sesotho-Setswana language group feature two semi-vowels namely /w/ and /y/. Table 3 presents examples of semi-vowels.

Table 3. Semi-vowels in Sesotho-Setswana language group

Semi-vowel	Sesotho/Sepedi/Setswana	English
w	wela wena	fall into you
y	boya moya	fur air

3.4. Diphthongs

In the Sesotho-Setswana languages there are no diphthongs, however, there are combinations of some basic vowels as illustrated in the examples in Table 4.

Table 4. Combination vowels in Sesotho, Setswana and Sepedi with English meanings

Combination vowel	Sesotho/Setswana/Sepedi	English meaning
ao	maoto	feet
oa	boatla	careless

ae	mae	eggs
éi	eiye	onion
ia	diatla	hands
oé	boele	return

3.5. Consonants

Sesotho-Setswana, like other African language groups, is characterised by a transparent orthography, which as discussed, implies that there is a high incidence of individual letters which represent only one sound.

For example; the word /bona/ can be broken down into four letters into four letters [b] + [o] + [n] + [a] that map on to four phonemes/b/+ /o/+ /n/+ /a/.

Table 5. Simple consonants in Sesotho-Setswana languages

Simple Consonant	Sounds like			Example	
				Sesotho/Sepedi/Setswana	English translation
b	b	in	bat	beke	week
d	d	in	die	dula	sit
f	f	in	fly	fofa	fly
g	g	in	gorrel (in Afrikaans)		gas
h	h	in	head	hema	breathe
j	j	in	June	ja	eat
k	k	in	keep	kolobe	pig
l	l	in	lay	loma	bite
m	m	in	man	motho	person
n	n	in	norm	nama	meat
p	p	in	pink	padi	novel
q	q	in	qala (in isiZulu)	qadile	started
r	r	in	rope	rata	love
s	s	in	sale	sekolo	school
š	sh	in	shine	lesela /lešela	cloth
t	t	in	time	tau	lion
w	w	in	wet	wa	fall
y	y	in	yell	yela	that one
x	-	-	click sound	nxa	disapproval
c	-	-	click sound	cecece	sympathy

However, the challenge comes with the complex consonants where a sequence of consonants follow one another and make it difficult for learners to learn to map sounds to orthography with ease, especially at the Foundation Phase. There is a larger code set of simple and complex consonants, represented as diagraphs (two letter sounds) and trigraphs (three letter sounds) as reflected in Table 6 and Table 7 below. Simple consonants consist of single sounds that are not complicated to read and write since they are represented by only one letter.

Complex consonants are types of consonants that are represented by two (digraph) or three letter-sounds (trigraph), that when combined with vowels and other sounds create words that give a specific sound and meaning.

Table 6. Complex consonants in Sesotho, Sepedi and Setswana

Example									
Complex consonant	Sounds like			Sesotho	English translation	Sepedi	English translation	Setswana	English translation
bj				bjabjaretsa	hit	bjang	grass		
hl				hlaba	stab	hlaba	stab		
hlw				hlwekisa	clean	hlweka	clean		
hw				hwama	solidify	hwa	die		
kg				kgaka	guinea fowl	kgaka	guinea fowl	kgomo	cow
kgw				kgwele	string	kgwele	ball		
kh	k	in	king	khora	become full	khora	become full	khiba	apron
kw				kwana	there	kwa	hear		
lw				lwana	to fight	lwala	be sick		
mm				mmele	body	mmele	body		
mph				mpho	gift	mpho	gift	mpha	give me
mp				mpa	stomach	mpa	stomach	mpa	stomach
mps				mpshe	ostrich	mpsha	new/young		
mpš						mpša	dog		
ng	ng	in	wrong	ngaka	doctor	ngaka	doctor	ngaka	doctor
ngw				ngwapa	scratch	ngwala	write		
nk				nko	nose	nko	nose		
nkw				nkwe	leopard	nkwe	leopard		
nkg				nkga	smell	nkga	smell		
nn				nne	four			monna	man
nng						nngele	left side	ngwana	child
nny						nnyane	small		
nt				nta	louse	nta	louse		

nth				ntho	something	ntho	wound		
ntl				ntlo	house	ntlo	house		
ntlw				ntlwana	small house	ntlwana	small house		
nts				ntsebe	know me	ntsebe	know me		
ntš						ntši	many		
ntsh				tsho	black	ntsho	black		
ntši						ntšhi	fly		
ntw				ntwa	fight	ntwa	fight		
nw				nwa	drink	nwa	drink		
ny	ny	in	canyon	nyala	marry	nyaka	want		
nyw						nywanywa	smile		
ph	p	in	plough			phadima	shine	phala	impala
rw				rwala	put on	rwala	carry		
tt				-	-	-	-	rra/rre	my father
shw				shwa	die	-	-		
sw				swaba	be disappointed	swa	burn		
šw				-	-	-	-	maši	milk
th	t	in	tin	hapa	wet	hapa	wet	thata	strong
tl				tlala	hunger	tlala	hunger	batla	seek
tlh						tlhago	nature	tlhako	hoof
ts				tsoha	wake up	tsoga	wake up	tsela	path
tš						tšea	take	ntšwa	dog
tsh				tshela	six	tshela	six	tshaba	run away
tšh	ch	in	church			tšhela	pour	setšhaba	nation
tshw				tshwanelo	appropriate	tshwanelo	appropriate	tshwana	same as
tšhw						tšhweu	white		
tšw						tšwafa	be lazy		
tsw						tswalela	close		

3.6. Consonants distinguishing between Sesotho, Sepedi and Setswana

Naturally, within the language group there are some differences in how some words with the equivalent meaning take on different consonants or pronunciations. A good example of this is how

in addition to the regular [s] in Sepedi, there is also the diacritic s [š] pronounced as the unvoiced fricative/ʃ/. However the latter does not occur in Sestswana or South African Sesotho orthography. Another interesting difference is how the sound [tlh] in both Sesotho and Sepedi is used for nouns e.g. /tlhokomelo/ (care) and /hl/ for verbs, e.g. /hlapa/ (bath). However, it is the opposite in Setswana because the sound [tlh] is used for verbs and [hl] or sound /ɬ/ is for nouns. Table 7 below exemplifies some of these other differences.

Table 7. Distinctive consonants for Sesotho, Sepedi and Setswana orthography

CONSONANT	SESOTHO	SEPEDI	SETSWANA
b - b´	/b/ bana (kids)	b (bilabial Fricative) – bana (kids)	b´ (bilabial plosive) > b´ana (kids)
j-bj – jw	/j/ jala (plough)	bj > bjala (plough)	jw > jwala (plough)
ng – kw	/ng/ ngola (write)	ng > ngwala (write)	kw > kwala (write)
tš – ts	/ts-/ > pitsa (pot)	tš > pitša (pot)	ts > pitsa (pot)
hl – tlh	/hl/ >hlapa (bath)	hl >hlapa (bath)	tlh > tlhapa (bath)
pš – psw		pš > bops' (built)	
ts – b	/-ts-/ > matsoho (arms)	ts > matsogo (arms)	b > mabogo (arms)
ts – tš	/ts-/ > tsoma (hunt)	ts > tsoma (hunt)	tš > tsoma (hunt)
š – s	/s/ > sala (remain)	š > šala (remain)	s > sala (remain)
tšh – tshš		tšh > tšhuma (set alight)	tš > tshuba (set alight)
šsw – š		fišwa (burned)	š > fiswa (burned)

3.7. Syllables

As a unit of spoken language, a syllable is no bigger than a speech sound and consists of one or more vowel sounds alone or of a syllabic consonant alone or of either one or more consonant sounds preceding or following.

The Sesotho, Sepedi and Setswana languages are characterised by an open syllable structure. A syllable requires a speaker to pause a little within a word when articulating the word slowly. The end of a syllable is often identified by a vowel and these languages are therefore referred to as syllabic languages (PanSALB, 2019).

A syllable may have the structure /CV/ which means that it can be any consonant followed by a vowel e.g. /ma/ from the above example. Sometimes it is represented as a digraph followed by a vowel (CCV) e.g. /sho/ and /tsa/. A vowel can stand alone as a syllable (V) which indicates that a consonant deletion has taken place. It can also be identified as a letter /n/.

3.8. Sentence structures in Sesotho-Setswana languages versus Nguni languages

The African languages spoken in South Africa are agglutinating, syllabic languages with a transparent orthography, as opposed to English being a partially analytic, stress-timed language with an opaque orthography (Spaull, Pretorius & Mohohlwane 2020). However, the Sesotho, Setswana and Sepedi

languages are written disjunctively as compared to Nguni languages which are conjunctive. This is illustrated in the example below showing the sentence meaning 'It is beautiful' in Sepedi, Sesotho, Setswana and isiZulu.

'Ke se sebotse' in Sepedi;

'Ke se setle' in Sesotho

'Ke se se pila' in Setswana

'yinhle' in isiZulu

3.9. The noun class

In Sesotho, Sepedi and Setswana, nouns are classified according to their prefixes, each class indicating whether the noun is singular or plural or a variant (Guma, 1981). The noun class prefixes play a huge role on the oral language development (discourse) of children and influence the development of literacy in terms of the acquisition of phonology, morphology, semantics and syntax. The nouns that are variants use a zero morpheme and belong to the main noun class prefix.

4. Benchmarking methodology

Our approach to establishing early grade reading benchmarks for the Sesotho-Setswana languages is based on a detailed exploration of large-scale reading assessment data that is grounded in reading development theory and guided by expert linguistic knowledge of each language. To ensure consistency in benchmarking approaches across South African languages, this is the same approach that was used to derive Nguni language benchmarks (Ardington et al. 2020, 2021). We draw on an understanding of curriculum demands and system realities to ensure that the benchmarks are contextually appropriate. In this section, we briefly summarize the key insights from reading development theory that motivate our approach and then describe the steps in our empirical strategy. See Moholwane, N., Wills, G. & Ardington, C. (2022) for a detailed discussion of the approach to benchmarking.

4.1. Conceptual underpinnings

Our understanding of reading development, outlined in Section 2, informs our approach to benchmarking in the following ways:

1. Reading is hierarchical, with the development of lower-level skills necessary for the development and application of higher order skills (Stanovich 2000). This supports establishing benchmarks for lower order skills to ensure that learners are on a successful trajectory for learning to read for meaning. Letter-sound knowledge has been shown to be predictive of later oral reading fluency. Benchmarking this foundational skill provides a means of identifying at-risk learners early on at the lower end of the hierarchy.
2. Reading comprehension is not a simple construct and is challenging to assess in an equivalent

or reliable manner. Yet ORF is an important skill and is a reasonable proxy for comprehension. ORF is easily understood and measured, making it an appropriate skill for benchmarking. The focus of the empirical work is to identify the fluency level where decoding skills are sufficiently established to support reading comprehension.

3. We explicitly acknowledge the importance of accuracy in reading by focussing on the relationship between accuracy and speed of reading before turning our attention to understanding the relationship between fluency and comprehension. For learners reading at speeds below the instructional level of accuracy (which is getting 95 of every 100 words correct)[5], decoding is likely to be laboured hindering the ability of the learner to make meaning from the text.
4. Reading development theory suggests there may be non-linearities in the relationship between fluency and comprehension. Initially, comprehension may increase steeply when fluency increases but comprehension improvements may start to get smaller at higher fluency levels. Our exploratory data analysis aims to identify critical decoding thresholds in learners' reading development. We specifically look for fluency points below which comprehension is unlikely to develop further and examine if there is evidence of an upper threshold where limited comprehension skills become a constraint and there are no further gains to increasing fluency.
5. Differences between languages necessitate language-specific benchmarks. African languages are understudied and we are careful not to impose any *a priori* assumptions on the accuracy-speed and fluency-comprehension relationships. Our empirical approach relies heavily on exploratory data analysis to uncover these relationships for early grade readers in each language.
6. Given the low levels of reading proficiency in our context, we use data from later grades to understand what a successful trajectory could look like. Longitudinal data allow us to examine the predictive validity of proposed benchmarks for an 'on track' successful reading journey.
7. Benchmarks need to be contextually appropriate and cognisant of curriculum requirements. On the one hand, setting benchmarks at a level that is out of reach for most learners limits their usefulness in tracking incremental improvements or guiding remediation or instruction. On the other hand, benchmarks need to be set high enough to encourage system improvements toward levels that are appropriate for the demands of the curriculum. While our approach is data driven, we are careful to examine the attainability of proposed benchmarks and engage in expert opinion on the appropriate grade level at which to set each benchmark.

Box 3. Exploratory non-parametric methods versus traditional benchmarking approaches

Exploratory non-parametric methods versus traditional benchmarking approaches

Typical approaches to benchmarking focus on identifying the fluency levels associated with achieving a fixed comprehension threshold, for example at least 80 percent of questions correct (Room to Read 2018; Abadzi 2012; RTI 2010). Our approach as described in Ardington et al. (2021) has a number of advantages over the traditional approach.

First, reading benchmarks are language and context specific and need to be set in way that is responsive to patterns emerging from the data. Non-parametric methods make no assumptions about the speed-accuracy or fluency-comprehension relationships which can be affected by both pedagogical and linguistic differences.

Second, our approach to identifying critical thresholds in the accuracy-speed and fluency-comprehension examines the full distribution of these relationships whereas traditional methods only focus on these relationships around the specific comprehension cut-off.

Third, traditional methods assume that comprehension is an easily defined and comparable construct across passages and languages. There is plenty of evidence to the contrary and the individual language technical reports (Ardington et al. 2020; Wills et al. 2022) highlight the serious challenges of establishing the appropriate level of comprehension questions. Our approach is much less sensitive to these challenges than traditional approaches that focus on a particular comprehension cut-off.

A disadvantage of our approach is that it requires some degree of expert subjective judgement. However, a prescriptive, formulaic approach to benchmarking runs the risk of setting benchmarks that are neither contextually appropriate nor informative for tracking incremental improvements or guiding remediation or instruction. For example, RTI International (2017) report that across African countries only around 5 percent of learners were reaching the established benchmarks. We instead are guided by both the patterns that emerge from the data and the current realities of South African classrooms. This developmental approach enables the measurement of incremental improvements over time in a low-literacy context.

Source: Extract adapted from Ardington et al. 2022

4.2. Empirical approach

The aim of this report is to establish appropriate letter-sound knowledge and oral reading fluency benchmarks to map out a successful reading trajectory for Sesotho-Setswana learners through the Foundation Phase.

4.2.1. Establishing ORF benchmarks

The steps in our empirical process are as follows:

1. **Examining the relationship between speed and accuracy.** Using locally weighted polynomial regressions, we investigate the relationship between speed and accuracy paying particular attention to the speed associated with the instructional level of accuracy identified by Betts (1946). The lower threshold is then set around the speed below which the average learner is below the instructional level of accuracy across all passages and grades.
2. **Examining the relationship between fluency and comprehension.** We then use the same non-parametric approach to explore the relationship between fluency (a measure of both speed and accuracy) and comprehension. We consider whether learners struggle to comprehend what they read when their fluency levels are below the lower threshold suggested by the accuracy-speed relationship. We then seek to establish the fluency level necessary to support comprehension, paying particular attention to whether there is evidence of an upper threshold below which there are limited improvements in comprehension with increased fluency.
3. **Concurrent and predictive validity.** Once these thresholds are identified, we use concurrent data on related reading skills to establish whether these potential benchmarks align with meaningful distinctions between learners and the stages of reading development. Next, we establish the concurrent validity of the fluency thresholds by examining how they align with the performance of the same learners on written comprehension assessments. For samples where we have longitudinal data, we investigate the predictive validity of the thresholds by examining whether meeting the thresholds at earlier grade points is predictive of learners' future fluency and comprehension levels.
4. **Contextual alignment.** Finally, we investigate whether the potential benchmarks are contextually appropriate by examining the proportion of current learners reading at these levels. The benchmarks need to be ambitious enough to support improvements in reading proficiency while at the same time being set at a level such that they can be used to measure incremental progress and inform instructional focus in the classroom. Our process of setting benchmarks therefore involves both backward and forward analyses of the data.

4.2.2. Establishing letter-sound benchmarks

Using longitudinal data and drawing on expert opinion, the Nguni benchmarking report identified 40 correct letter-sounds per minute as an appropriate minimum benchmark for the end of

Grade 1 (Ardington et al. 2020). Reaching this level was predictive of reaching later oral reading fluency benchmarks and data indicated that there were diminishing improvements in letter-sound knowledge once learners had reached 40 correct letter-sounds per minute. Despite differences in pronunciation, one wouldn't expect significant differences in the process of letter-sound acquisition across alphabetic languages. For the Sesotho-Setswana languages we focus on the extent to which the Grade 1 letter-sounds benchmark is appropriate in terms of reachability and predictive validity.

5. Benchmarking analysis

5.1. Data

The establishment of Sesotho-Setswana early grade reading benchmarks was based on reading assessments of 24,686 unique learners across 429 no-fee schools in the North West, Free State and Limpopo provinces (Table 8).

Table 8. Grades at which assessments were conducted, by individual study

Study	Grades	Language	No. of schools	No. of learners
Room to Read	1 & 2	Sepedi	67	3,450
Funda Wande	1, 3 & 6	Sepedi	120	4,729
EGRS I and RSP	1, 2, 3, 4 & 7	Setswana	230	15,851
Save the Children	2	Sesotho	12	656
		Total	429	24,686

The Setswana data are from the Early Grade Reading Study (EGRS I) and the Reading Support Programme (RSP). The EGRS I is an impact evaluation of a three year (2015-2017) early literacy intervention involving teacher training and coaching together with structured lesson plans and materials provision (DBE 2017). In 2015 the EGRS I started tracking the reading outcomes of a cohort of Grade 1 learners. These same learners were assessed five times between Grade 1 and 7 (2015 to 2021). In addition, a sample of Grade 3 learners were assessed in 2018 to measure the sustainability of the programme. As the EGRS I intervention ended in 2018, the RSP commenced in a subset of the original EGRS I schools enrolling a new cohort of Grade 1 learners. These learners were re-assessed in Grade 4 (2021). Additional or 'top-up' samples of Grade 3, 4 and 7 learners were added in the 2021 data collection to support the establishment of Setswana benchmarks (Wills et al. 2022).

The Sepedi data come from the Funda Wande and Room to Read projects. Room to Read conducted Sepedi reading assessments at the Grade 1 and 2 level over the period 2016 to 2019. The impact evaluation of the Funda Wande programme in Limpopo planned to conduct early grade reading assessments with Grade 1 learners in 2021 (Ardington & Henry, 2021). At the request of DBE, the data collection was expanded to include Grade 3 and 6 for benchmarking purposes.

The Sesotho data are drawn from a Save the Children programme called Literacy Boost for Sesotho Benchmarking Research. The programme was implemented in the Free State during the periods 2012

to 2014 with Grade 2 learners. The dataset does not include letter-sound knowledge or measures of accuracy for passage reading. We are therefore restricted to examining the fluency-comprehension relationships and the percentage of learners reaching fluency benchmarks.

Before 2021, data were mostly collected for evaluation or monitoring purposes and reading assessments were not specifically designed to support the establishment of reading benchmarks. The new EGRS I and Funda Wande data collection in 2021 presented an opportunity to conduct a rigorous instrument development process led by African language specialists collaborating with DBE and quantitative researchers. The Setswana assessments were developed through three rounds of piloting and revision. The Sepedi assessments drew on the Setswana design process, and also allowed for two small pilot studies.

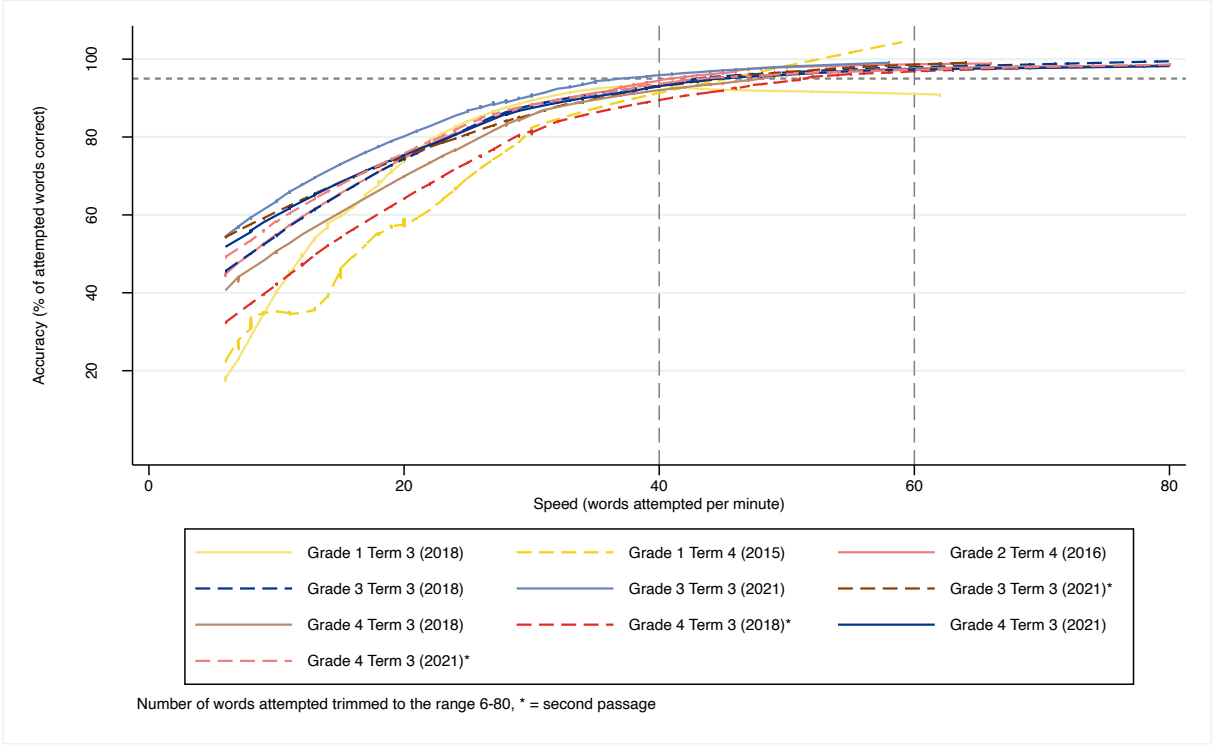
5.2. Establishing Oral Reading Fluency benchmarks

5.2.1. The relationship between accuracy and speed

The relationship between reading speed and comprehension is moderated by accuracy in reading, with errors (i.e. inaccuracy) both reducing speed and cluttering working memory. We therefore begin with an analysis of the accuracy-speed relationship using non-parametric methods to visually examine the levels of reading accuracy associated with each level of reading speed in Setswana (Figure 2) and Sepedi (Figure 3) with separate lines for each unique combination of grade, term and reading passage. Reading speed is measured by the number of words attempted in one minute while accuracy is measured by the percentage of those words attempted that are correctly read. The figures include grey dashed vertical lines at 40 and 60 words per minute as well as a dashed horizontal line representing 'the instructional level' of accuracy of 95 percent (i.e. for every 100 words attempted, the learner gets 95 words correct).

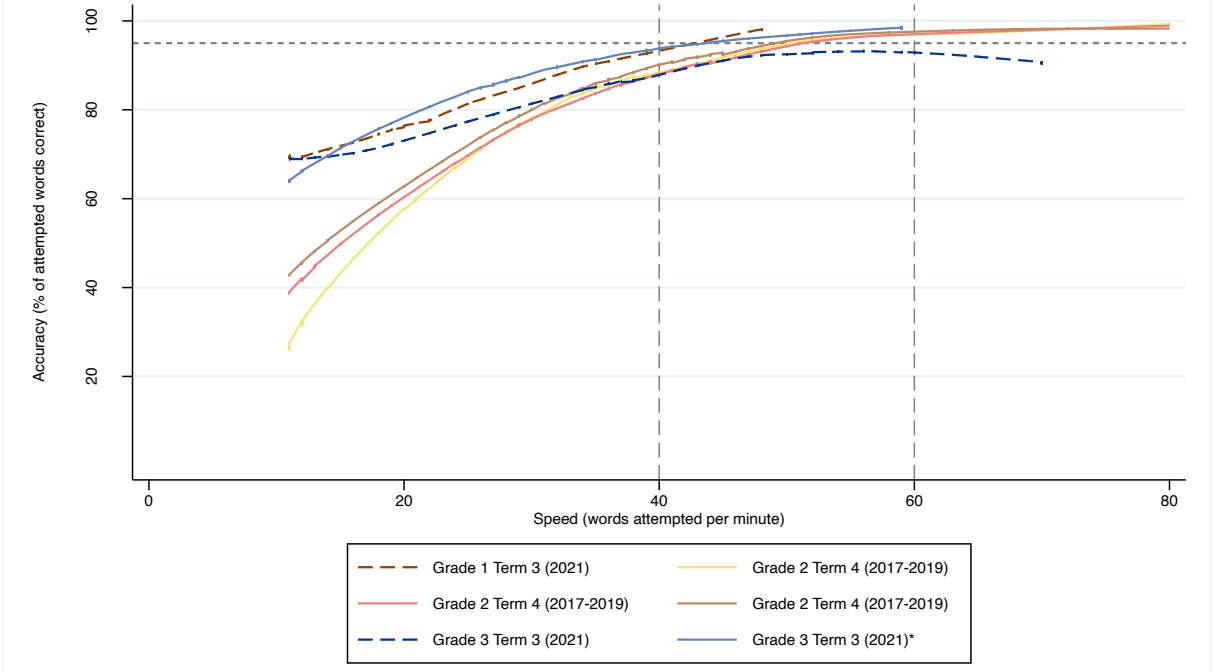
We observe a consistent pattern where initially accuracy and speed increase steeply together and then the relationship tends to flatten off at accuracy levels around 95 percent. For example, by the end of Grade 3 in 2018, Setswana learners attempting around 8 words per minute are making an error on every second word. Learners attempting 20 words per minute read three out of four words correctly. Accuracy and speed rapidly improve together and learners that reach 95 percent accuracy are usually reading at a speed of around 44 words per minute. After this there are very little changes in accuracy observed with increasing speed, suggesting that an accuracy threshold has been reached. This analysis suggests that learners reading at speeds below 40 words per minute are reading at 'the frustration level' getting less than 90 of every 100 words correct. This decoding point is then the lower threshold.

Figure 2. Speed and accuracy in reading, Setswana



Number of words attempted trimmed to the range 6-80, * = second passage

Figure 3. Speed and accuracy in reading, Sepedi



Number of words attempted trimmed to the range 11-80, * = second passage

5.2.2. The relationship between fluency and comprehension

Next, we consider the relationship between oral reading fluency and comprehension (identified from verbal answers given in response to questions asked about the oral reading fluency passage a child reads). The aim is to establish the fluency level necessary to support comprehension and to examine whether learners at fluency levels below the lower threshold are indeed at the frustration level, struggling to comprehend what they read. Again, we use non-parametric methods to visually summarise the comprehension level associated with each level of fluency for all the unique grade, term and passage combinations.

Although we observe differences in the average comprehension level between samples, the fluency-comprehension gradient is remarkably similar across languages, grades and reading passages (Figure 4, Figure 5 and Figure 6). The gradient is very steep below 40 correct words per minute (cwpm) with learners reading below this lower threshold having very poor comprehension skills. For the learners that read between 40 and 60 cwpm, increased fluency is associated with improvements in comprehension. The fluency-comprehension gradient tends to flatten out at around 60 cwpm. This flattening occurs at fairly low comprehension levels (between 50 and 80 percent of comprehension questions correct, in most of the samples) suggesting that underdeveloped comprehension skills become the key hurdle for learners at these higher levels of fluency.

Figure 4. Fluency and comprehension, Setswana

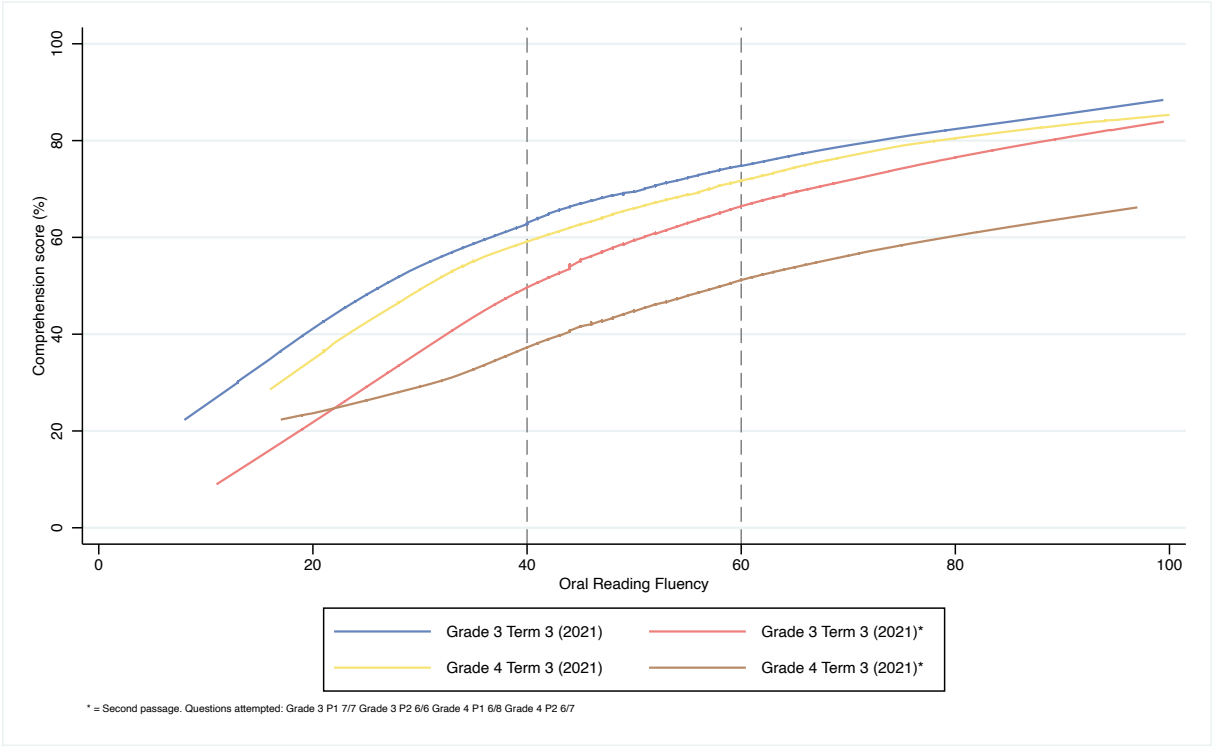
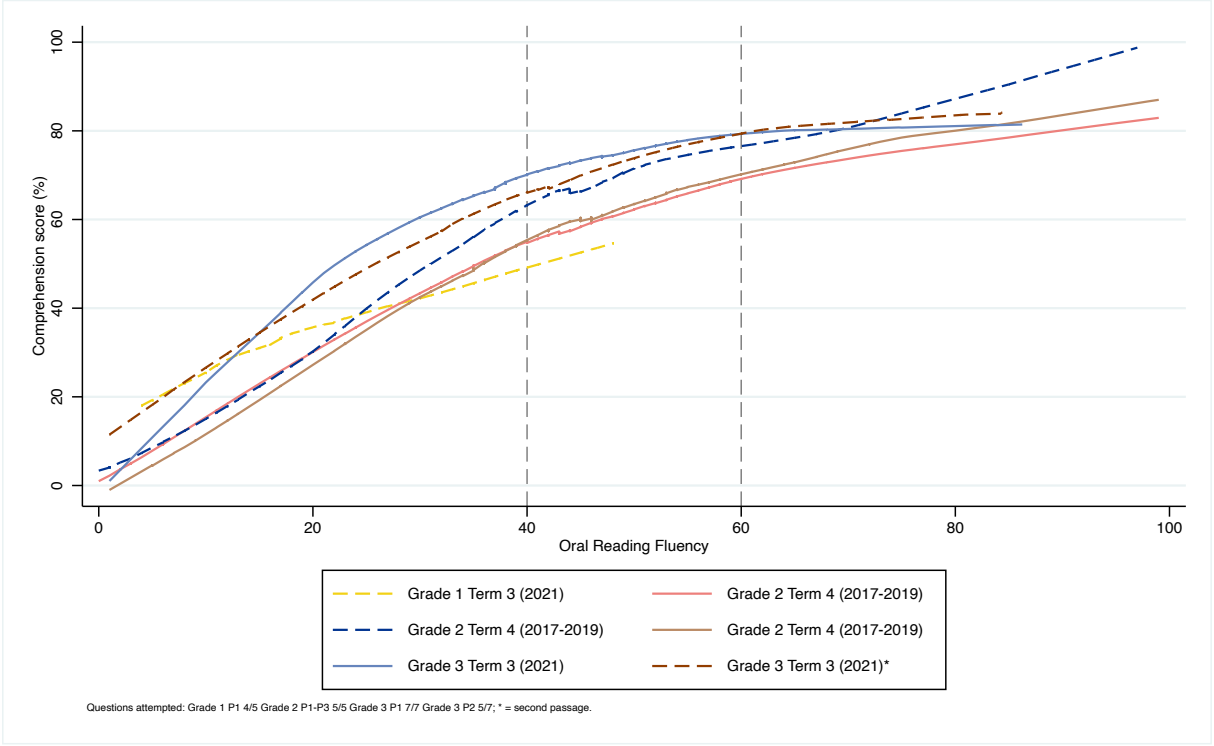
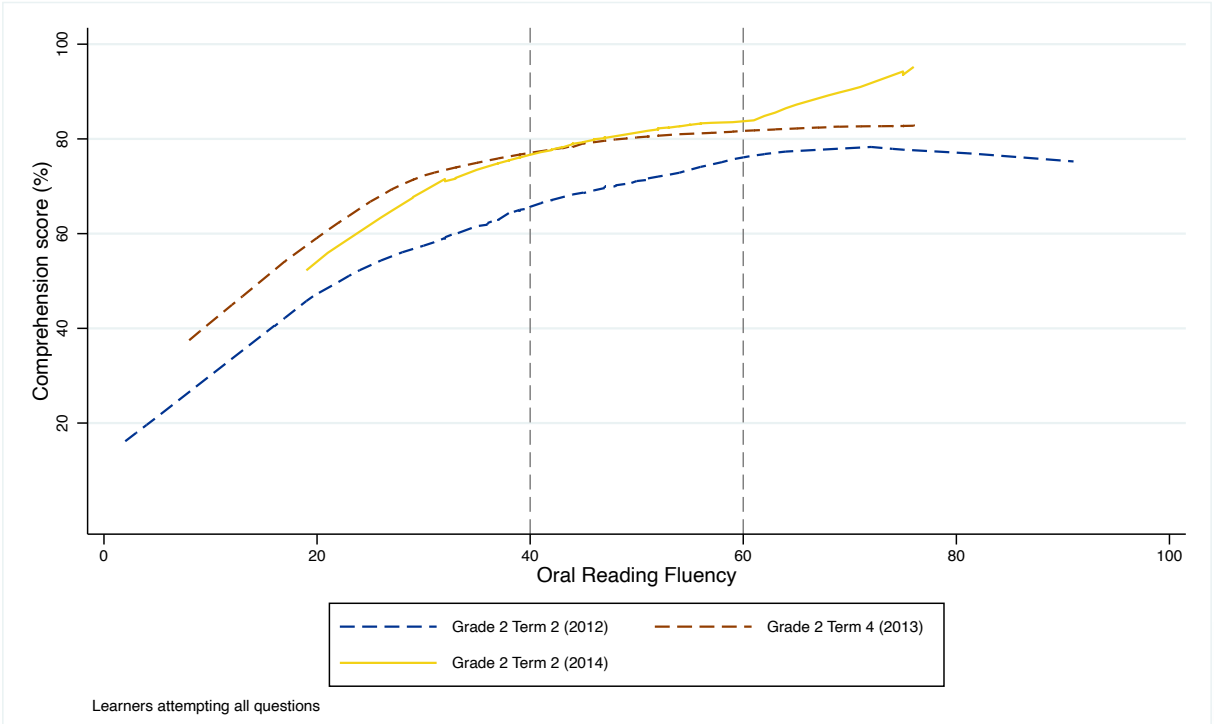


Figure 5. Fluency and comprehension, Sepedi



Questions attempted: Grade 1 P1 4/5 Grade 2 P1-P3 5/5 Grade 3 P1 7/7 Grade 3 P2 5/7; * = second passage.

Figure 6. Fluency and comprehension, Sesotho



Learners attempting all questions

This analysis reveals regular patterns across languages, Foundation Phase grades and reading passages supporting the identification of:

- a lower threshold at around 40 cwpm
- a higher threshold at around 60 cwpm

We now turn to consider the concurrent and predictive validity of these thresholds, before examining whether they are contextually appropriate.

5.2.3. Predictive validity of fluency thresholds

The longitudinal nature of the EGRS I sample allows us to investigate the predictive validity of the lower and upper thresholds in terms of future oral reading fluency levels and future written comprehension outcomes in Setswana. We begin by tracking the fluency levels of EGRS I learners from the end of Grade 2 to the end of Grade 4 (2016-2018), and then from Grade 4 to Grade 7 (2018-2021). We separate learners into initial fluency categories: non-readers (reading 0 cwpm), those not meeting the lower threshold (reading 1-39 cwpm), those meeting the lower threshold (reading 40-59 cwpm) and those meeting the upper threshold (at least 60 cwpm). By initial fluency category, we then identify learners' fluency category at a later grade assessment (Figure 7). Learners who were already reaching the upper threshold at the initial point are excluded from the figure. Three clear patterns can be seen when we consider the Grade 2-4 and Grade 4-7 transitions as discussed in Box 4 drawing on findings in Wills et al. (2022) and in reference to Figure 8.

Box 4: How meeting 'thresholds' predicts later reading acquisition

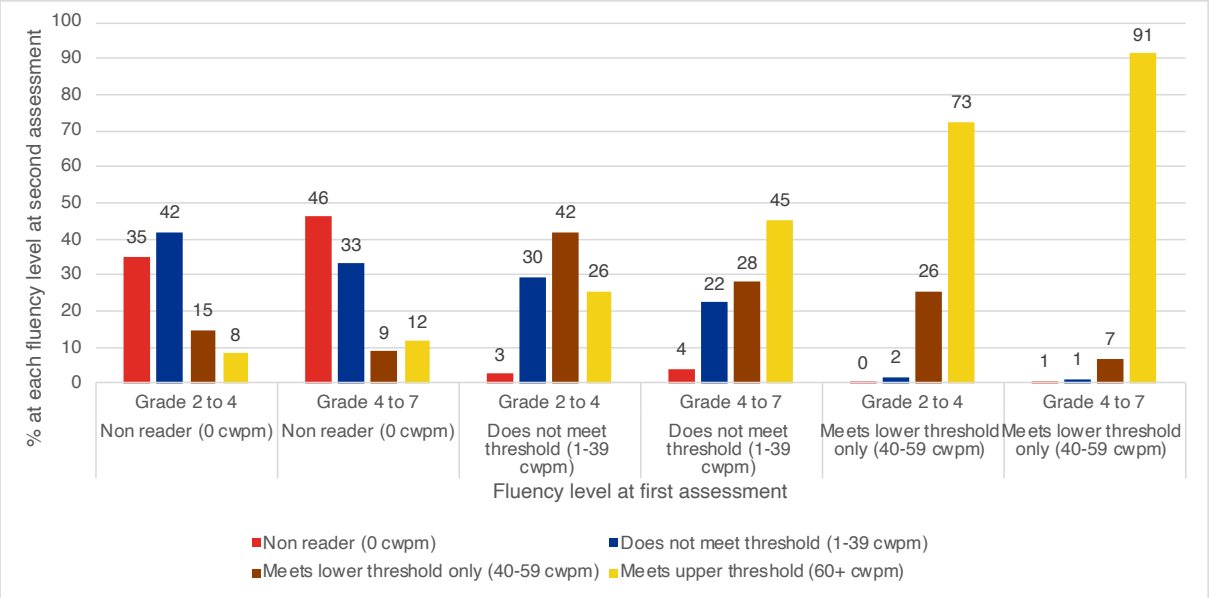
Non-readers stagnate. About 35 percent of Grade 2 learners who were non-readers were still unable to read one word by Grade 4. A sizeable portion of these Grade 2 non-readers begin to read slowly by Grade 4, but most are not yet reaching the lower threshold (of 40 cwpm). By Grade 4, only 23 percent of Grade 2s reach the lower threshold, and just 8 percent meet the upper threshold. A similar picture holds in the Grade 4-7 transition. About 46 percent of Grade 4 learners who were non-readers were still unable to read one word by the end of primary school. However, a sizeable portion of these non-readers have begun to read slowly by Grade 7, but most are not yet reaching the lower threshold (of 40 cwpm). Only a small percentage (12 percent) meet the upper threshold by the time they leave primary school.

Slow readers can attain the lower threshold. Among Grade 2 learners who were reading below the lower threshold (1-39 cwpm) in Grade 2, the majority (68 percent) had reached that threshold by Grade 4, with just over a quarter (26 percent) meeting the upper threshold. Among learners who were reading below the lower threshold (1-39 cwpm) in Grade 4, the majority (73 percent) had reached that threshold by Grade 7, and 45 percent meet the upper threshold. However, such a slow pace of reading development is unlikely to support learning in primary school.

Meeting the lower threshold is highly predictive of meeting the upper threshold. An encouraging picture emerges for those meeting the lower threshold by the end of Grade 2. By the time they reach Grade 4, 73 percent of this group are meeting the upper threshold. At the Grade 4 level, of learners meeting the lower threshold, almost all (91 percent) meet the upper threshold by the end of Grade 7. The lower threshold of 40 cwpm clearly signals a point at which reading development can take off, and a key milestone in being able to meet the upper threshold of 60 cwpm.

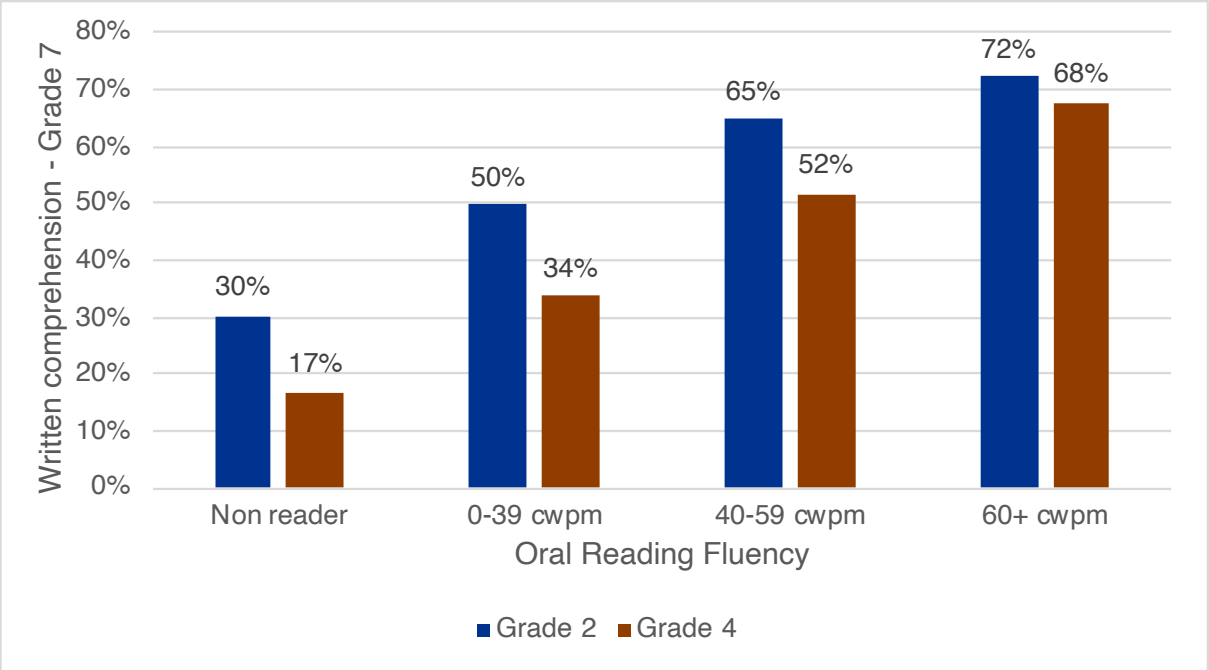
Source: Extract adapted from Wills et al. 2022

Figure 7. Fluency level at second assessment by fluency level at first assessment, an example in Setswana (EGRS I waves 3, 4 and 5)



In addition to investigating the predictive validity of the thresholds in relation to future fluency, we investigate the validity of the fluency thresholds in predicting learners' future comprehension skills. Figure 8 shows the relationship between the average written comprehension score in Grade 7 (2021) and initial fluency category in Grade 4 (2018) and Grade 2 (2016).

Figure 8. Performance on Grade 7 written comprehension by fluency category in Grade 2 and 4.



A learner's written comprehension in Setswana in the final year of primary school (Grade 7) is strongly related to their oral reading fluency in the first year of the intermediate phase (Grade 4). Learners reading below the lower threshold of 40 correct words per minute in Grade 4, score on average only 34 percent for written comprehension in Grade 7. Learners meeting this lower threshold (but not the upper threshold) in Grade 4 are getting more than half (52 percent) of the written comprehension questions correct in Grade 7, whilst those reading at or above the upper threshold in Grade 4 are, on average, scoring 68 percent for their written comprehension.

The majority of learners who have not met the lower fluency threshold by the end of Grade 2 fail to develop the reading comprehension skills that they need at the end of primary school. In contrast, those reading at or above the lower threshold of 40 cwpm (but not yet reaching the upper threshold) at the end of Grade 2, are scoring above 65 percent for written comprehension by the end of primary school. They are clearly equipped with the reading skills they need to be able to understand what they are reading in later grades. Reading at or above the lower threshold by the end of Grade 2, is an important milestone to be able to read for meaning and to learn in later grades.

5.2.4. Contextual appropriateness of the fluency thresholds

The preceding analysis has shown that meeting the lower threshold around the end of Grade 2 is indicative of being on track to reach the upper threshold by Grade 4 and to perform competently on written comprehension in Grade 7. Reaching the upper threshold by Grade 4 was associated with good written comprehension outcomes in Grade 7. We now turn to consider whether the lower and upper thresholds would serve as contextually appropriate benchmarks for the end of Grade 2 and Grade 3. Benchmarks should not be set at a level that is out of reach of the vast majority of learners as this would limit their usefulness in measuring incremental progress and assisting teachers in targeting their instruction. However, they need to be set high enough to encourage system-wide improvements and to meet the demands of the curriculum (Ardington et al. 2021).

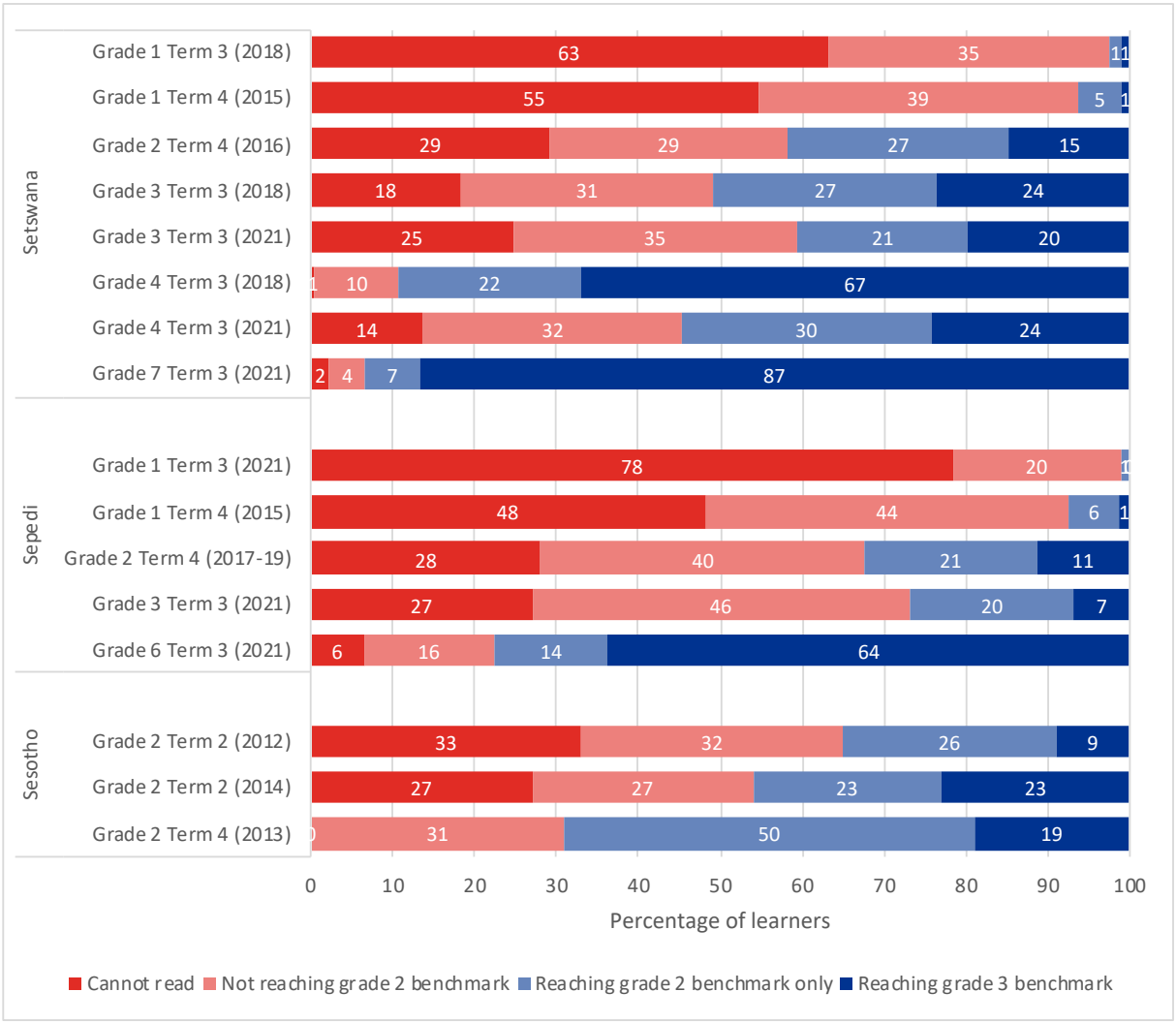
We set the benchmark for the end of Grade 2 and 3 at the lower and upper thresholds respectively and then investigate their attainability in our samples in

Figure 9. Learners are classified as 1) not being able to read (cannot correctly read one word), 2) reading at less than the Grade 2 benchmark of 40 cwpm, 3) reaching the Grade 2 benchmark of 40 cwpm or 4) reaching the Grade 3 benchmark of 60 cwpm. Although there are differences between the studies, the general progression is clear.

Attainability of the Grade 2 benchmark: Pre-pandemic, by the end of Grade 2, 42 percent of Setswana learners, 32 percent of Sepedi learners and 51 percent of Sesotho learners were reaching the Grade 2 fluency benchmark[6]. By the end of Grade 3 and 4, 51 percent and 89 percent of Setswana learners were reaching the Grade 2 fluency benchmark. In 2021, results for Setswana and Sepedi learners at the end of Grade 3 are similar to those of Grade 2 learners pre-pandemic in line with the estimated Covid-19 learning losses in the range of 1 to 1.4 years of normal learning (Ardington, Wills & Kotze. 2021).

Attainability of the Grade 3 benchmark: Pre-pandemic, by the end of Grade 3, 24 percent of Setswana learners were reaching the Grade 3 fluency benchmark. This improves to 51 percent of the Setswana learners in Grade 4. Once again, 2021 results reflect learning losses of around one year. For Sepedi learners, we do not have pre-pandemic data for the end of Grade 3. In the third term of 2021, only 7 percent of Grade 3 learners reach the Grade 3 fluency benchmark. By Grade 6, the percentage of Sepedi learners reaching the Grade 3 benchmark has risen to 54 percent suggesting that the Grade 3 benchmark is ambitious yet attainable, particularly in the absence of Covid-19 related schooling disruptions. Given South Africa's poor reading outcomes, we look to the higher Grades to examine were 'on-track' learners should be. Across both Setswana and Sepedi samples, in higher primary Grades (6 and 7), it is clear that the Grade 3 benchmark is attainable by the majority of learners.

Figure 9. Percentage of Setswana, Sepedi and Sesotho learners reaching fluency benchmarks



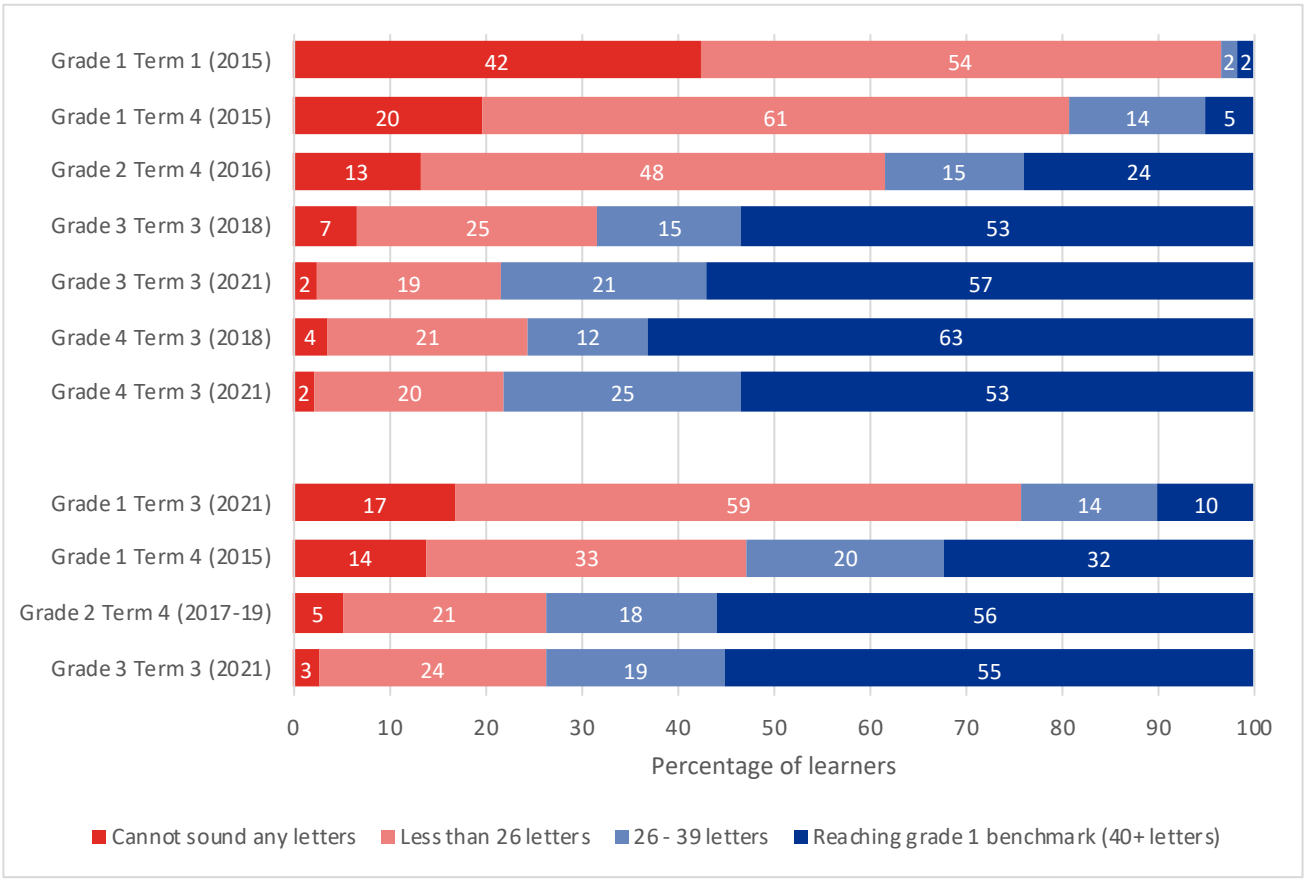
5.3. Establishing a letter-sound benchmark

The Nguni language benchmarking report (Ardington et al. 2020) established a letter-sound benchmark for the end of Grade 1 of 40 correct letter-sounds per minute. We repeated the analyses in the Nguni language report using the Setswana and Sepedi reading data and found consistent patterns in line with those for the Nguni languages data. Specifically, letter-sound accuracy and speed initially improve together steeply but the speed-accuracy gradient tends to flatten out above 40 letter-sounds per minute. In the longitudinal Setswana samples, there are diminishing improvements in letter-sound knowledge over time with limited improvements after 40 letter-sounds per minute. We now consider the attainability of the letter-sound benchmark in our Setswana and Sepedi samples.

In Figure 10 we distinguish learners, by grade samples, into four categories: cannot sound any letters correctly, sounding less than 26 letters, sounding 26–39 letters and meeting the benchmark (at least 40 letters). Pre-pandemic, by the end of Grade 1, 24 percent of Setswana learners and 32 percent of Sepedi learners attained the letter-sound benchmark of 40 correct letters per minute. By the end of Grade 2, 53 percent of Setswana learners and 56 percent of Sepedi learners are meeting this Grade 1 benchmark. This confirms that the benchmark is attainable, but the majority of learners are acquiring letter-sound knowledge too slowly and considerable progress will need to be made for all learners to reach the benchmark by the end of Grade 1. Indeed 13-14 percent of learners are unable to sound one letter at the end of Grade 1 and between 33 and 48 percent are sounding less

than 26 letters in a minute. Beyond Grade 2, the letter-sound distribution does not improve very much. Teachers are required by the curriculum to move on towards teaching higher order skills with each grade, yet this basic skill is not being mastered by learners with around 36-46 percent unable to meet the benchmark by the end of Grade 3 (and Grade 4).

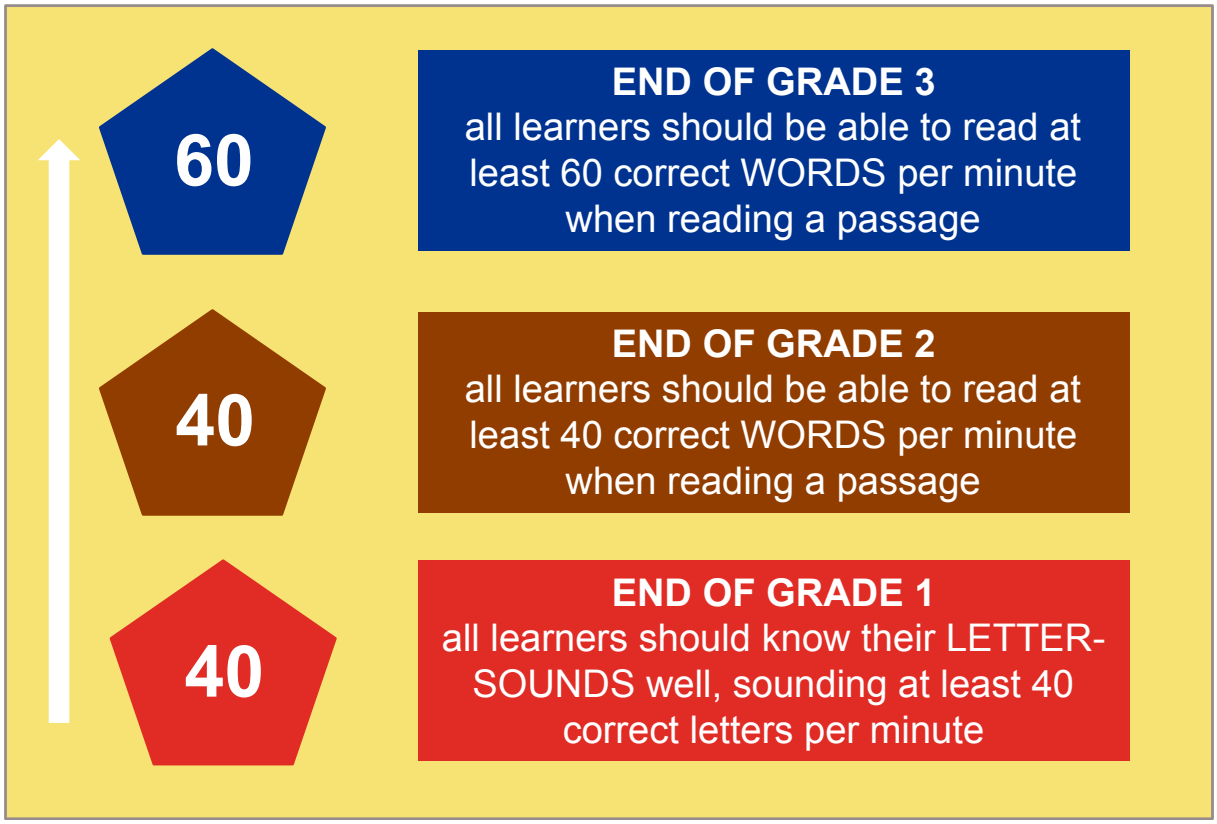
Figure 10. Percentage of Setswana and Sepedi learners reaching the letter-sound benchmark



6. Grade-specific Sesotho-Setswana reading benchmarks

The analysis supports the establishment of the following grade specific benchmarks in Sesotho-Setswana languages:

Figure 11. Reading benchmarks for early grade reading in Sesotho-Setswana languages



- By the end of Grade 1, *all* learners should be able to correctly sound 40 letters per minute.
  - Letters are a good early predictor of oral reading fluency (ORF) levels attained by the end of Foundation Phase. Improvements in letter-sound speed stagnate at 40 letters.
  - Once learners have achieved this level of letter-sound knowledge, phonics instruction should focus on blending sounds and knowledge of complex consonants while decoding instruction should focus on helping learners apply word attack strategies.
  - Pre-pandemic, between 24 and 32 percent of learners in our samples met this benchmark by the end of Grade 1.

- By the end of Grade 2, *all* learners should be able to correctly read at least 40 words per minute.
  - Below this threshold, accuracy is poor and we find little evidence that learners can understand what they have read. Quite simply, they are making too many mistakes and reading too slowly to comprehend what they are reading. For learners not meeting the Grade 2 benchmark, instruction should focus on improving decoding skills.
  - Once learners have reached this level, they would benefit from instruction that focuses on developing fluency and that exposes them to a wider range of texts.
  - Pre-pandemic, between 32-51 percent of learners in our samples met this benchmark by the end of Grade 2.

- By the end of Grade 3, *all* learners should be able to correctly read at least 40 words per minute.
    - At this level of fluency, reading comprehension becomes increasingly possible when learners read on their own. Once learners reach this level of fluency, it appears that poor comprehension skills become the limiting factor to further literacy development.
    - Once learners have reached this level of fluency, instruction should shift to strengthening comprehension skills through continued development of vocabulary, language skills and encouraging learners to engage critically with text.
    - Pre-pandemic, 24 percent of learners in our samples met this benchmark by the end of Grade 3.

7. Sector plan, alignment and next steps

By the end of 2023, benchmarks for all South African languages will have been established and released by the DBE. Effective collaboration between government, funders, research organisations and African language specialists has not only accelerated the benchmarking agenda, but also resulted in methodological innovations, established best practices and supported capacity building. That said, the value of benchmarks lies in their use.

Reading benchmarks articulate what a successful reading looks like and provide a metric against which to measure progress and develop strategies at various levels of the education system. At the national level, they provide an early indicator of whether the system is on track to deliver sustainable development goals such as all children reading for meaning by age 10. Similarly, at a provincial level they allow for the setting and monitoring of targets and feed into strategic planning for necessary support to achieve such targets. At the school level, benchmarks provide a framework to set learning expectations across grades and classrooms. Within the classroom, benchmarks signal curriculum expectations and have the potential to provide a simple way to determine learner reading levels and consequently guide instructional focus. The next priority for DBE's broader benchmarking efforts is research on how, when and where such strategies can be implemented.

With support from the Zenex Foundation and J-PAL Africa, the DBE and SALDRU are currently conducting a pilot study with teachers from 40 schools across four provinces to explore how best to support teachers in effectively using the newly established benchmarks in their classrooms.

8. Bibliography

Abadzi, H., 2012. Developing Cross-Language Metrics for Reading Fluency Measurement, Developing Cross-Language Metrics for Reading Fluency Measurement. https://doi.org/10.1596/26819

Allington, R. L., McCuiston, K., & Billen, M. (2015) What research says about text complexity and learning to read. *The Reading Teacher, 68*(7), 491-501.

Ardington, C. & Henry, J. (2022). Funda Wande Eastern Cape Evaluation. Baseline Report. Cape Town: SALDRU, University of Cape Town.

Ardington, C., Wills, G., Pretorius, E., Deghaye, N., Mohohlwane, N., Menendez, A., Mtsatse, N. & Van der Berg, S. (2020). Technical Report: Benchmarking Early Grade Reading Skills in Nguni Languages. Stellenbosch: ReSEP, Stellenbosch University. Cape Town: SALDRU, University of Cape Town. Chicago, ILL: NORC at the University of Chicago. Pretoria: DBE.

Ardington, C., Wills, G. & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development,* 86 (October 2021), 102480.

Ardington, C., Wills, G., Pretorius, E., Mohohlwane, N. & Menendez, A. (2021). Benchmarking oral reading fluency in the early grades in Nguni languages. *International Journal of Educational Development,* 84, 102433. https://doi.org/10.1016/j.ijedudev.2021.102433

Ardington, C., Mohohlwane, N., Sebaeng, L., Jodar, P., Beggs, C., Makgabo, C., Maledu, A. & Zwane, Z. (2022) *Sepedi Early Grade Reading Skills Benchmarks. Technical Report.* Pretoria: DBE.

Batibo, H.M. (1999) A lexicostatistical survey of the Setswana dialects spoken in Botswana. *South African Journal of African Languages,* 19(1), 2–11. https://doi.org/10.1080/02572117.1999.10587376

Betts, E. A. (1946) *Foundations of Reading Instruction, with Emphasis on Differentiated Guidance.* New York, NY: American Book Company.

Chard, D.J, Pikulski, J.J & McDonagh, S. (2006) Fluency: The link between decoding and comprehension for struggling readers. *Fluency Instruction: Research-based best practices.* Guilford Press, 2006. 39-61.

Constitution of the Republic of South Africa. (1996) As adopted on 8 May 1996 and amended on 11 October 1996 by the Constituent Assembly.

DBE, 2017. Summary Report: Results of Year 2 Impact Evaluation. The Early Grade Reading Study (EGRS). Department of Basic Education, South Africa. Pretoria. https://www.education.gov.za/Portals/0/Documents/Reports/EGRS%20Summary%20Report.pdf?ver=2017-08-17-090215-583

De Vos, M., Van der Merwe, K. & Van der Mescht, C. (2014) A linguistic research programme for reading in African languages to underpin CAPS. *Journal for Language Teaching,* 48(2), 149–177. https://doi.org/10.4314/jlt.v48i2.7

Demuth, K. (1993). Issues in the acquisition of the Sesotho tonal system. Journal of Child Language, 20(2), 275-301.

Guma, S. M. (1981) *An outline structure of Southern Sotho.* Pietermaritzburg: Shuter and Shooter.

Howie, S., Combrinck, C., Roux, K., Tshele, M., Mokoena, G.M. & Palane, N. (2017) PIRLS LITERACY 2016: South African Highlight Report. Pretoria: Centre for Evaluation and Assessment, University of Pretoria.

Machobane, M.E., Matlosa, L. & Mokitimi. M.I. (2003) Harmonisation of orthography in the Sotho Group of languages. In: Chebanne, A., Jokweni, M., Mokitimi, M.A. & Ngubane, S. (eds), *Unifying Southern African Languages: Harmonisation and Standardisation,* 135–142. Cape Town: CASAS.

Malda, M., Nel, C., & van de Vijver, F.J.R. (2014) *Learning and Individual Differences* 30, 34-45.

Messerschmidt, J., Ramabenyane, J., Venter, R. & Vorster, C. (2008) The facilitative role of adults in the language development of Afrikaans- and Sesotho-speaking preschool children. *European Early Childhood Research Journal,* 16(3), 283-296.

Mohohlwane, N., Wills, G. & Ardington, C. (2022) A review of recent efforts to benchmark early reading skills in South African languages. In: Spaull, N. & Pretorius, E. *Early Grade Reading.* Cape Town: Oxford University Press.

Mojela, V. M. (1999). Prestige terminology and its consequences in the development of Northern Sotho vocabulary (Doctoral dissertation, University of South Africa).

Phaahla, P. (2006) *The Feasibility of Northern Sotho as a Language of Commerce and Industry in the Limpopo and Gauteng Provinces.* Johannesburg: University of Johannesburg.

Poulos, G. & Msimang, T. (1996) A linguistic analysis of Zulu. Pretoria: Via Afrika Limited.

Ramaphosa, C. (2019) *President Cyril Ramaphosa: State of the Nation Address 2019.* 20 June 2019. South African Government. Available at: https://www.gov.za/speeches/2SONA2019

Room to Read, 2018. Setting Data-Driven Oral Reading Fluency Benchmarks Guidance Note. Research Triangle Park, NC.

RTI International. 2015. Early Grade Reading Assessment (EGRA) Toolkit, Second Edition. Washington, DC: United States Agency for International Development.

RTI International. 2017. All Children Reading-Asia: EGRA Benchmarks and Standards Research Report

Schoeneborn, A. (2008) *The Sesotho Book. A Language Manual.* Mazenod: CHE

Share, D. (2021) Is the Science of Reading Just the Science of Reading English? *Reading Research Quarterly* 56(S1), S391-S402. https://doi.org/10.1002/rrq.401

South African Government, The Presidency. (2019) The State of the Nation Address. Available at: https://www.gov.za/speeches/president-cyril-ramaphosa-2019-state-nation-address-7-feb-2019-0000

Spaull, N. & Pretorius, E. (2019) Still Falling at the First Hurdle: Examining Early Grade Reading in South Africa. In: Spaull N., Jansen J. (eds) *South African Schooling: The Enigma of Inequality. Policy Implications of Research in Education.* Springer, Cham. https://doi.org/10.1007/978-3-030-18811-5_8

Stanovich, KE. 2000. Progress in Understanding reading: Scientific foundations and new frontiers. New York: The Guildford Press.

Statistics South Africa. (2012) Census 2011 Statistical release – P0301.4. Pretoria: Statistics South Africa.

Spaull, N., Pretorius, E. & Mohohlwane, N. (2020). Investigating the comprehension iceberg: Developing empirical benchmarks for early-grade reading in agglutinating African languages. *South African Journal of Childhood Education,* 10(1), a773.

Taljard, E., & Bosch, S. E. (2006). A comparison of approaches to word class tagging: Disjunctively vs. conjunctively written Bantu languages. *Nordic journal of African studies,* 15(4).

Wang, Z., Sabatini, J., O'Reilly, T., & Weeks, J. (2019) Decoding and Reading Comprehension: A Test of the Decoding Threshold Hypothesis. *Journal of Educational Psychology,* 111(3), 387–401. https://doi.org/10.1037/edu0000302

Wills, G., Ardington, C. & Sebaeng, M.L (2022) Foundational skills in home language reading in South Africa: Empirical evidence from 2015-2021. In: Spaull N. & Pretorius E. (eds) *Early Grade Reading in South Africa.* Cape Town: Oxford University Press.

Wills, G., Ardington, C., Pretorius, E., Pooe, E. & Ramagoshi, R. (2022) Benchmarking Early Grade Reading Skills: Setswana and English First Additional Language. Technical Report. Johannesburg: Khulisa Management Services.

Wilsenach, C. (2019) Phonological awareness and reading in Northern Sotho – Understanding the contribution of phonemes and syllables in Grade 3 reading attainment. *South African Journal of Childhood Education.* 9(1): 1-10.

9. Endnotes

1 The Global Alliance to Monitor Learning describes meeting this benchmark as follows: “Students independently and fluently read simple, short narrative and expository texts. They locate explicitly stated information. They interpret and give some explanations about the key ideas in these texts. They provide simple, personal opinions or judgements about the information, events and characters in a text[i].” Reading proficiency is not only easily understood but a lack thereof is also “usually a clear indication that school systems are not well organised to help children learn in other areas such as maths, science and the humanities[ii]”.

2 A review of recent evidence supports the continued use of these levels (Allington, McCuiston & Billen 2015).

3 The term Bantu is not only linguistic, objectified almost immediately and used as an ethnic label for ethnographic purposes. The term remains controversial due to its politicized nature. However, the linguistic label remains official.

4 These tonal marks are not typically represented in orthography. The tone is derived from the context.

5 The levels developed by Betts should be easily attainable for learners reading in the Sesotho-Setswana language as accuracy tends to develop more readily in transparent languages than in English.

6 The Sesotho sample includes only 12 schools, all of which were receiving the Literacy Boost intervention. These schools are possibly less informative about average learning outcomes in Sesotho LOLT schools than the Sestwana and Sepedi samples. Within the 12 schools, there is considerable variation in reading proficiency and the data are useful in demonstrating that the fluency-comprehension relationship in Sesotho is very similar to that of the other languages in the family.

Published by the Department of Basic Education

222 Struben Street

Private Bag X895, Pretoria, 0001

Telephone: 012 357 3000 Fax: 012 323 0601

ISBN Number: 978-1-4315-3412-8

website: www.education.gov.za

facebook: www.facebook.com/BasicEd

twitter: www.twitter.com/dbe_sa

basic education
Department:
Basic Education
REPUBLIC OF SOUTH AFRICA

ZENEX FOUNDATION

Funda Wande
Reading for Meaning

Room to Read

SALDRU
Southern Africa Labour and Development Research Unit

University of Cape Town · Iyunivesithi yaseKapa · Universiteit van Kaapstad

RESEP
Research on Socio-Economic Policy

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

UNIVERSITY OF LIMPOPO

UNISA
university of south africa

UFS·UV
UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA

NORTH-WEST UNIVERSITY
YUNIBESITI YA BOKONE-BOPHIRIMA
NOORDWES-UNIVERSITEIT