February
2018

RESEP WORKING PAPER

Investigating the Comprehension Iceberg: *Developing empirical benchmarks for early grade reading in agglutinating African languages*

Nic Spaull (Stellenbosch University),

Elizabeth Pretorius (UNISA),

Nompumelelo Mohohlwane (Stellenbosch University)

RESEP Working Paper Series No. WP01/2018

Keywords: reading benchmarks, African languages, EGRA, literacy, South Africa

RESEP
Research on Socio-Economic Policy

UNIVERSITEIT STELLENBOSCH UNIVERSITY

Physical Address
Research on Socio-Economic Policy (RESEP);
Department of Economics, Matieland;
7602; South Africa.

The RESEP Working Paper series
(https://www.resep……)
is a sister publication of the Stellenbosch Economic Working Paper series
(https://www.ekon.sun.ac.za/wpapers)

Investigating the Comprehension Iceberg: *Developing empirical benchmarks for early grade reading in agglutinating African languages*

Authors: Nic Spaull (Stellenbosch University), Elizabeth Pretorius (UNISA), & Nompumelelo Mohohlwane (Stellenbosch University)

Abstract

The importance of learning to read in mother-tongue is widely acknowledged in the linguistics literature yet reading acquisition in African languages remains under-researched and under-theorized. While numerous studies have highlighted the low levels of comprehension among learners reading in African languages in South Africa, little has been done to understand what lies beneath this 'comprehension iceberg.' In this paper we present new empirical evidence on reading outcomes and the sub-components of reading for 785 Grade 3 learners across three languages (Northern Sotho, Xitsonga and isiZulu), drawn from 61 primary schools in South Africa. This is the largest sample of such learners to date. Using an adapted EGRA-type assessment we assessed letter-sounds, single-word reading, non-word reading, oral reading fluency and oral comprehension. From this data we present results on fluency, accuracy and comprehension and how these might relate to each other in these morphologically rich agglutinating languages. We also show that there are large differences in reading sub-components between languages with conjunctive and disjunctive orthographies. Our results suggest that there are minimum thresholds of accuracy and oral reading fluency in each language, below which it is virtually impossible to read for meaning. These are 52-66 wcpm in Northern Sotho, 39-48 wcpm in Xitsonga and 20-32 wcpm in isiZulu. We argue that there is a strong need for empirical language-specific norms and benchmarks for indigenous African languages and present our benchmarks for these three languages as a move in that direction.

Abbreviations: ORF - Oral reading fluency; wcpm – words correct per minute

Funding: The authors would like to acknowledge funding from the Economic and Social Research Council (ESRC) as part of the Raising Learning Outcomes (RLO) initiative (Grant: ES/N01023X/1).

Introduction

Given the important role that reading plays in scholastic performance, it is important to launch children on successful reading trajectories from the start of schooling. The Progress in International Reading Literacy Study (PIRLS) assesses reading comprehension internationally at the Grade 4 level, by which time children have already been launched on their reading trajectories during the first three years of schooling. The PIRLS and prePIRLS data from 2006 and 2011 indicate that Grade 4 children in South Africa perform very poorly in reading comprehension, particularly when reading in their African home language. More than half of Grade 4 learners have not learned to read for meaning in any language by Grade 4 (Spaull, 2016). While such outcomes clearly signal challenges within the education system regarding comprehension and the need for learners to develop meaning making skills in the written form, they also raise questions about the development of early reading skills, and how these enable comprehension, particularly in the African languages. To remedy this situation, we need a clear idea of what a successful reading trajectory looks like, what factors underpin its success, and how it is similar or different across languages.

Decades of research into reading in English - probably the most widely researched language in the world – has provided education stakeholders with an evidence-based framework for profiling what successful reading in English looks like (Adams, 1990; National Reading Panel, 2000). For example, by the end of Grade 3 children at the 50th percentile can on average read 107 words correct per minute (wcpm) in English (Hasbrouck & Tindal, 2006), while children reading slower than 40 wcpm at the end of Grade 1 are considered to be at reading risk. Notwithstanding the importance of this contribution to our general understanding of reading in alphabetic languages, identifying what is universal and what is language specific in early reading development calls for a research base that includes alphabetic languages that are typologically different and have different orthographic systems. 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. What would an average Grade 3 or an at risk Grade 1 reader in an African language look like? Very little reading research has been done in these languages. Currently anecdotal experience, intuitions and linguistic hunches tend to underlie educational judgements about how young African language readers are faring. In many cases, teachers are poorly trained and do little reading themselves (Pretorius & Knoetze, 2012).

To its credit, South Africa has prioritised the large-scale measurement and monitoring of reading comprehension outcomes across the country[1]. While there are several nuances in the successive results of the large-scale comprehension assessments undertaken in South Africa, what is lacking is not accurate information on reading outcomes but accurate information on what is less visible beneath the comprehension iceberg. As De Vos, van der Merwe, and van der Mescht (2014, p. 168) point out, very

[1] These include PIRLS, the Southern and East African Consortium for Monitoring Education Quality (SACMEQ), and the Annual National Assessments (ANA) undertaken nationwide by the Department of Basic Education.

little has been done on the 'cognitive-linguistic processes involved in reading in African languages'. A strong empirical base is needed to gain insight into early reading development in African languages and make sound judgments about ways to reduce the literacy inequalities within the education system.

Given the paucity of research on decoding in African languages, this article uses Grade 3 reading data from three African languages in South Africa and examines the role of alphabetic knowledge, word reading and oral reading fluency in early reading success in these languages. Before turning to the research itself, we first identify three attributes of early reading in alphabetic languages, briefly outline ways in which African languages differ from English and the implications this may have for reading, and then we look at the role of alphabetic knowledge, word reading and oral reading fluency in early reading development.

Early reading development in alphabetic languages

The first three years of schooling are typically dedicated to laying a sound foundation for the development of numeracy and literacy skills on which all subsequent schooling depends. By the end of Grade 3 readers are expected to read accurately, rapidly and with comprehension. Why are these three attributes regarded as desirable reading outcomes?

- ***Comprehension is the sine qua non of reading***. Reading is a form of communication; we read to comprehend the information in the written text. The aim of reading instruction is for children to understand what the written alphabetic code conveys in any text.
- ***Accuracy supports comprehension***. The ability to identify letters and read words accurately reduces comprehension complications (Adams, 1994; Spear-Swerling, 2006), e.g. it is important to distinguish *three* from *tree* in English, or *bafunda* 'they read' from *bafundile* 'they have read' in isiZulu.
- ***Speed matters in cognitive-linguistic processing, and hence in reading*.** A difference of a few milliseconds can signal difficulty or success in cognitive functioning. Processing speed tends to be strongly associated with word reading and reading comprehension (Fuchs et al. 2001; Wolf & Katzir-Cohen, 2001). The more effort expended on processing the alphabetic code and words, the less attentional capacity there is for comprehension.

Research into the acquisition of literacy shows that individual differences between learners in accuracy, speed and comprehension can emerge early and can persist throughout their schooling, impacting negatively on their scholastic performance (Spear-Swerling, 2006). If some children find reading effortful and frustrating, they will not perceive it as meaningful or pleasurable, and be less inclined to actively engage in it. The relationship between accuracy, speed and comprehension may play out in different ways in languages with different typologies or orthographies. Before looking at research on early reading, we digress now for a brief overview of agglutinating African languages.

Typological and orthographic features of agglutinating African languages

This section highlights some features that distinguish agglutinating African languages and their orthographies from English, and identifies in what ways these features might impact on early reading development.

Agglutinating languages: morphological complexity

The nine African languages spoken in South Africa belong to the family of Southern African Bantu languages. In terms of linguistic typology, they are all agglutinating languages with a complex morphology whereby prefixes, infixes and suffixes are added to noun and verb stems. The verbal elements in a sentence are especially complex, marking tense, aspect and mood and added as infixes and suffixes.

Other agglutinating languages are Finnish, Turkish and Basque. Morphological complexity is a distinctive feature of all these languages, and a single orthographic word with a stem and morphemes stacked onto it can represent a whole sentence. For example, the word *Andizithandi* in isiXhosa ('I don't like them') has the stem -*thand*- 'like' with the separate morphemes *a-ndi-zi* and -*i* attached.

Transparent orthography

Orthography is transparent in all nine African languages – letters represent specific sounds in a one-to-one mapping relationship. This is unlike English with its opaque orthography, where one letter can represent different sounds (a is sounded differently in *car, call, cane, alone*), or where the same sound can be represented by different letters (/f/ can be written as f, ph, or -gh in *frog, phone* and *cough)*. Seidenberg (2017) points out that that languages with complex morphological systems, all have transparent orthographies; an inconsistent orthography would make reading 'intolerable' (p.136) in agglutinating languages.

Although the orthography is transparent, a distinction is made between conjunctive and disjunctive orthographies. This distinction coincides with language family groupings. Within the Southern Bantu language family, the nine South African languages are further divided into the Nguni (isiZulu, isiXhosa, Siswati and isiNdebele) and Sotho (Northern Sotho, Southern Sotho and Setswana) subfamilies, and two smaller minority subfamilies (Tshivenda and Xitsonga, related to languages in Zimbabwe and Mozambique), as shown in Figure 1 below. The reading data presented in this article was collected from isiZulu (n=514), Northern Sotho (also called Sepedi) (n=143) and Xitsonga (n=128) Grade 3 readers, and thus reflect the three main linguistic subgroups, as highlighted below.

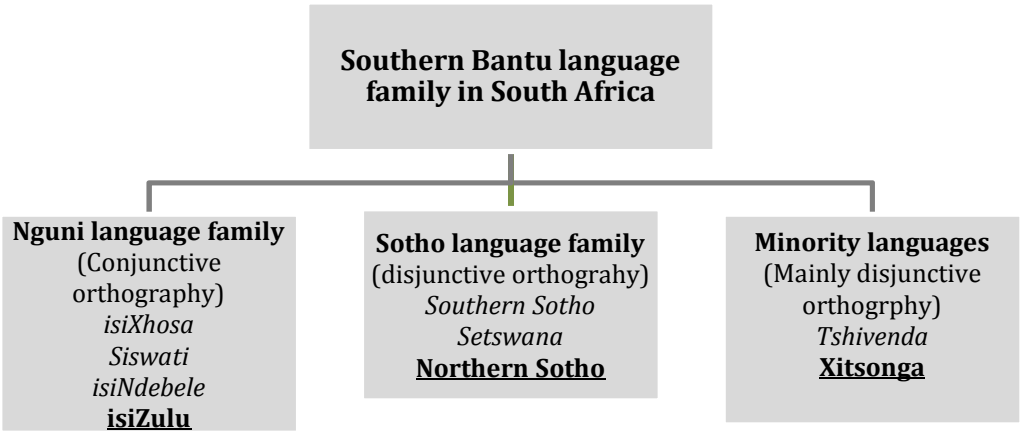

Figure 1. The Southern Bantu language families in South Africa

During the 19th century, the work of codifying these languages was initially undertaken mainly by missionaries, with training in different philological schools.

Morphophonological features specific to the different African languages (e.g. vowel elision in the Nguni languages) resulted in the development of different transparent orthographies for these languages. For example, the Nguni languages have a conjunctive orthography, where nominal and verbal elements in a sentence tend to be written together as single orthographic 'words'. In contrast, the Sotho languages evolved a disjunctive orthography, where some of the verbal elements in a sentence (e.g. noun class markers and suffixes) are written separately. For example, 'They used to read it' is written conjunctively as a single orthographic word *Bebayifunda* in isiZulu, while it is written disjunctively as three separate words *Ne ba ebala* in Northern Sotho. Xitsonga orthography is somewhere in between, having elements of both conjunctive and disjunctive orthography. The conjunctive/disjunctive distinction has implications for early reading, measurement and benchmarks.

Conjunctive orthography gives rise to long word units which create 'dense' texts; conversely, disjunctive orthography results in much shorter word units (often single syllables comprising V or CV). Because of its conjunctive orthography, there are typically few free morphemes in an Nguni language sentence – bound morphemes by way of prefixes, infixes and suffixes are added to noun and verb stems. Single syllable words are practically non-existent (they are mainly exclamations) and two syllable words are not common in the conjunctive orthography. Because of the noun class prefix attached to a noun stem, nouns typically contain three or more syllables. In terms of text length, equivalent texts translated into the conjunctive Nguni texts will yield short texts with long words, while the same text in a disjunctive Sotho language will yield longer texts with many short words. To illustrate these orthographic differences, examples taken from the first three sentences in a Grade 3 reader, in isiZulu, Northern Sotho and Xitsonga respectively are given in Table 1 below.

Table 1. Words per sentence in conjunctive/disjunctive orthographies

Language	**Text**						
N Sotho	*Ka le lengwe la matšatši mosepedi yo a bego a na le tlala. O fihlile motseng wo mongwe a kgopela dijo. Go be go se na yo a bego a na le dijo.*						
Xitsonga	*Siku rin'wana mufambi loyi a ri na ndlala. U fikile emugangeni. A kombela swakudya, kambe a ku nga ri na loyi.*						
isiZulu	*Kunesihambi esasilambile kakhulu. Sahamba sicela ukudla emizini yabantu. Abantu abengenakho ukudla.*						
Gloss	There was a stranger who was very hungry. He came to a village and asked for food. Nobody had any food.						
	Words in Sentence1	Words in Sentence2	Words in Sentence3	Total words	Words per sentence	Letters per word	Total single syllable words: V/ CV
N Sotho	13	8	12	33	11	3.2	21
Tsonga	8	3	10	21	7	4	9
Zulu	3	5	3	11	3.6	8	0

As can be seen, the same three sentences yield texts with different profiles. The three sentences in isiZulu comprise a total of 11 words only, but these are long words (average of 8 letters per word), averaging 3.6 words per sentence. The shortest word in the isiZulu text has three syllables. In contrast, the same three sentences comprise 33 words in Northern Sotho; these are mainly shorter words (average of 3.2 letters per word), averaging 11 words per sentence. The Xitsonga text profile is in between: the three sentences comprise 21 words, with 7 words per sentence and 4 letters on average per word. There are 21 single syllable words in the Northern Sotho text, 9 in the Xitsonga text and none in the isiZulu text.

The agglutinating nature of African languages, their complex consonants and the conjunctive/disjunctive orthographies may have important implications for reading development in these languages.

Foundational reading skills

In order to optimise reading instruction and to look out for those who fall behind their grade peers, it is important to understand the dynamics of how the different components of decoding and comprehension interact and mesh, and where and why reading fallout happens. Skills that are key to learning to read the alphabetic code are foregrounded in the initial stages of learning to read and may predict early reading skill in Grades 1 or 2. When mastery is achieved, these skills become automatised and so recede to the background, while qualitatively different processes and skills become foregrounded and push reading development to another level (Adams 1990; Spear-Swerling, 2006). The ways in which these components interact may be sensitive to specific linguistic and orthographic constraints associated with different languages that share the same alphabetic code.

Alphabetic knowledge

Alphabetic knowledge refers to knowledge that written symbols stand for the phonemes of spoken language. Inability to grasp this principle negatively affects the development of decoding (Nieto, 2005).

During their first year of formal schooling children become acquainted with different aspects of letters, their names, shape in lowercase and uppercase, the sounds they represent, and later too, how their shape changes in different fonts and writing styles.

Letter-sound knowledge is also related to phonological awareness, especially at the phonemic level. Phonological awareness has been found to be important in learning to read across alphabetic languages. It follows a large-to-small developmental path: awareness of larger units such as words, rhymes and syllables occurs in preschool, while developing an awareness of the smallest unit, the phoneme, usually happens with formal reading instruction. When children learn letter-sound relationships, they develop an awareness of individual sounds within words (Ziegler & Goswami, 2005). Vihman (1996) argues that alphabetic knowledge enables phonological representations to become more precise and that letter-sound knowledge is thus predictive of phonemic awareness. However, some researchers regard the relationship to be reciprocal (e.g. Perfetti, Beck, Bell & Hughes, 1987).

Letter-sound knowledge is a critical foundational skill of early literacy acquisition (e.g. Muter & Diethelm 2001) and becomes the main processing stage in word reading, where children use their letter-sound knowledge to sound out new words not previously

encountered. Blaiklock (2004) suggests that the role between phonological awareness and reading development is mediated by letter knowledge.

Because of its strong link to early reading instruction, alphabetic knowledge seems to have a narrow developmental window (Ouelette & Haly, 2013). Using measures of alphabetic knowledge with preschool children can lead to floor effects (Burgess & Lonigan, 1998), while using it with older learners can produce ceiling effects (Wise, Sevcik, Morris, Lovett & Wolf, 2007). However, given the slow rate of reading development and the low literacy levels that usually obtain in developing countries, assessing alphabetic knowledge with older learners may help to distinguish readers from non-readers who have not yet grasped the relationship between print and sound.

Word and non-word reading

The most basic task of reading is being able to process the meaning of individual words from print and construct the overall meaning of the text in which the words occur. Although the ability to read words quickly and accurately is one aspect of reading, Adams argues that unless word reading operates properly, "nothing else in the system can either" (1994:838). In alphabetic scripts, this is not possible without initial letter-sound knowledge (Adams, 1994; Share, 1995). However, to build fluency children need to become aware of recurring letter patterns in their own language, based on morphological and orthographic information, incorporating smaller and larger word chunks until full word recognition is reached (Ehri, 2005; Share, 1995). After several encounters, words become known and familiar, readers recognise word chunks and so build up word-specific knowledge (Kilpatrick, 2015) which helps to speed up and automatise the reading process so that attention is freed up for comprehension.

There has been a long history of word reading research and its relationship to reading development in general and reading comprehension specifically. There is a strong association between speed and accuracy of word reading and reading comprehension (Adams, 1990; Stanovich 1986).

Assessing children's word reading ability is a good way to assess their decoding ability. Context free word reading by way of word lists containing increasingly longer and more complex words is a significant predictor of reading (Jenkins, Fuchs, et al. 2003). The use of non-words is also a commonly used to assess decoding ability. Non-words meet the phonological criteria of a language but don't exist, e.g. *brillig, slithy, toves* in English. Because these words lack meaning and readers have no orthographic representations of such words, non-words eliminate lexical processing and reveal a reader's phonological recoding ability. Research shows that real words are processed faster and more accurately than non-words. This seems to apply not only in opaque orthographies but also in transparent agglutinating languages such as Turkish (Miller, Kargin & Guldenoglu , 2014).

Because of its opaque orthography, and high occurrence of common, short words, many of which are not conventionally decodable (e.g. *are, could, there*), English readers need to build up a sight vocabulary of high frequency words that they can recognise quickly and accurately. Research suggests that this process takes longer in English than in languages with transparent orthographies, where rapid and accurate word reading can be achieved far more quickly, as in Greek, Welsh, German and Spanish, where letter-sound mapping occurs without much difficulty because of its regularity, and children can become efficient decoders within a year or so (Ellis & Hooper, 2001; Wimmer, 2003;

Ziegler & Goswami, 2005). This has also been found in agglutinating languages such as Turkish (Öney & Durgunogu, 1997; Babayağit & Stainthorp, 2007). In their study of differences in reading long, inflected words in Basque (an agglutinating language) Acha, Laka & Perea (2010) found that while Grade 3 children relied mainly on letter-sound decoding, word identification was faster and more efficient with Grade 6 readers, who besides phonological decoding seemed also to rely on basic orthographic and inflectional patterns in the language as they became exposed to less frequent words during reading.

Oral reading fluency

Oral reading fluency (ORF) reflects the speed, accuracy and naturalness that readers display when reading a text aloud, following the intonation and rhythm of spoken language. ORF is seen as a general indictor of reading competence (National Reading Panel, 2000). Because intonation is more difficult and subjective to assess, speed and accuracy form the main focus of ORF assessment. Typically, readers are given a text to read within a minute, with errors subtracted from the total number of words read in a minute, giving a score of words correct per minute (wcpm). To control for decoding without understanding a short oral reading comprehension follows.

Research shows a strong association between ORF and reading comprehension (Spear-Swerling, 2006; Fuchs et al. 2001) despite differences in socioeconomic status and instructional programmes; it occurs in children without reading difficulties as well as those with learning disabilities (Wolf & Katzir-Cohen, 2001). It has also been found in second language reading (Al Otaiba et al., 2009; Jimerson, Hong, Stage & Gerber, 2013), also in South Africa, the country of analysis here (Draper & Spaull, 2015; Pretorius & Spaull, 2016).

The greatest growth in ORF seems to occur in the early school years, between Grades 1-4. ORF is useful for measuring small increases in improvement, unlike many other standard measures of performance which can only detect large changes in the outcome (Blachowicz, Moskal et al. 2006). Typically, from Grade 4 onwards the effects of ORF start to level off (Fuchs et al., 2001; Spear-Swerling 2006). Once reading becomes relatively fast and accurate, other variables account for differences in reading comprehension, such as vocabulary knowledge, inferencing abilities and text, genre and background knowledge.

ORF norms have been established for English readers, showing how children at different grades and at different percentiles typically perform. However, very little research has been done on ORF in the African languages. For example, if Mpumi in Grade 3 reads at 28 wcpm in isiZulu we currently have little empirical evidence of whether or not she is a good reader.

Research on early reading development in African languages

Approximately 70% of children in South Africa complete the first three years of schooling in their home language (an African language) with English taught as an additional language (Pretorius & Spaull, 2016: 1450). The situation then flips from Grade 4 onwards, with 90% of learners now learning in English, with African languages taught as a home language subject. Since these learners need to be not only bilingual but also biliterate, much of the research on early reading thus focuses on reading in two languages.

There are currently not many studies on early reading in African languages and a rather uneven picture emerges from them as not all studies focus on the same factors, use the same measures, or use similar measures in the same way (e.g. some studies use timed word reading measures, other do not). Research findings from the Nguni (isiZulu and isiXhosa) and Sotho (Northern Sotho and Setswana) languages are available, but often come from small scale studies, and as yet no research seems to have been done in Xitsonga.

Letter knowledge: Because there are many consonants in African languages, with many digraphs (hl, ph, tj), trigraphs (ngw, kgw) and also 4-letter consonant sequences (mpfh, ntlh), it is important that children learning to read in African languages master these consonants. Children learning to read in African languages need to be able to distinguish between the different letter shapes, their sounds and combinations in order to get on with the task of learning to read words that combine single consonants, digraphs and trigraphs. Surprisingly, however, only a few studies have included measures of alphabetic knowledge in their assessment of early reading skills in African languages. These include Swahili (Alcock, Ngorosho, Deus & Jukes, 2010), Setswana (Lekgoko & Winskel, 2008) and Northern Sotho (Wilsenach, 2015), but the sample sizes are relatively small ranging from 36-108 learners. Although these studies show a relationship between letter-sound knowledge and early literacy in African languages, the relationship has not yet been examined closely.

Word reading and ORF: Results on word reading and ORF in both Nguni and Sotho languages can be gleaned from a few studies. In a small study of Grade 4 isiZulu learners (n=44) Pretorius (2015) found that only half (53%) of the words could be read correctly with a mean ORF score of 19 wcpm, indicating very slow reading in isiZulu. There was a strong correlation between word reading and ORF (r=.79). The findings suggested that Grade 4 children had not yet mastered phonics principles in isiZulu.

Another Nguni language, isiXhosa, was studied by Diemer (2015) and Rees (2016) at Grade 3 level, with Diemer focussing on phonological awareness and Rees on morphological awareness. In Diemer (2015) the Grade 3 (n=55) ORF mean was 19 wcpm, and the comprehension mean 23%. Despite the low and slow literacy levels, speed and accuracy increased together in the ORF scores, and a strong correlation of .69 was found between ORF and comprehension. However, in Rees' study (2016) of Grade 3s (n=74), a lower correlation of .46 was found between ORF and comprehension.

In the Sotho language group with its disjunctive orthography, Wilsenach (2013, 2015) looked at features of early reading of Grade 3 bilingual Northern-Sotho/English learners, half of whom had Northern Sotho as the language of learning in the first three years of schooling, while for the other half, early reading instruction had been in English. The ORF scores vary across the cohorts. Although in the 2016 study 67% of the words were read accurately, reading was very slow. Like Wilsenach, Maukare (2017) also looked at Grade 3 bilingual Northern Sotho/English readers (n=98). Although the untimed word reading scores of the Northern Sotho children showed 79% accuracy, the children read slowly, averaging 35 wcpm. Here too, performance on single word and text word reading was highly correlated (r= .78). Readers interested in Setswana and Herero are directed to Malda, Nel and Vijver (2014) and Veii & Everatt (2005) respectively for similar findings. These are not discussed here due to space constraints.

It is clear that while interest in early reading in African languages is emerging, there are still many issues that need to be further researched.

- There are surprisingly few studies that directly examine the role of alphabetic knowledge in African language reading.
- Although English reading research shows strong correlations between word reading and ORF measures with comprehension, in early African language reading, the relationship varies from mild to robust.
- Despite their transparency, conjunctive/disjunctive orthographies seem to affect early reading rates differentially. The reading rates from the Nguni studies are uniformly slow, while the reading rates from the Sotho languages are relatively faster. However, there is as yet no clear picture of the range of performance at the different percentiles within the different languages.
- Nearly all the studies reviewed involve fairly small sample sizes from a small number of schools (never more than 4-5 schools), so generalisation is constrained. A much larger and more varied empirical base is needed for theory building and for benchmarking.

This article looks at Grade 3 reading data from 61 schools across three provinces in South Africa, representing both conjunctive and disjunctive transparent orthographies. There are two main aims: (1) an analysis of the relationship between letter-sound knowledge, word and non-word reading, ORF and oral reading comprehension, and (2) on the basis of these relationships, determining minimum thresholds of accuracy and fluency within the three different language groups, below which comprehension is compromised.

Background to the study

The data presented in this article draws on a larger study formally known as the "*Leadership for Literacy*" project. The schools selected for the study are typical of those which serve the majority of learners and come from three South African provinces – Gauteng, KwaZulu-Natal, and Limpopo. Of the 61 schools in the study 56 are from the poorest 60% of schools in the country (Quintile 1-3) which are no-fee schools, and 5 are from Quintile 4 where some charge relatively low fees (<R3000/year). The aim of the sampling process was to ensure that there was the full range of performance across Quintile 1-3 schools in these provinces. To that end we sampled 29 of the highest achieving quintile 1-3 schools in these three provinces and 29 'matched' typical schools (typical schools were all low-achieving schools). Schools were selected based on their achievement in the Annual National Assessments (ANA) across four years (2011-2014). The ANAs are a yearly national assessment at the primary school level in South Africa. All allegedly higher-performing schools were validated by at least two recommendations from experts in the provincial departments of education, non-governmental organisations and/or other researchers. The five non-Quintile 1-3 schools were included towards the end of the project in order to ensure that there were some learners reading at higher levels. This was because even the highest achieving Quintile 1-3 schools were found to be achieving at very low levels. In total there were 21 schools from Gauteng, 21 in KwaZulu-Natal, and 19 in Limpopo.

The data used in this article was collected between February-March 2017 in all three provinces, with three fieldworkers per team. The tests were administered one-on-one by the fieldworker, with information captured electronically on tablets using Tangerine, an open source software programme primarily designed for Early Grade Reading Assessments. Each test was designed to be completed within 15 minutes. In all, 785 Grade 3 learners were assessed: 514 in isiZulu, 143 in Northern Sotho and 128 in

Xitsonga. All fieldworkers were native speakers of the language they were assessing, held at least a bachelor's degree or 3 year diploma and received three-days of intensive training.

Grade 3 reading assessment

The Grade 3 learner assessment was an adapted form of the the Early Grade Reading Assessment (EGRA) that already existed for these three African languages[2].

Each Home Language assessment consisted of five subtests: (1) a timed letter-sound subtest containing rows of letters that learners must sound aloud; (2) a timed word subtest, consisting of a list of words that learners must read out aloud; (3) a timed non-word subtest; (4) reading the title of the ORF passage story, and (5) an ORF passage read aloud within one minute. Following the ORF subtest, learners were asked oral reading comprehension questions, based on the passage. Various opt-out rules were applied in the various subtests to protect learners who could not read at all, as part of the ethical practices of the study.

In each of the assessed languages, the letter sound section had 110 items. In addition to the standard EGRA test of lowercase and uppercase letters, this subtest was adapted to include digraphs, trigraphs and 4-5 letter phonemes. The isiZulu subtest included 27 digraphs such as "ng" and 6 trigraphs such as "ncw" and "nhl". The Northern Sotho subtest included 23 digrahs, 6 trigraphs and a 4-letter phoneme. The Xitsonga subtest included 21 digraphs such as "dy" and "hl"; 8 trigraphs such as "mbh" and "mpf "and a 4 letter phoneme.

Across the three languages, for both the word and the non-word reading tasks, there were 60 words per task, with the words ranging from two to seven syllables, starting with shorter words and ending with longer words (e.g. from *ikati* to *intothoviyane* in isiZulu; from *pula* to *kanagelokopana* in Northern Sotho; from *teka* to *mpfampfarhuta* in Xitsonga). In order to keep the word tests comparable across the three African languages, no single syllable function words that are common in the disjunctive Sotho orthographies were included in the Northern Sotho word lists (e.g. *a, na, go, le* etc., as shown in Table 1 earlier). The words in all three languages were nouns or infinitive forms of verbs.

The ORF passage was 'Rock Soup', a narrative text from a South African graded-reading series (Vula Bula) and was already available in each the three African languages. Although this was the same story, given the conjunctive/disjunctive features of the three languages, there were 120 words in the Northern Sotho passage, 105 in the Xitsonga passage and 67 in the isiZulu passage.

Data results and analysis

Table 3 reports a range of descriptive statistics for each of the EGRA subtasks by language group, including the number of learners in the sample, the 10th, 25th, 50th, 75th and 90th percentiles of the distribution as well as the minimum, mean, maximum and standard deviation (SD). Some of the notable findings are listed below:

[2] RTI International, together with reading experts, developed the Early Grade Reading Assessment known as EGRA, funded by USAID, the World Bank and other donors. It is composed of subtasks designed to systematically assess foundational reading skills in the early grades in low-income countries. It is increasingly being used in developing countries to monitor early reading development.

- *Letters correct per minute*: On the whole, letter-sound knowledge was low. Of the 740 learners assessed, only a quarter of learners could name at least 40 letter-sounds correctly per minute. Across all languages, 25% could only sound out at most 15 letters correctly in one minute.
- *Word reading:* Word reading (which excluded single-syllable words), irrespective of orthography, was fairly similar across the three languages, ranging from 22 wcpm in Northern Sotho to 19 wcpm in isiZulu. Interestingly, when single syllable function words typical of the disjunctive orthographies are excluded from a word reading list, then learners in Northern Sotho and Xitsonga read at similar rates as learners in isiZulu. Predictably, reading non-words was slower than reading words.
- *Oral reading fluency*: The ORF scores in isiZulu (21wcpm at the 50th percentile) were considerably lower than those in Northern Sotho (41wcpm) and Tsonga (47wcpm). The longer words in written isiZulu texts results in slower reading rates. The occurrence of several short grammatical morphemes that are written separately in the more disjunctive orthographies of Northern Sotho and Xitsonga result in faster reading rates in ORF passages in those languages.
- *Oral reading comprehension:* Reading comprehension was generally low. As seen in the analysis below, reading comprehension was a function of reading speed and accuracy.

Table 4 shows correlations between various subcomponents of reading across the three African languages.

Table 4. Correlations between subcomponents of reading

Correlations *r*	Northern Sotho	Xitsonga	isiZulu
letter-sound x word reading	.74	.76	.60
letter sound x nonword reading	.69	.75	.58
letter-sound x ORF	.68	.75	.55
word reading x nonword reading	.91	.92	.91
word reading x ORF	.92	.92	.91
ORF x comprehension	.87	.78	.81

The results show robust and significant correlations between all the subcomponents of reading. Knowledge of letter-sounds is strongly associated with ability to read words and nonwords, as well as with oral reading fluency, although to a lesser degree in the conjunctive reading of isiZulu. Oral reading fluency and comprehension also show a strong relationship. These relationships can clearly be seen in the box plots in Figure 2, showing increasing skill across the deciles. The analysis below provides a more nuanced view of skill in these subtasks.

Fluency and accuracy

Table 5 below shows the mean for letter sounds attempted and the percentage of letters sounded incorrectly. It would seem that while those learners in ORF Decile-1 make more errors than those in the higher ORF Deciles, almost the entire sample read 15-20% of the letter-sounds attempted incorrectly. This low level of letter-sound knowledge and accuracy might be a reflection of early reading instructional practices, where teachers may not be spending enough time on systematic phonics instruction, especially of the complex consonant system. This result may also reflect lower levels of accuracy in

letter-sound reading than in word reading, where words provide a context for the letter sounds. Perhaps most importantly those who read the ORF passage at very slow rates (0 wcpm or 1-10 wcpm), also have exceedingly high rates of inaccuracy, making mistakes on every second letter-sound attempted. This is not an insignificant percentage of the sample, accounting for 27% of all learners (202/740). If learners are as likely to get letter sounds right as wrong, it will be almost impossible for them to read words or connected text with understanding.

Table 5. *Mean letter sounds attempted and percentage correct by decile of ORF Words Correct Per Minute*

	Northern Sotho			Xitsonga			isiZulu		
	Letters attempted	% incorrect	Sample	Letters attempted	% incorrect	Sample	Letters attempted	% incorrect	Sample
0 wcpm	11	46%	24	10	47%	22	11	55%	101
Decile 1 (1→10)	21	41%	9	7	60%	3	20	38%	43
Decile 2 (11→20)	26	26%	11	21	25%	4	26	29%	95
Decile 3 (21→30)	27	28%	9	32	14%	4	29	25%	104
Decile 4 (31→40)	35	18%	13	34	12%	10	34	19%	97
Decile 5 (41→50)	37	19%	26	39	16%	26	42	15%	46
Decile 6 (51→60)	41	19%	18	44	13%	21	40	18%	6
Decile 7 (61→70)	37	16%	17	43	23%	9	43	18%	2
Decile 8 (71→80)	36	11%	4	57	13%	7			
Decile 9 (81→90)	48	7%	2	56	16%	4			
Decile 10 (91→100)	43	7%	2	53	43%	1			

Table 6 below provides the same information but for the mean number of words attempted by learners in the ORF task, as well as the percentage of words read incorrectly. This is reported for deciles of Words Read Correctly Per Minute (wcpm) in the ORF passage. For example it shows that the 9 Northern Sotho learners in Decile-1 (reading at 0→10 wcpm) actually attempted 16 words on average but read half (52%) of these words incorrectly. Across all three language groups, faster readers are more accurate than slower readers. Comparison across the languages shows that accuracy seems to be more important for fluent reading in isiZulu than in Northern Sotho or Xitsonga. The isiZulu learners reading at 21 wcpm or faster are reading with 95% accuracy or higher. In contrast, 95% accuracy is only associated with reading at 51 wcpm or faster in Northern Sotho and 31 wcpm or faster in Xitsonga. One of the reasons why Decile-1 learners are reading so slowly is that they are making mistakes on every second or third word. The fastest Northern Sotho readers (wcpm=107) and Xitsonga readers (wcpm=91) in the sample made no mistakes whatsoever.

Table 6. *Mean oral reading fluency (ORF) words attempted and percentage correct by decile of ORF words correct per minute*

	Northern Sotho			Xitsonga			isiZulu		
Words Correct Per Minute	Words attempted	% incorrect	Sample	Words attempted	% incorrect	Sample	Words attempted	% incorrect	Sample
Decile 1 (1→10)	16	52%	9	6	39%	3	9	36%	43
Decile 2 (11→20)	21	26%	11	19	24%	4	18	10%	95
Decile 3 (21→30)	34	24%	9	29	7%	4	26	4%	104
Decile 4 (31→40)	40	8%	13	36	2%	10	36	3%	97
Decile 5 (41→50)	49	7%	26	47	2%	26	46	2%	46
Decile 6 (51→60)	57	4%	18	55	0%	21	53	1%	6
Decile 7 (61→70)	67	3%	17	63	1%	9	68	3%	2
Decile 8 (71→80)	75	2%	4	76	0%	7			
Decile 9 (81→90)	86	3%	2	87	1%	4			
Decile 10 (91→100)	107	0%	2	91	0%	1			

Letter-sounds, word-reading and oral reading fluency

Figure 2 below shows the strong and predictable relationship between both *letters-read-correctly-per-minute* and ORF (panels A, C and E), as well as between *single-words-read-correctly-per-minute* and ORF (panels B, D and F). Decile-0 in the graph represents learners who scored zero on the ORF task; Decile-1 represents those scoring 0→10 WCPM; Decile-2 those who scored 11-20 WCPM and so on. Looking across the three language groups we can see that approximately 75% of the learners in Decile-0 could only pronounce 15 or fewer letter sounds in a minute and less than 5 single words in a minute.

The similarities between Northern Sotho and Xitsonga are clear, particularly when looking at the right-panel graphs (single-words-correct-per-minute and ORF). There is a tight interquartile range of approximately 5-10 single-words per ORF decile. This shows the lock-step relationship between reading single-words correctly and connected text fluently.

The 'slope' of the right panel graphs is clearly steeper for isiZulu showing a tight relationship where the interquartile range of single-words roughly maps to the ORF decile, i.e. for the ORF Decile-3 (ORF scores of 20-30WCPM) the single-word interquartile range is about 19-25. This in contrast to both Northern Sotho and Xitsonga which exhibit flatter slopes, i.e. these learners are reading fewer single-words correct in a minute than ORF words correct in a minute. For example, in Northern Sotho learners in ORF Decile-5 (ORF scores of 40-50 WCPM), are only reading 22-30 single-words correct per minute. While this may initially seem surprising, closer inspection of the EGRA assessment provides a logical explanation: the single-word assessment included only lexical words and excluded all function words, as explained earlier.

Table 3: *Descriptive statistics for EGRA sub-components by language*

	RCPM = read correctly per minute	Sample	10th Perc	25th Perc	50th Perc	75th Perc	90th Perc	Min	Mean	Max	SD
Northern Sotho	Total letters RCPM	135	11	19	**31**	41	48	1	30,4	62	14,0
	Total single-words RCPM	135	1	8	**22**	29	35	0	19,3	40	12,1
	Total non-words RCPM	135	0	4	**13**	20	25	0	12,5	30	8,9
	Total ORF RCPM	135	0	11	**41**	55	67	0	36,2	109	25,7
	Story Title	132	0	1	**2**	2	2	0	1,4	2	0,8
	Oral comprehension	135	0	0	**1**	3	4	0	1,7	6	1,7
Xitsonga	Total letters RCPM	111	5	17	**38**	48	60	0	34,3	69	19,5
	Total single-words RCPM	111	1	9	**20**	27	31	0	18,2	48	11,5
	Total non-words RCPM	111	1	6	**16**	22	26	0	14,8	42	9,9
	Total ORF RCPM	111	0	13	**47**	57	71	0	39,8	91	25,9
	Story Title	111	0	2	**2**	2	2	0	1,6	2	0,8
	Oral comprehension	111	0	2	**4**	5	6	0	3,4	8	2,2
isiZulu	Total letters RCPM	494	4	13	**25**	39	49	0	26,4	85	17,1
	Total single-words RCPM	494	0	8	**19**	27	33	0	17,8	44	11,5
	Total non-words RCPM	494	0	5	**14**	21	25	0	13,5	45	9,2
	Total ORF RCPM	494	0	6	**21**	33	42	0	21,0	67	15,6
	Story Title	488	0	1	**2**	2	2	0	1,4	2	0,8
	Oral comprehension	494	0	0	**2,5**	4	5	0	2,4	6	2,0
Total	Total letters RCPM	740	5	15	**28**	41	51	0	28,3	85	17,2
	Total single-words RCPM	740	0	8	**20**	27	33	0	18,2	48	11,6
	Total non-words RCPM	740	0	5	**14**	21	26	0	13,5	45	9,3
	Total ORF RCPM	740	0	8	**26**	42	54	0	26,6	109	21,1
	Story Title	731	0	1	**2**	2	2	0	1,4	2	0,8
	Oral comprehension	740	0	0	**2**	4	5	0	2,4	8	2,1

Developing a framework for early reading development in African languages

When developing benchmarks for languages or grades one can take the approach of norming to the population as a whole. For example, Hasbrouck & Tyndal (2006, p.637) collected ORF data from students across the performance spectrum included gifted as well as dyslexic readers. The benchmarks help teachers to identify learners at risk for reading failure and requiring additional support. However, this approach becomes problematic in South Africa where the level of reading achievement in the country is so low that any population norms would be unacceptably low. For example, both South African and American learners participated in the 2006 round of the Progress in International Reading Literacy Study (PIRLS). The results showed that while 96% of American Grade 4 learners reached the Low International Benchmark, only 22% of South African Grade 4 learners reached this rudimentary benchmark (Mullis et al., 2007, p.69).

If one cannot benchmark to national norms, what are the alternatives? As in earlier work (Draper & Spaull, 2015) we argue that benchmarking to comprehension outcomes is a feasible and justifiable alternative. Given that comprehension is the goal of reading, linking reading benchmarks to this outcome seems logical, and this is the approach we take in the present study. As part of the adapted EGRA there were 8 oral comprehension questions presented to learners after their minute of ORF reading. Using the total scores on these comprehension questions as a rough classification tool we group learners into one of four categories, (1) Non-readers (those who could not read the title of the story properly), (2) Pre-readers (1-2 on comprehension; <25%), (3) Emergent-readers (3-4 on comprehension; 26-50%), and (4) Basic readers (5+ on comprehension; 62,5%+).

While these are somewhat arbitrary categories, and a short oral comprehension assessment is not ideal as the metric of comprehension, we argue that this is less of a problem for our purposes. Ultimately, we are trying to establish nascent benchmarks for reading letter-sounds, single words, non-words and connected text for previously unexamined languages. Part of this is identifying the levels of each sub-component that are typically found together for the same learner. We believe there is a similar underlying cognitive-linguistic data generating process that is consistent within a language. Our descriptive statistics seem to support this given the relatively narrow range of letter-sound and single-word scores associated with certain ORF Deciles. Table 7 shows a similarly narrow interquartile range for ORF scores relative to comprehension categories.

Figure 2. *Boxplots of total letters read correctly per minute and total single-words read correctly per minute by Oral Reading Fluency Deciles. (Note: for ORF Deciles 0 = 0WCPM; 1=0 →10WCPM; 2=11 →20WCPM; 3=21 →30WCPM etc.)*

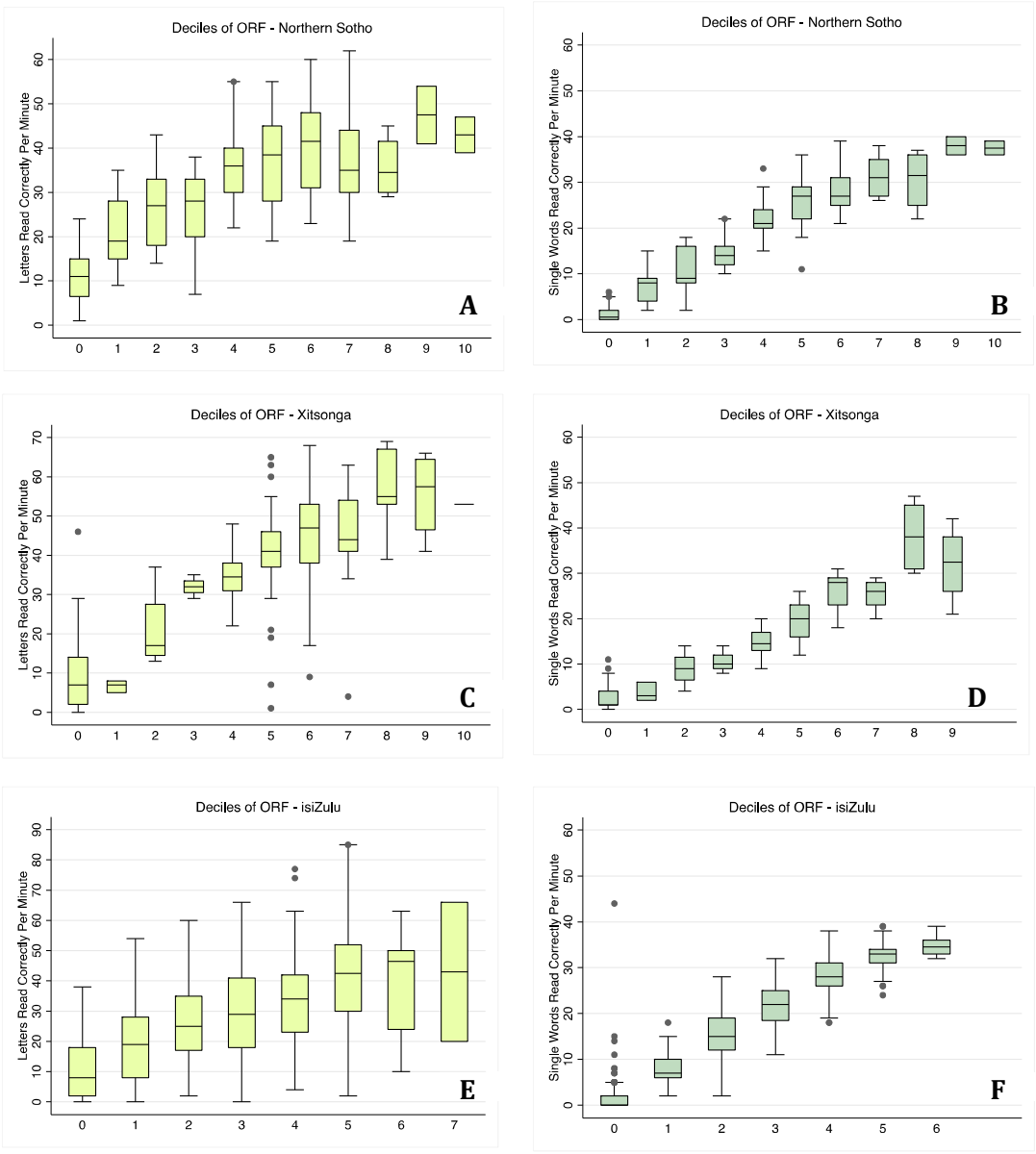

Table 7. *EGRA sub-test distributions by comprehension categories showing median scores with interquartile ranges presented in brackets*

		…read correctly per minute					
		Letters	**Single words**	**Non-words**	**Connected text (ORF)**	**Sample**	
Northern Sotho	*Non-readers*	**24** (17-31)	**8** (3-11)	**4** (2-8)	**14** (7-25)	15	*11%*
	Pre-readers	**25** (28-41)	**23** (18-28)	**14** (9-19)	**43** (34-48)	48	*36%*
	Emergent	**42** (29-49)	**30** (27-35)	**21** (18-26)	**58** (52-62)	27	*20%*
	Basic	**43** (39-46)	**33** (27-36)	**24** (21-26)	**70** (66-84)	12	*9%*
Xitsonga	*Non-readers*	**16** (10-25)	**6** (3-11)	**6** (4-9)	**12** (7-18)	4	*4%*
	Pre-readers	**33** (18-41)	**16** (13-20)	**15** (9-18)	**40** (32-50)	16	*14%*
	Emergent	**39** (34-48)	**19** (15-23)	**16** (11-20)	**48** (39-51)	30	*27%*
	Basic	**46** (38-55)	**28** (21-31)	**21** (15-26)	**57** (48-71)	43	*39%*
isiZulu	*Non-readers*	**19** (9-23)	**6** (3-13)	**5** (3-9)	4 (1-15)	37	*7%*
	Pre-readers	**26** (15-38)	**15** (10-20)	**11** (7-16)	13 (9-22)	83	*17%*
	Emergent	**34** (20-43)	**23** (18-29)	**17** (13-23)	**28** (20-35)	145	*29%*
	Basic	**34** (24-48)	**30** (26-33)	**22** (19-26)	**37** (32-43)	102	*21%*

What Table 7 seems to show is that there are certain 'minimum thresholds' below which one cannot find learners that have the requisite comprehension outcomes. To identify these, we look at the 25th percentile score for the Emergent-readers category. For example, to get 25% or more on the comprehension questions (Emergent-reader) one would need to read at least 53 WCPM in Northern-Sotho, 39 WCPM in Xitsonga and 20 WCPM in isiZulu. We will refer to these as the '*Minimum* Fluency Thresholds' for reading in these languages. Interestingly these figures are very similar to the lowest levels at which learners had 95% accuracy in reading connected text (ORF). These were 51+ WCPM (N-Sotho), 31+ WCPM (Xitsonga), and 21+ WCPM (isiZulu) – see Table 6. If one takes a more reasonable comprehension metric – that learners should achieve 62,5% or more, then learners need to be reading at least 66 WCPM in Northern Sotho, 48 WCPM in Xitsonga and 32 WCPM in isiZulu. We will refer to these as the '*Minimum* Comprehension Thresholds' for reading in these languages.

Concluding remarks

The concern about low literacy levels in developing countries such as South Africa is a valid and urgent one. Factors such as reduced time on task, inadequate access to reading materials in African languages, and poor quality early reading instruction in high poverty contexts all contribute to low literacy levels (e.g. De Stefano et al. 2012). In this article we have probed beneath the comprehension iceberg to better understand how different components of reading promote or hinder reading in agglutinating African languages with transparent disjunctive and conjunctive orthographies. The results show that across all three languages, accuracy and speed matter in reading. This finding is supported by research into reading in other alphabetic languages elsewhere in the world (Jenkins et al., 2003; Siedenberg 2017). Accuracy and speed were reflected in all the subcomponents of the reading test, with a knock-on effect from the most basic reading level, namely, letter-sounds, through word reading to ORF passage reading. The

best comprehenders were learners who read faster and more accurately than their peers.

Knowledge of letter-sounds showed strong relationships to both word and non-word reading, suggesting that readers in transparent orthographies rely on letter-sound conversion to decode words accurately. Although performance was better on the word than nonword reading tasks, as predicted by research elsewhere (e.g. Adams, 1990; Miller et al. 2014), performance on the two subtasks was highly correlated, as in the Veii & Everatt (2015) study with Herero children. Learners who could not sound out, minimally, 25-30 letters correctly per minute on this subcomponent of the test fell into the non-reader or pre-reader categories, suggesting that although they were entering their third year of schooling, they had not yet been launched on a successful reading trajectory. Letter-sound knowledge of the complex consonant system in African languages may help to fine-tune phonological awareness, enabling readers to make finer distinctions at the phonemic level, which in turn improves word processing. Systematic phonics instruction early in the foundation phase may help to mitigate this backlog in grasping the alphabetic principle.

Although reading scores did not differ much across languages in the word and nonword subtasks when function words were excluded, large differences in ORF scores showed up when learners read extended text. Differences in word length in the disjunctive and conjunctive orthographies of Northern Sotho and isiZulu respectively affect reading rate. This has important implications for benchmarking and for identifying at-risk readers at different grade levels.

Although more research is still needed, the differential reading rates in the conjunctive/disjunctive orthographies have implications for streamlining the benchmarking process; rather than establish benchmarks for each individual African language (a costly and time-consuming process), benchmarks for the conjunctive/disjunctive orthographies may suffice. Separate, intermediate benchmarks for languages that show features of both orthographies, such as Xitsonga, should also be established.

Although some reading studies in African languages have not shown a strong relationship between ORF and comprehension (e.g. Malda et al. 2014; Rees 2016), a strong relation obtained in this study. Comprehension was compromised when speed and accuracy dropped below minimum thresholds. Reading below 50wcpm and 40wcpm in Grade 3 seem to signal at-risk readers in Northern Sotho and Xitsonga respectively, while reading below 20wcpm signals an at-risk reader in isiZulu. If a comprehension threshold of at least 60% is desired, then learners should be reading at least 10 wcpm faster than the above scores in the respective languages.

Irrespective of whether languages are analytic or agglutinating, have transparent or opaque scripts, systematic phonics instruction tailored to language-specific orthographic characteristics can provide children learning to read an alphabetic script with letter-sound knowledge that forms accurate building blocks pertinent for word reading in their language. Easy access to reading material will also be critical. Fluency in word and passage reading is built up through daily opportunities to practise reading extended texts in and out of the classroom (Spear-Swerling, 2006; National Reading Panel, 2000).

It is also important to identify learners who get off to a slow start in reading in the first three years of schooling. Thal, Bates, Goodman and Jahn-Samilo (1997, p.241) argue that 'if there are no clear criteria for identifying what is 'normal', then it is especially difficult to be certain that a child is delayed or precocious'. There is no 'one size fits all'; reading benchmarks are language specific. In order to reduce inequalities in literacy, it is important for teachers in developing countries to be aware of appropriate reading benchmarks in different languages in which reading is taught. We argue here that we need to move beyond a repetitive focus on low comprehension outcomes; this is simply the tip of the iceberg. Below the surface there is widespread evidence that most children have not acquired the basic 'tools' for reading success – the ability to accurately and fluently decode letters and words and move from an effortful activity to an automated skill.

Acknowledgements

The study is funded by the Economic and Social Research Council (ESRC) and the Department for International Development (DFID) in the United Kingdom under the *Raising Learning Outcomes in Education Systems 2015* research grants.

References

Acha, J., Laka, I., & Perea, M. (2010). Reading development in agglutinative languages: Evidence form beginning, intermediate and adult Basque readers. *Journal of Experimental Child Psychology,* 105, 359-375.

Adams, M. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

Adams, M.J. (1994). Modeling the connections between word recognition and reading. In Ruddell RB, Ruddell MR & Singer H. (eds) *Theoretical models and processes of reading* (pp 838-863). Newark: International Reading Association.

Alcock, K.J., Ngorosho, D., Deus C. & Jukes, M.C.H. 2010. We don't have language at our house: Disentangling the relationship between phonological awareness, schooling, and literacy. *British Journal of Educational Psychology* 80(1), 55-76.

Al Otaiba, S.A., Petscher, Y., Williams, R.S., Pappamihiel, N.E., Dyrlund, A.K., & Connor, C. (2009). Modelling oral reading fluency development in Latino students: A longitudinal study across second and third grade. *Journal of Educational Psychology, 101(2)*, 315–329.

Babayağit, S. & Stainthorp R. (2007). Preliterate phonological awareness and early literacy skills in Turkish. *Journal of Research in Reading*, 30, 394-413.

Blaiklock, K. (2004). The importance of letter knowledge in the relationship between phonological awareness and reading. *Journal of Research in Reading*, 27, 36-57.

Blachowicz, C.L.Z., Moskal, M.K., Massarelli, J.R., Obrochta, C.M., Fogelberg, E., & Fisher, P. (2006). "Everybody reads": Fluency as a focus for staff development. In T. Rasinksi, C. Blachowicz, & K. Lems (Eds.), *Fluency instruction: Research-based practices* (pp. 141–154). New York, New York: The Guilford Press.

Burgess, S.R. & Lonigan, C.J. (1998). Bi-directional relations of phonological sensitivity and pre-reading abilities: Evidence from a pre-school sample. *Journal of Experimental Child Psychology*, 70, 117–141.

De Stefano, J. (2012). *State of the art knowledge in Education: Opportunities to learn*. USAID/EQUIP2, Washington, DC: FHI360.

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.

Diemer, MN. 2015. The contributions of phonological awareness and naming speed to the reading fluency, accuracy, comprehension and spelling of Grade 3 isiXhosa readers. Unpublished MA dissertation. Grahamstown: Rhodes University.

Draper, K., & Spaull, N. (2015). Examining oral reading fluency among Grade 5 rural ESL learners in South Africa: An analysis of NEEDU 2013. *South African Journal of Childhood Education, 5*(2), 44-77.

Ehri, L.C. (2005). Learning to read words: Theory, findings and issues. *Scientific Studies of Reading* 9(2),167-188.

Ellis, N.C., & Hooper, M. (2001). Why learning to read is easier in Welsh than in English: Orthographic transparency effects evinced with frequency matched tests. *Applied Psycholinguistics*, 22(4), 571-599.

Fuchs, L.S., Fuchs, D., Hosp, M.K., & Jenkins, J.R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading,* 5(3), 239–256.

Hasbrouck, J. & Tindal, G.A. 2006. Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, 5, 636-644.

Jenkins, J.R., Fuchs, L.S., Van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and fluency. *Journal of Educational Psychology, 95(4)*, 719–729.

Jimerson, S.R., Hong, S., Stage, S., & Gerber, M. (2013). Examining oral reading fluency trajectories among English language learners and English speaking students. *New Approaches in Educational Research, 2(1)*, 3–11.

Kilpatrick, D.A. (2015). *Essentials of assessing, preventing and overcoming reading difficulties.* Hoboken,NJ.: Wiley.

Lekgoko, O. & Winskel, H. 2008. Learning to read in Setswana and English: cross-language transference of letter-knowledge, phonological awareness and word reading skills. *Perspectives in Education*, 26(4), 57 - 73.

Malda, M., Nel, C. & van de Vijver, F.J.R. (2014). The road to reading for South African learners: The role of orthographic depth. *Learning and Individual Differences,* 30, 34-45.

Makaure, ZP. 2017. Phonological processing and reading development in Northern-Sotho – English bilinguals. Unpublished MA dissertation. University of South Africa, Pretoria.

Miller, P., Kargin, T. & Guldenoglu, B. (2014). Differences in the reading of shallow and deep orthography: developmental evidence from Hebrew and Turkish readers. *Journal of Research in Reading,* 37(4), 409–432. Doi: 10.1111/j.1467-9817.2012.01540.x

Mullis, I., Martin, M., Kennedy, A., & Foy, F. (2007). *PIRLS 2006 International Report: The Progress in International Reading Literacy Study in Primary Schools in 40 Countries*. TIMSS & PIRLS International Study Center: Boston.

Muter, V., & Diethelm, K. (2001). The contribution of phonological skills and letter knowledge to early reading development in a multilingual population. *Language Learning,* 51, 187-219.

National Reading Panel. 2000. *Teaching children to read: An evidence-based assessment of the scientific literature on reading and its implications for reading instruction*. Washington, DC: US Government Printing Office.

Nieto, J.E. (2005). Analysis of difficulties in understanding and applying the alphabetic principle. *Journal of Research in Educational Psychology* 2(2), 75-104.

Oney, B. & Durgunoglu, A.Y. (1997). Beginning to read in Turkish: A phonologically transparent orthography. *Applied Psycholinguistics*, 18, 1–15.

Ouellette, G.P. & Haley, A. (2013). One complicated extended family: the influence of alphabetic knowledge and vocabulary on phonemic awareness. *Journal of Research in Reading* 36(1), 29-41.

Perfetti, C., Beck, I., Bell, L., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly*, 33, 283-319.

Pretorius, E.J. 2015. Failure to launch: Matching language policy with literacy accomplishment in South African schools. *International Journal for the Sociology of Language* 234, 47-76.

Pretorius, E.J. & Knoetze, J.J. (2012). The teachers' book club: Broadening teachers' knowledge and building self confidence. *Musaion* 31(1), 27- 46.

Pretorius, E.J. & Spaull, N. (2016). Exploring relationships between oral reading fluency and reading comprehension amongst English second language readers in South Africa. *International Journal: Reading and Writing,* DOI 10.1007/s11145-016-9645-9.

Rees, S.A. 2016. Morphological awareness in readers of isiXhosa. Unpublished MA dissertation, Rhodes, University, Grahamstown.

Share, D.L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55, 151–218.

Siedenberg, M. (2017). *Language at the speed of sight*. New York: Basic Books.

Spear-Swerling L. (2006). Children's reading comprehension and oral reading fluency in easy text. *Reading and Writing: An Interdisciplinary Journal, 19,* 199-220.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–406.

Thal, D.J., Bates, E., Goodman, J. & Jahn-Samilo, J. (1997). Continuity of language abilities. An exploratory study of late- and early- talking toddlers. *Developmental Neuropsychology*, 13, 239-273.

Veii, K. & Everatt, J. 2005. Predictors of reading among Herero-English bilingual Namibian school children. *Bilingualism: Language and Cognition*, 8(3), 239-54.

Vihman, M.M. (1996). *Phonological development*. Oxford: Basil Blackwell.

Wilsenach, C. 2015. Receptive vocabulary and early literacy skills in bilingual Northern Sotho-English bilinguals. *Reading and Writing* 6(1), Art.#77, 11 pages. http://dx.doi.org/10.4102/rw.v6i1.77

Wise, J.C., Sevcik, R.A., Morris, R.D., Lovett, M.W. & Wolf, M. (2007). The growth of phonological awareness by children with reading disabilities: A result of semantic knowledge or knowledge of grapheme–phoneme correspondences? *Scientific Studies of Reading,* 11, 151–164

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading, 5(3),* 211–239.

Ziegler, J. & Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin,* 131, 3-29.