Review of research on reading acquisition and analyses of the main international reading assessment tools

Liliane Sprenger-Charolles, Souhila Messaoud-Galusi

<div style="border:1px solid black">

Review of research on reading acquisition and analyses of the main international reading assessment tools

Report, IIEP-UNESCO
(Accepted, November 10th, 2009)

Liliane Sprenger-Charolles* & Souhila Messaoud-Galusi**
(QLIP "Quality Learning Indicators Project"; Coordinator: Dan Wagner)

*LPP (Laboratoire Psychologie de la Perception), CNRS, Paris-Descartes and ENS
**UCL, Department of Speech, Hearing and Phonetic sciences

</div>

CONTENT

Acknowledgements

Table of Contents

English Executive Summary

Several international tools are used to conduct surveys and measurements of literacy: for instance, PIRLS (Progress in International Reading Literacy Study, International Association for the Evaluation of Educational Achievement), PISA (Program for International Student Assessment, OECD), and more recently EGRA (Early Grade Reading Assessment, World Bank and Research-Triangle-Institute, RTI). These tools have similar purposes. However, they differ in the methodology that they use and in the abilities that they assess. This is due to the fact that reading is not a unitary construct and testing batteries most often measure only some of the many abilities that should be mastered to become literate, from the mastery of lower level processing (decoding skills) to higher level processing (reading comprehension). In addition, it is not with exactly the same tools that should be assessed the reading abilities of beginning readers (the target population of EGRA, for instance) and of more experienced readers (the target population of PISA, for instance).

There are many studies on this topic in countries from OECD, not in developing countries, where designers of reading assessments are faced with specific challenges. Especially, in such countries learning to read is often based on a non-native language (mainly English, French or Spanish). Consequently, familiarity with the language of instruction may differ among individuals assessed, as well as between enumerators, and this might affect the outcome of the assessment

In the present review, the outcomes of research literature on reading acquisition are first examined (Part 1). Then, available reading assessments tools and reports are described and appraised (Parts 2-6). Particular attention will be paid to how to best assess reading acquisition in different linguistic contexts, and in the context of developing countries.

_____

In the text, an individual's native language is designated by the term "L1", and "L2" is used for a second language. "Lit1" and "Lit2" (reading in a native versus non-native language) are used to solve ambiguities (e.g. incidence of L1 spoken language abilities on Lit2 and vice versa).

Résumé en français

Les tests de lecture à la disposition de la communauté internationale comme PIRLS (Progress in International Reading Literacy Study, International Association for the Evaluation of Educational Achievement), PISA (Program for International Student Assessment, OCDE), et plus récemment EGRA (Early Grade Reading Assessment, Banque Mondiale et Research-Triangle-Institute, RTI) ont un but similaire. Néammoins, ils diffèrent dans la méthodologie qu'ils utilisent, et dans les compétences qu'ils mesurent. Cela peut s'expliquer par le fait que la lecture n'est pas un concept unitaire. Les instruments qui mesurent cette capacité ne capturent donc le plus souvent que partiellement les dimensions multiples impliquées dans le savoir lire : de la maîtrise des processus de bas niveaux (décodage) à la maitrise des processus et de hauts niveaux (compréhension écrite). En outre, ce n'est pas avec exactement les mêmes outils qu'il est possible d'évaluer les compétences en lecture des lecteurs débutants (la population cible de EGRA, par exemple) et celles de lecteurs plus confirmés (la population cible de PISA, par exemple).

Il existe de nombreux travaux de recherche sur ce sujet dans les pays de l'OCDE, mais pas dans les pays en développement, dans lesquels les créateurs de tests sont confrontés à des difficultés spécifiques parce que, dans ces pays, l'apprentissage de la lecture se fait le plus souvent dans une langue seconde, principalement l'anglais, le français ou l'espagnol. Par conséquent, la connaissance de la langue d'instruction varie chez les individus évalués, tout comme chez les évaluateurs, ce qui peut affecter les résultats.

Dans la présente revue, les travaux de recherche sur l'acquisition de la lecture sont d'abord présentés (Partie 1). Cette partie est suivie par une présentation et une analyse critique des tests internationaux de lecture disponibles (Parties 2-6). Une attention particulière est portée à l'acquisition de la lecture dans différents contextes linguistiques, et dans les pays en voie de développement.

_____

Dans le texte, la langue maternelle est dite « L1 » (langue 1), et « L2 » est utilisé pour une langue seconde. « Lit1» et « Lit2 » sont respectivement utilisés quand l'enseignement de la lecture est en langue maternelle (L1) ou en une langue seconde (L2), mais uniquement quand il peut y avoir une ambiguité: par exemple, incidence de la maîtrise de la langue orale en L1 sur la langue écrite en Lit2 et vice versa.

1. Review of research on reading acquisition (psycholinguistic approaches)

Since the beginning of the 1980's, many studies have been published on reading acquisition. These studies have been reviewed in articles, chapters or books designed to help develop new educational policies. Some of these reviews have mainly considered studies with English speakers[1]. Others have considered studies with students learning to read in different languages (mostly languages with an alphabetic writing system such as English, French, and Spanish)[2]. In some of these reviews, special attention has been paid to studies with L2 learners[3]. In the present review -- mainly based on the previously quoted reviews -- the research conducted so far on reading acquisition is examined in L1 and in L2 learners[4].

1.1. Introduction

1.1a. Specific, and non specific, reading skills

To be able to understand a written statement, it is necessary to be able to understand that statement in its spoken form, and thus to master the spoken language in which it is written. In addition, it is necessary to master some skills specific to reading. Processes specific to reading can be understood when compared to those required to read a music score. Indeed, trained musician are able to automatically associate the written notes to the corresponding melody in their mind. Alternatively, untrained listeners who are able to hear and understand the same melody played onto the piano will nevertheless be unable to read this melody on a music score. Likewise, normally intelligent students with accurate vision and audition who master the spoken language can only read if they have automated a process specific to print, namely written word identification.

As for music, reading comprehension thus depends on two processes: (1) "a high level process" involving different levels of linguistic comprehension, from the lexical level (vocabulary), to the understanding of short utterances (orders…) and up to the understanding of long texts; (2) "a low level process" involving automatic written word identification. The relationships between automatic written word identification, oral and written comprehension abilities can be summarised as followed: the level of reading comprehension depends on the level of oral language comprehension and on the level of automaticity in written word identification (accuracy and speed)[5].

This statement is supported by the fact that reading comprehension is highly correlated with oral comprehension. For instance, a child unable to understand the spoken language in which s/he learns to read (a native speaker suffering from specific language impairment[6a] or a non-native speaker with a low level of mastery of L2[6b]) will have difficulty in understanding what s/he is reading. Alternatively, for a child without oral comprehension problems, reading

---

1. Adams, 1990; Castles & Coltheart, 2004; National Reading Panel, 2000 (US); Ehri et al 2000a-b; Rayner et al., 2001; Share, 1995; Vellutino et al., 2004.
2. Share, 2008; Sprenger-Charolles, 2003; Sprenger-Charolles et al., 2006; Ziegler & Goswami, 2005 and 2006.
3. August & Shanahan, 2006a-b; Snow & Kang, 2006. As explained in August and Shanahan (2006b), this issue is important since there are a growing number of L2 students (in US: from 6 million by 1979 up to 14 million in 1999). In addition, some L2 students are not faring well in school (according to US states reports, only 19% of L2 learners score above the states-established norm for reading comprehension). Finally, whereas few L1 failed to complete high school, this is the case for 31% L2 students.
4. For studies recently published, we have consulted the Web of Sciences (we do not found studies on reading acquisition involving L1 or L2 learners from developing countries). In addition, it is to be noted that this review considers research on reading acquisition from the beginning of reading instruction, and not research on the impact of pre-reading skills (before reading instruction) on future reading skills.
5. Adapted from Hoover & Gough, 1990; See also Carver, 1998; Catts et al., 2006; Joshi & Aaron, 2000; Laberge & Samuels, 1974.
6. 6a: Bishop, 1997; Cain, & Oakhil, 2007; 6b: Lesaux & Geva, 2006; Lesaux et al. 2006a. We will return to this issue later (Section 1.6).

comprehension difficulty will mainly be attributable to the fact that s/he has not automated his/her word identification skills. Indeed, the single most salient and universal fact about skilled readers is "the remarkable speed and apparent effortlessness of word identification" which reflect an "automated process"[7]. Using an automatic process to read words allows freeing up cognitive resources that will help the reader focus on reading comprehension[8].

1.1b. Relationship between written and spoken words

One basic task facing a beginning reader is to learn to associate the written words corresponding to the spoken words that are already part of his/her oral vocabulary (at least for L1 learners). For example, s/he will have to learn to convert the orthographical form of the word "speed" into sounds -- its phonological form -- and into meaning. In an alphabetic writing system, the conversion from the orthographic form of a word to its phonological form and its meaning can be done in two ways.

❖ It can be done indirectly, by a phonological route (also called alphabetic route or decoding skills) based on meaningless sublexical units, i.e., grapheme-phoneme correspondences. For instance, the graphemes 's', 'p', 'ee', and 'd' will be respectively matched to the phonemes /s/, /p/, /i:/ and /d/ and then assembled to produce the word "speed". Note that graphemes are the smallest written units in an alphabetic written system which correspond to phonemes[9]; and phonemes are the smallest units of sound allowing for differentiation of two spoken words within a specific language (e.g. "top" and "mop" differ by only one phoneme, but the meaning changes).

❖ It can also be done through the orthographic route. That reading route can be defined as being a reading route using the lexical principle of an alphabetic writing system. In such systems, written words not only encode the meaningless sublexical units of spoken language (phonemes) but also lexical units such as morphological marks ('sing' and 'ing' in singing, 's' for the plural…) and marks that recall the history of words (for example, the 'th' of "theatre", from the Greek, which is pronounced /t/ in French, unlike in English). This reading route relies on lexical reading units, with a meaning (i.e., word and morpheme). It is also to note that the development of the orthographic route depends on the efficiency of the phonological reading route (see 1.4); and that expert readers have automated the orthographic reading route (see 1.3).

1.1c. Presentation of the different sections of the review

The following sections focus on alphabetic writing systems. First, results showing how the two reading routes function are presented (Section 1.2: Decoding skills or phonological reading route; Section 1.3.: Orthographic reading route and Automatic written word identification). The other sections focus on the stages of acquisition of the two reading routes (Section 1.4) and on the incidence of the depth of the orthography on the use of these reading routes (Section 1.5). The focus of the last section (1.6.) is on the relationship between reading comprehension versus spoken language comprehension and written word identification. The available evidence in L1 and L2 learners will be discussed in relationship to their implication for developing countries.

---

7. Share, 2008 (p.590). Automatic word identification should not be confused with "speed reading", in the perspective of Goodman (1967). In this perspective, reading is conceived as being a guessing game: readers, and especially good readers, are supposed to read only some of the words in a text and to guess the others based on the context. Subsequent studies (particularly those in which eye movements have been examined during reading) have shown that, contrary to that assumption, good readers tend to look at each word of a text and identify them automatically (Rayner, 1998; Rayner et al. 2001 and 2006).
8. See for results with adults: Kintsch & Mross, 1985; Till et al., 1988; for results with children, see Perfetti et al., 1979; Perfetti & Hogaboam, 1975; Stanovich, 1980; Stanovich et al., 1981; Stanovich & West, 1979; West & Stanovich, 1978.
9. A grapheme might be composed of one or more than one letter (e.g. "e" or "ee" in the English words "be" and "bee"), or of a letter with a diacritic mark (e.g. "é" vs "e" in the French words "de" and "dé").

1.2. Decoding skills (or phonological reading route)

1.2a. Linking sounds (phonemes) with print (graphemes)

The main challenge facing beginning readers in alphabetic writing systems deals with learning how to decode print. The decoding process involves breaking down written words into graphemes which are associated with the corresponding phonemes.

Decoding skills (i.e. the use of grapheme-phoneme correspondences) are assessed using new words, which have never been encountered by the reader before. In every day life, we have to rely on our decoding skills: for example, when we are reading in a story the name of an unknown character or of an unknown city.

Using tasks involving unknown words (called "pseudowords"), it has been shown that the level of decoding skills allows identifying good and poor readers. Such results have been obtained at various ages or grades and whatever the language or the learner's status: L1 or L2 learners[10]. Decoding skills are also highly predictive of future reading abilities: for instance, a child who is very slow in the acquisition of decoding skills at the beginning of reading acquisition is likely to become a poor reader; and the reverse is true for a child with an early good level of decoding skills, whatever the language or the learner status[10]. In addition, phonics programs aimed at improving decoding skills are extremely effective programs to improve reading in young L1 and L2 learners[11]. And with such reading program L2 learners can reach a level of decoding skills equivalent to that of L1 learners, at least for accuracy[12].

Longitudinal data also indicate that from childhood to adolescence decoding skills allows word identification in beginners or experienced readers (section 1.4). However in some older readers, decoding skills have sometimes been shown to be accurate but slow[13]. Therefore, an assessment that relies on accuracy only would incorrectly classify an accurate but slow reader as being a good reader. It is thus necessary to take into account both accuracy and speed in an assessment of decoding skills[13].

1.2b. Breaking down the alphabetic code: Phonemic awareness

The mastery of grapheme-phoneme correspondences depends on the ability to segment spoken words into phonemes, which is not a straightforward task. Individual phonemes are rather abstract segments, due to the fact that spoken words (such as "cat", "need", "thus"…) are pronounced in a single articulation. This renders smaller units of the size of the phoneme difficult to isolate and can make grapheme-phoneme correspondences extremely difficult to grasp. Importantly, the awareness that words can be divided up into phonemes (i.e. phonemic awareness) is crucial for beginning readers in L1 and L2.

It has been shown by more than 30 years of research that the relationships between phonemic awareness and learning to read are reciprocal[14]. For instance, phonemic awareness is acquired gradually, in conjunction with learning to read[15a]. However, early phonemic awareness skills

---

10. See for reviews of studies with L1 learners: Siegel, 1993; Share, 1995 and 2008; Sprenger-Charolles, 2003; Sprenger-Charolles et al., 2006, pp. 78-90; Ziegler & Goswami, 2005; for reviews of studies with L2 learners, see: Genesee et al., 2006; Geva, 2006; Lesaux & Geva, 2006; Lesaux et al., 2006a; see also Chiappe et al., 2002a-b; Lesaux & Siegel, 2003.

11. For reviews of studies with L1 learners, see Ehri et al., 2001a; Share, 1999; see also Wagner et al., 1997; For reviews of studies with L2 learners: Lesaux & Geva, 2006; Lesaux et al., 2006a; Shanahan & Beck, 2006. See also Lesaux & Siegel, 2003. However, as highlighted by August & Shanahan (2006b, p.3) "while approaches that are similar to those used with native language populations are effective, the research suggests that adjustments to these approaches are needed to have maximum benefit with language minority students", for example, work on particular phonemes that do not exist in their own language (see also Shanahan & Beck, 2006).

12. See Lesaux & Geva, 2006; Lesaux et al., 2006a; Shanahan & Beck, 2006.

13. Shaywitz et al., 1999; Share, 2008.

14. Perfetti et al., 1987; Stanovich & Siegel, 1994. See also Bradley & Bryant, 1983; Lundberg et al., 1988;

15a. Morais, 1987; Morais et al., 1979.

15b. See for reviews of studies with L1 learners: Share, 1995 and 2008; Sprenger-Charolles et al., 2006 (pp. 56-65 and 134-139); Ziegler & Goswami, 2005; see also Caravolas et al., 2005.

constitute a reliable predictor of future reading achievement in L1 learners[15b], as in L2 learners[15c-d]. In addition, a training involving phonemic awareness exercises is very effective to improve decoding skills, most of all when coupled with training on grapheme-phoneme correspondences, which has been evidenced in L1 and L2 learners[16].

1.2c. The role of phonological short-term memory (STM)

Phonological STM allows us to hold a small amount of information in mind for a short period of time. This memory is involved when storing verbal material as when we are told to dial a 10 digits phone number: the string of numbers is temporarily stored in phonological STM, the time necessary to press the digits on the phone pad. To avoid forgetting, the string of digits is frequently repeated over and over. However, one rarely remembers the digits rehearsed as stored in STM once the number is dialled and stops being of any use.

It has been suggested that phonological STM is crucial to setting up decoding skills[17a]. Indeed, decoding skills requires the ability to connect each individual grapheme with the corresponding phoneme, a process followed by the assembly of the resulting grapho-phonemic units to form a word. The first operation requires a good phonemic awareness and the second relies mostly on an adequate phonological STM to sustain the assembly process. A child unable to correctly handle phonemes, and/or suffering from a deficit in STM, will hardly be able to use decoding skills and will experience severe reading difficulties[17b].

Poor phonological STM skills have typically been observed in disabled readers[17c]. However, phonological STM skills are less reliable predictors of reading level than phonemic awareness skills, at least for L1 learners[18]. In addition, poorer phonological STM has been found in L2 learners relative to L1 learners, although the difference appears to decrease over time[19]. Poorer STM in L2 relative to L1 learners has implication for reading acquisition as it can impact on decoding skills. Moreover, STM contributes to vocabulary learning, which not only has an impact on L2 learning but on L2 reading too[19].

1.3. Orthographic reading route and automatic word identification

A clear explanation of what is automatic written word identification has been proposed by Snow and Kang (2006), who emphasize that good readers "are looking at and processing most of the letters in almost all words on a page" whereas "skilled reading fells as it goes too fast for such detailed attention". They explained this apparent contraction by the fact that "good readers manage to process frequently occurring sequence of letters very efficiently because they have seen them and converted them into phonological form so often. Thus sequence such as *-ation*, *-itude*, can be processed as units. Although readers are not aware of this automatic chunking and rapid processing, it becomes obvious because of the difficulties of reading words in which unfamiliar sequences occur, for example, *Ghazi Ajil al-Yawar*, *Tblisi*, or *diyelthyl-m-toluamide*.This aspect of skilled reading is referred to as automaticity"[20].

15c. See for reviews of studies with L2 learners, see: Lesaux & Geva, 2006; Lesaux et al., 2006a; Genesee et al., 2006; See also: Chiappe et al., 2002a; Chiappe & Siegel, 1999; Everatt et al, 2000; Lesaux & Siegel, 2003.

15d. However, there are differences between the results of experimental studie and those of some field studies (Samuels, 2007; Sprenger-Charolles, 2008a-b) which suggest that phonemic awareness is not a good predictor of reading level. This could be due to the fact that, in experimental research, the enumerators are well trained to deliver the task and encode the results, not those involved in field studies who, in addition, might not be aware of some subtle characteristics of the sound pattern of the language in which the testing is done: e.g. the presence in the word "charge" of silent graphemes (r, e), and complex phonemes (ch, g). See Labov, 1995; Genesee et al., 2006.

16. For reviews with L1 learners, see Ehri et al., 200b; Castles & Coltheart, 2004; for a review with L2 learners, see: Shanahan & Beck, 2006.

17a. The repetition of short vs long pseudowords is often used to assess phonological STM.

17b. Baddeley et al., 1998; see also Siegel & Ryan, 1988;

17c. For instance, Ramus, 2003; Ramus et al., 2003.

18. For instance, Chiappe et al., 2002c; Pennington et al., 2001.

19. For instance, Chiappe et al., 2002b; Lesaux & Siegel, 2003; see also Genesee, 2006; Genesee & Geva, 2006.

20. Snow & Kang, 2006 (p.82).

Reading through the orthographic reading route is based on this automatic word recognition process which allows freeing useful memory capacity that becomes available for comprehension processes. Comprehension depends on long lasting storage of information in long term memory, as well as on the temporary storage of information waiting to be processed in short term memory. Written word identification is not in place in beginning readers who instead heavily rely on decoding. In beginning or inexperienced readers, the slow decoding process will use up most of their phonological memory capacity: as a result, few cognitive resources will be available for comprehension. On the other hand, skilled readers (who rely on automatic word identification) will be able to process words at a lesser cognitive cost. As a consequence, they will be able to allocate greater resources to reading comprehension[21].

As indicated by three decades of research, early phonological reading skills (decoding skills) and phonological reading-related skills (phonemic awareness) are the main predictors of automatic written word identification for L1 learners. For L2 learners early phonological reading and reading related skills are also the main predictors of written word identification, and better predictors than oral language proficiency[22]. In addition, for written word identification, L2 learners can reach a level equivalent to that of L1 learners[23].

1.3a. The issue of speed in automatic written word identification

What characterises an automatic process is mainly the speed at which it occurs. Thus, in order to determine whether written words are automatically accessed when using a reading test, it is crucial to take processing speed into account. This has almost always been done in studies conducted in languages with a more transparent orthography than English (e.g. Spanish, Italian, German or French)[24]. In such languages, accuracy scores are not sufficient to characterise and detect poor performers, as they very quickly reach ceiling level (at the end of the first grade in most cases, see section 1.5). Processing speed is thus the only possible measure that allows differentiating poor from good readers[25].

Automatic written word identification is a very fast process which requires only a few milliseconds, and can discriminate amongst good and poor readers. In studies that took into account, for example, the latency of vocal responses (i.e. the delay between the appearance of the word on the computer screen and the start of its pronunciation by the participant) the differences between good and poor readers are of almost 200 milliseconds per word. Such a delay corresponds to 1 minute of difference between poor and good readers for the reading of a text including 300 words, i.e. a one half page text, the length of most of the texts included in school books for elementary school. The additional time taken by poor readers to read a short text should impede their understanding of what they have read, not only because by the time they reach the end of the text, they probably have forgotten its beginning, but also because they have less memory capacity available for the comprehension process.

Speed in written word identification is nonetheless subject to variation in good as well as in poor readers. These variations depend not only on the reading level of the student but also on his/her language (for example, the transparency of the orthography, see Section 1.5). It is

---

21. Perfetti et al., 1979; Perfetti & Hogaboam, 1975; Stanovich et al., 1981; Stanovich & West, 1979; West & Stanovich, 1978. These experimental studies have clearly demonstrated the importance of the mastery of isolated word reading in reading comprehension. Alternatively, the results of some assessments based on DIBELS (Good, & Kaminski, 2002) suggest that the reading of word in context (Oral reading fluency, ORF) is a good indicator of reading comprehension (Fucks et al. 2001; Riedel, 2007). However, in the study by Riedel, it has been found that a striking characteristic of students with satisfactory ORF skills but poor reading comprehension was poor vocabulary skills and that vocabulary level is linked to reading comprehension scores. In this study, vocabulary level has been assessed with written words; and thus with the reading of isolated word.

22. For reviews with L1 learners, see: Share, 1995; Sprenger-Charolles, 2003; Sprenger-Charolles et al., 2006 (pp. 55-65); Stanovich & Siegel, 1994; For reviews with L2 learners, see: Geva, 2006; Lesaux & Geva, 2006; Lesaux et al., 2006a; see also Chiappe et al., 2002a-b; Lesaux et al., 2007; Verhoeven, 1990 and 2000.

23. See Lesaux & Geva, 2006; Lesaux et al., 2006a; Shanahan & Beck, 2006.

24. Share, 2008; but see Olson et al. 1994

25. French: Sprenger-Charolles et al., 2000; German: Landerl & Wimmer, 2008; Spanish: Jimenez et al., 2009; see also Section 1.5.

therefore necessary to systematically assess the speed of isolated word reading in reading tests.

1.3b. Type of tasks and automatic written word identification

Another issue about word recognition is related to the type of task used to assess that skill: silent reading versus reading aloud. Direct and indirect evidence converge on the conclusion that the outcomes of oral and non-oral modes of reading assessments have common features. For instance, strong phonological effects have been observed in skilled readers when reading words silently. Indeed, in silent reading, the phonological code of a word is activated very early and automatically[26a], even before the activation of its semantic code[26b].

However, to measure word reading skills with young children, it is more useful to rely on reading aloud tasks. Indeed, the results obtained with such tasks are more straightforward to interpret than those from silent reading tasks because the assessor knows exactly what has been read, which makes reading aloud tasks more reliable to mark than silent reading tasks, especially with beginning and poor readers[27]. In developing countries where a large proportion of those assessed will have poor reading abilities, it will therefore be important to know precisely what has been read and favour reading aloud to silent reading tasks[28].

1.3c. Brain correlates of automatic written word identification

Several evidence points to three important neural systems in the left hemisphere for reading in children and adults: the first, around the inferior frontal gyrus (Broca's area), is involved in linguistic analysis; the second (a parieto-temporal region), is involved in phonological processing; the third, an occipito-temporal region labelled the "visual-word form area", is responsible for automatic word reading[29a].

Studies that have compared disabled and non-disabled readers have consistently demonstrated that the left occipito-temporal region does not function properly during reading in disabled readers[29b]. Language related changes in disabled and non-disabled adult readers have been examined in some studies[29c]. The main behavioural results are that there are clear differences related to the opacity of the orthography (English, French, and Italian) and the participants' reading level (disabled vs. non-disabled readers). In sharp contrast, brain results indicated that, whatever the language, the same occipito-temporal region is less activated in disabled readers as compared to non-disabled readers. These data very clearly show that a "cultural diversity" is associated with a "biological unity".

Age related changes (from 7 to 18 years of age, for instance) have also been examined[29d]. The main finding is that a single reading system develops with age for non-disabled readers: the visual-word form area, which is associated to the ability to read words automatically, the hall mark of a skilled reader. Conversely disabled readers who struggle to read especially new or unfamiliar words come to rely on an alternate system, which functions through memory networks (for instance, sight word recognition, i.e. recognition of words learned by rote).

These data indicate that in non-disabled readers new connections are created in brain areas that become consolidated as a function of the level of reading expertise. They help to understand why and how the written word identification mechanism becomes automatic.

---

26. 26a: Frost, 1998; Perfetti & Zhang, 1995; Van Orden & Kloos, 2005; see also Ferrand & Grainger, 1993; Grainger & Ferrand, 1996. 26b: Sprenger-Charolles et al., 2006 (pp.10-14); see also Perea & Gotor, 1997; Plaut & Booth, 2000.
27. See for instance, Sprenger-Charolles et al., 2003
28. Another issue concerns the unit used: number of words or of syllables correctly read? This issue has been investigated in a study with English and Czech children. Both groups did not differ in terms of number of words read in one minute, whereas Czechs read more syllables than English children. These results could be due to the fact that the length of the words is longer in Czech than in English. Therefore, in order to take into account the linguistic specificities of the different languages, it seems better to calculate the number of syllables correctly read (See also about the role of the syllable in learning to read, Colé et al., 2000).
29. 29a: Deheane et al., 2002; 29b: Demonet et al, 2004; 29c: Paulesu et al., 2001; Ziegler, 2006; 29d: Shaywitz, et al., 2007; see also Parviainen et al., 2006.

1.4. From phonological to orthographic processing

According to developmental models[30], reading skills are acquired in stages. The first step is assumed to be based on a logographic reading route involving recognition of visual cues at a global and/or local level (for instance, the length of a word and/or its first letters). The second step is assumed to be based on a phonological reading route (described in section 1.2) which involves the sequential decoding of words and requires phonemic awareness skills. More complex reading skills acquired later are assumed to involve the use of an orthographic reading route (described in section 1.3) based on meaningful units larger than the phoneme, such as the morpheme (i.e. 'most' and 'ly' in "mostly") or the word (as "skill" or "table", which cannot be decomposed in smallest meaningful unit).

Abundant evidence suggests that the development of, at least, the phonological and orthographic reading routes, occurs in successive stages[31]. First, the relationship between phonemic awareness, linked to the phonological reading route, and reading level is very strong at the beginning of reading acquisition[32a]. Alternatively, more global parsing into larger units, linked to the orthographic reading route, arises at a later stage[32b]. Furthermore, beginning readers have difficulty reading exception words such as "women" which cannot be read through grapheme-phoneme correspondences. Alternatively, they can easily read regular words such as "dog" which can be read by grapheme-phoneme correspondences. These results indicate that phonological reading skills develop prior to orthographic reading skills. Other results indicate that phonological reading skills are the building blocks allowing the setting-up of orthographic reading skills both for L1 and L2 readers[32c].

In contrast to the results about the phonological and orthographic reading route, the first step described in stage models -- the logographic stage -- has been heavily criticized. Counter-evidence comes mainly from languages with a writing system more transparent than English (French and German): in such writing systems, the phonological reading route has been shown to be used from the beginning of reading acquisition[33a]. Nevertheless, the role of logographic strategies in learning to read has been questioned even in a non-transparent writing system such as English, where these strategies have been found to be marginally used by learners and to be of little use in teaching children how to read new words[33b].

1.5. How does the writing system affect learning to read?

One of the major results of the current research on reading is that the transparency/opacity of grapheme-phoneme correspondences (GPC) has a great impact on learning to read[34]. Consistency defines the degree of transparency of GPC in a given orthography. In a transparent system (as Spanish), most of the relationships between graphemes and phonemes are one to one (e.g. the grapheme 'a' systematically corresponds to the phoneme /a/), thus GPC consistency is near 100%. In contrast, in an opaque system as English, the grapheme 'a' corresponds to several phonemes (as in the words "a", "cat" and "date"[35]) thus the consistency of GPC is very low, specifically for vowels (less than 50%). This is an important finding given that most data on reading come from studies conducted in English[36].

---

30. For instance, Morton's model, that takes into account the role on reading acquisition of environmental factors (social and linguistic background), and of cognitive skills (reading and reading-related skills), as well as the neurobiological bases sustaining that acquisition.
31. See for reviews of studies conducted in different orthographies with L1 learners: Share, 1995 and 2008; Sprenger-Charolles, 2003; Sprenger-Charolles et al., 2006 (pp.36-55); Ziegler & Goswami, 2005 and 2006; see also Casalis & Colé, 2009.
32a. See note 31. See for reviews of studies with L2 learners: Geva, 2006; Lesaux & Geva, 2006; Lesaux et al., 2006a
32b. For a review, see Sprenger-Charolles et al., 2006 (pp.46-52 and 65-67); see also Casalis & Louis-Alexandre, 2000; Casalis & Colé, 2009
33a. For a review, see: Share, 2008; 33b: Ehri & Wilce, 1985; Laing & Hulme, 1999; Stuart & Coltheart, 1988.
34. For reviews with L1 learners, see: Sprenger-Charolles, 2003; Sprenger-Charolles et al., 2006 (pp.29-41); Ziegler & Goswami, 2005 & 2006.
35. According to Coulmas (2003), English is even a system 'without a predominant principle for the interpretation of their basic units'
36. Note that many languages in developing countries use a recent – and thus transparent – orthography. See for traditional Sub-Sahara African languages (e.g. Swahili or Wolof…): International Institute of African Languages and Cultures, 1930; See also Sebba, 2007.

1.5a. Consistency of the English and French orthography

The low consistency of the English orthography compared to that of the French orthography, for example, emerges from a comparison which has been focused on monosyllabic words that began and ended with a consonant in both languages[37].

English is characterized by a low consistency of grapheme-phoneme correspondences (GPC, used to read), especially for vowels (48%). However, the consistency of the pronunciation of English vowels doubles when the rime of the word is taken into account (from 48 to 91%)[38], which is not found in French where GPC consistency for vowels is very high (94%). This helps to understand why the use of rhyming words (deep, sleep…; night, fight ...) can facilitate the learning of reading in English, not in French.

Alternatively, the consistency of phoneme-grapheme correspondences (PGC, used to spell) is low, particularly for French consonants located at the end of words (58%). This result stems from the fact that the morphological marks written at the end of words are mostly silent in French (for instance, the 't' at the end of the word "chat" /$a/, "cat" in English).

1.5b. Impact of the consistency of the orthography on reading

Several studies support the assumption that reading skills are acquired at a faster pace in languages with a transparent orthography, for instance, a study conducted in 13 languages: English, Danish, French, German, Spanish and Portuguese, among others[39a]. First graders have been assessed in all languages, and second graders in English and French only. English second graders have approximately the same age as French and Spanish first graders (6 years and half). For word reading, English second graders score similarly to Danish, Portuguese and French first graders. Conversely, for pseudoword reading, English second graders scores are lower than those of all first graders, except Danish children. These results are obtained in spite of the fact that the English children enrolled in this study had a reading level higher than the national standards. Similar trends have been reported in earlier studies with English and French students (children or adults)[39b]. Differences due to the transparency of the orthography have also been observed with L2 learners[39c].

Differences due to grapheme-phoneme correspondence consistency are not only quantitative but also qualitative. Indeed, English-speaking students (children and adults) use more than non-English-speaking students their word knowledge to read. This is indicated by the fact that, compared to non-English speakers, the difficulties of English readers are always more pronounced in pseudoword reading than in word reading.

The opacity of the orthography also has an impact on the reading units used to process words. For instance, in a study involving English, French and Spanish children, the students had to read pseudowords which rhymed or did not rhyme with words[40]. The presence of rime facilitates more greatly the reading of English children than that of French children whereas the presence of rime has no incidence on reading for Spanish children. This probably is due to the fact that the inconsistency of English vowels is reduced when the rime of the words is taken into account, and could explain why teachers use rime units to teach reading in English.

---

37. About 3500 English words and 1800 French words (Peereman & Content, 1998; see also Venezky, 1966).
38. See Peereman & Content, 1998; See also Treiman et al., 1995; Ziegler et al., 1997.
39. 39a: Seymour et al., 2003; 39b: For other results with L1 learners, see Bruck et al., 1997; Goswami et al., 1997; Paulesu et al., 2001. 39c: For L2 learners, see: Geva & Siegel, 2000.
40. Goswami et al., 1997 (e.g. the pseudoword "fape" that rime with "tape"). For similar results with English vs German students, see Goswami et al., 2003; Ziegler et al., 2001. See also note 38. The inconsistency of GPC for English vowels also leads to the production of a higher number of reading errors on vowels than on consonants in English (Bryson & Werker, 1989; Fowler, et al.,1979; Frith et al., 1998; Siegel & Faux, 1989), not in more transparent orthographies (e.g. German: Frith et al., 1998; Wimmer, 1993; Italian: Cossu et al., 1995).

1.5c. Impact of the consistency of the orthography on spelling

Grapheme-phoneme correspondences (GPC) which are used to read are not always comparable to phoneme-grapheme correspondences (PGC) that are used to spell. Indeed, GPC are more consistent than PGC especially – but not only – in French (see 1.5a). As a consequence, learning to read is supposed to be easier than learning to spell.

Several studies support the assumption that learning to read is easier than learning to spell. For instance, regular words are very rapidly read more accurately than pseudowords, while they are not better written, particularly in French[41]. This may be explained by the fact that a regular word (like "route") can be read only one way whereas it can be written in different ways (with one or two 't', for instance), however only one form will respect the word standard spelling. On the other hand, pseudowords which have no standard spelling can be written which ever way that respects French PGC and are therefore easier to spell than words.

Beyond these differences between reading and spelling, it is essentially the phonological route which is used by beginning readers/spellers, both in L1 and L2[42]. Therefore, as in reading, regular words are better processed than irregular words in spelling, even when irregular words are of a very high level of frequency.

1.5d. Consequences for test design

The orthography of a language has a strong impact on learning to read and spell, both quantitatively and qualitatively. Thus, we must take into account a system's linguistic characteristics (especially GPC/PGC consistency) when designing assessments for beginning readers/spellers. We should also keep in mind that level of transparency might differ in writing systems when designing the same assessment in several languages. In addition, reading measurement relying on written responses should be considered with great caution, given the asymmetry between GPC (used to read) and PGC (used to spell). These issues are of a particularly importance for students learning to read and spell in L2[43], and in the context of developing countries where reading and spelling instruction in L2 will take place in a writing in which GPC and/or PGC are not transparent, as English for both reading and spelling, and French for spelling.

1.6. Reading Comprehension

Reading comprehension depends on both oral language comprehension and decoding skills[44]. However, the strength of the relationship between these skills depends on several factors: the reader's cognitive skills (for instance, memory); his/her linguistic skills (especially his/her status: native or non-native speaker); his/her background knowledge; and his/her goals when reading. They also depend on specific text properties: its complexity (at the level of the vocabulary and of the word order [syntax], and at more global levels: relationship between its different parts); its length (stories as in Appendix A versus sentences as "It is forbidden to use this website for any purpose other than personal use"); and its nature (for instance, story versus nonfictional prose compositions appearing in newspapers, magazines, encyclopedia…).

More specifically, to understand a text it is necessary to make connections between its different parts. This process depends, for instance, on the relations between the different elements of a text through the use of "anaphora" (e.g. "Wash five cooking apples. Put them into a fireproof dish") and of different marks of connection: additive (and, in addition…),

---

41. Alegria, & Mousty, 1996; Sprenger-Charolles et al., 2003.
42. Foorman et al. 1991; Juel, 1988; Lété et al., 2008; Sprenger-Charolles et al. 2003; Waters et al., 1985. But see Martinet et al., 2004. For reviews of studies with L2 learners, see: Lesaux & Geva, 2006; Lesaux et al., 2006a; Shanahan & Beck, 2006.
43. See Genesee et al., 2006; Labov, 1995.
44. For L1: Cain, & Oakhill, 2007; Catts et al., 2006; For L2, see: Lesaux & Geva, 2006; Lesaux et al., 2006a; Snow & Kang, 2006.

causal (thus, therefore…), adversative (alternatively, otherwise…), sequential (before, after…)[45a].

These connections are also based on the backgroung knowledge of the reader. Five crucial dimensions have been shown to accurately represent the mental representations built by adult readers, at least when they read narratives: time (when), space (where), protagnonist (who), causality (why), and intentionallity (why)[45b]. Experimental evidences in support to these statements can be found in the seminal book published by Bartlett in 1932, in which are also taken into account the requirement for memory and the cultural dimension of text understanding (mainly through several experiments showing how an old North American folk tale is understood by students from England)[45c].

What we know about comprehension also depends on the type of measure. The results of "on line" measures such as the exam of the student's eye movements when s/he is silently reading sentences such as "The old man the boat" versus "The old man drives a caleche" are not the same as the results of "off-line" measures such as responses to questions after the silent reading of the same sentences[46]. And, for large scale assessments, only some types of measure can be used: responses to questions, sentence or text to complete (cloze tests)[47] or to match with picture, retelling or summarizing a story...

Within the present section, it is not possible to develop all the issues just introduced. Only those that should be taken into account for the setting-up of large-scale tests will be presented.

1.6a. Differences due to the learner's status (L1 versus L2 speaker)

There is an important difference between word level skills and reading comprehension in L1 and L2 learners: L2 learners might catch up L1 learners for word-level reading skills (at least for accuracy)[48] while being unable to do so in reading comprehension. Indeed, in sharp contrast with word level reading skills, highly consistent results indicate that reading comprehension performance of L2 learners falls well below that of the L1 peers[49]. The reason for that disparity is the level of L2 oral language proficiency: reading comprehension is compromised when oral language is insufficient to support understanding. De facto, well-developed oral proficiency in L2 (but also in L1) has been found to be associated with comprehension in Lit2[50]. More generally, there are strong relationships between Lit1 and L1, and between Lit2 and L2/L1 spoken language skills[51]. Alternatively, oral language proficiency is not a strong predictor of word-level skills for L2 and L1 learners.

1.6b. Differences due to the type of test

Tests assessing reading comprehension differ greatly such that correlations between different tests are quite low[52]. This suggests that these tests do not assess the same skills. Therefore, the assumption that reading comprehension tests are fairly comparable is not correct. Reading comprehension is not a unitary phenomenon, and thus it should not and cannot appropriately

---

45a. Halliday & Hasan, 1976; see also Graesser et al., 1994; Mann & Thompson, 1986. 45b: Tapiero, 2007 (pp.183-188); 45c: Barlett, 1967 (pp.63-94, 120-129, 171-176).

46. Or "The police arrested the sister of the nursemaid who recently gave birth to twins" versus "The police arrested the brother of the nursemaid who recently gave birth to twins": See Rayner, 1998; Rayner et al. 2001 and 2006.

47. Cloze tests include at least one missing word. Such tests could be used to assess relations between comprehension and different linguistic skills: level of mastery of grapheme-phoneme correspondences ("I knew the weather was cold but I forgot to take my … [coat / goat / shorts]"); and semantic knowledge at a local level (vocabulary: "A king usually lives in a … [castle / house / flat]") or at a more global level ("I knew it was going to rain, but I forgot to take my… [raincoat / coat / pant /]".

48. See Lesaux & Geva, 2006; Lesaux et al., 2006a; Shanahan & Beck, 2006.

49. For reviews, see: Lesaux & Geva, 2006; Lesaux et al., 2006a; Snow & Kang, 2006.

50. See the previous note. See also: Crosson et al., 2008; Droop & Verhoeven, 1998; Verhoeven, 1990 and 2000; Zwaan & Brown, 1996.

51. For L1 learners, see: Chall, 1987; Nation & Snowling, 2004. For L2 learners, see for reviews: Lesaux & Geva, 2006; Lesaux et al., 2006a; see also: August et al., 2005; Carlisle et al. 1999; Francis et al., 2006; Leseman & De Jong, 1998; Proctor et al., 2005 and 2006.

52. Keenan et al., 2008: from .31 to .70 (9 out of 10 inferior to .54). See also Cutting & Scarborough, 2006; Fletcher, 2006.

be measured by a single instrument. However, as suggested by some results (see here below), some reading comprehension tests seem more reliable than others.

The scores obtained after having read aloud a long text depend more on oral language comprehension than on decoding skills, the opposite trend being found after the silent reading of short texts[53a]. As explained in one of these studies[53b], these results are likely to be due to the fact that questions asked after the reading of long texts can be answered by guessing from background knowledge. On the other hand, guessing is less likely to occur in the reading of short texts; therefore, the scores will depend more on decoding abilities.

When silent reading tests are used, it is thus necessary to rely on short texts and questions it is not possible to answer to without having read the test. A solution would be to use cloze tests, with short sentences. This type of test allows the assessment, with similar constraints, of different aspects involved in reading comprehension: from level of mastery of grapheme-phoneme correspondences to levels of vocabulary, morpho-syntax and semantic, for instance. This type of test is frequently used with L2 learners[54], because the correlations between different cloze tests have been found to be high; and performance on cloze tests is significantly correlated with teachers' judgements of L2 learners reading performances[55].

1.6c. The issue of speed in the assessment of reading comprehension

The very rapid activation of words allows freeing useful memory capacity that becomes available for comprehension processes, and individual differences in very rapid word identification account for reading comprehension results[56]. However, one should not conclude that increasing the time to process a text will always improve a reader's comprehension.

This has been investigated directly in a study which has examined the effect of extra time on the reading comprehension performance of four groups of adults with or without reading disabilities (average and above-average readers versus poor and very poor readers). The students were instructed to read at their own pace (untimed condition, but no more than 40 minutes) or in a timed condition (20 minutes)[57]. All of the reading disabled students benefited from extra time, but the normally achieving readers performed similarly under the timed and untimed conditions. In addition, very poor readers increased their scores in the untimed condition, but that condition was not enough to enable them to perform at the same level as average readers, which is the case for the poor readers.

This study thus suggests that pedagogical practices aimed at decreasing the processing time during reading, or those providing extra time during testing, are appropriate to help poor readers begin to compensate for their reading difficulties. It is not the same for very poor readers. Accordingly -- and this seems obvious -- it is not possible to use reading comprehension test with very poor readers or very beginning readers. An assessment of their level of spoken comprehension will be enough. Another solution is to use with very poor readers very short texts (for instance, cloze tests with no more than 10 words).

1.6d. Developmental changes

The role and importance of spoken comprehension and decoding skills in reading comprehension changes according to the developmental moment considered. Amongst the younger readers, reading comprehension is mostly explained by decoding skills. Later in development, reading comprehension is mostly explained by spoken comprehension. Thus as

---

53. 53a: Keenan et al., 2008; Cutting & Scarborough, 2006; See also Francis et al., 2006, for a study with L2 learners; 53b: Keenan et al., 2008.
54. Around 100 quotations of this type of test in August and Shanahan, 2006a. See also Garcia et al., 2006a-b. For some examples, see note 47.
55. Garcia et al., 2006a-b; see also Laesch & Van Kleeck, 1987.
56. Perfetti, 1985; Perfetti et al., 1979; Perfetti & Hogaboam, 1975; Stanovich, 2000; Stanovich et al., 1981; West & Stanovich, 1978;
57. Lesaux et al., 2006b. See also 6.1a

word recognition becomes more fluent and automated, reading comprehension starts being strongly predicted by listening comprehension rather than by decoding skills. Similar results have been observed with L2 learners[58].

When designing reading assessments tests, it is thus crucial to consider these changes in the relations between reading comprehension versus oral comprehension and decoding skills, changes that occur depending on the age and the level of reading expertise of the reader.

1.7. Summary

The findings reported in the present section allow us to draw relatively firm conclusions about reading development at the word level (decoding, written word identification, and word spelling) for L1 learners. This appears to be true across languages, locations and age groups over a variety of measures. In addition, the outcomes of more recent studies investigating L2 predictors of L2 word-level literacy skills are similar to those identified by decades of research on reading acquisition in L1 learners. More specifically, phonological processing (including decoding, phonological awareness and phonological memory) predicts word identification skills in L1 and L2 learners. Furthermore, reading disabled learners in L1 or L2 have poor phonological skills. Also, for both L1 and L2 learners, phonological training is seen as the most effective intervention. However, adjustments to these approaches that are helpful for L1 learners are needed to have maximum benefit with L2 learners: for example, work on the specificities of the language in which they learn to read, especially at the phonological level. Finally, the level of word-level skills of L2 learners (decoding, word identification, word spelling) can attain that of L1 learners, at least with adequate training.

In contrast, reading comprehension in L2 learners has been consistently found to fall well below that of L1. This is due to the fact that reading comprehension is compromised when the level of oral language is not sufficient to support understanding, which is very often the case for L2 learners. In addition, the conclusions to be drawn for the findings of the scattered lines of research on reading comprehension and on the factors that influence this multidimensional skill seem to be much less definitive for both L2 and L1 learners. However, the factors that influence reading comprehension seem to be largely the same for L1 and L2 learners. Some are related to the participant: his/her written word identification skills; his/her spoken language knowledge (at the level of word, sentence, and discourse); his/her cognitive abilities (e.g. memory); his/her cultural background; and his/her goal and motivation when reading. Other factors are related to the text: its complexity at different levels (vocabulary, syntax [word order], and relationships between the different parts of the text); its length; and its nature (for instance, story versus expository texts appearing in encyclopedia). Additional factors are related to the reading modality (silent reading versus reading aloud) and the way reading comprehension is assessed[59].

It is also important to stress a fact often put forward in the previous sections, which is also very often highlighted in reviews of reading acquisition, especially those focusing on L2 learners: "measurement of language minority children's language abilities requires linguistic expertise"[60]. This is also true for measurement of language in L1 learners.

---

58. L1 learners: see Keenan et al., 2008; Sticht & James, 1984; L2 learners: see Verhoeven, 1990 and 2000.
59. One possible means of bridging the scattered approaches on this domain is to setup different cloze tests (or word/sentence-picture matching tasks), with short texts. This type of tests requires an entire reading of the text so that the reader cannot provide an answer after a simple overview of the text. In addition, this type of test can be easily adapted to the reading level of a beginning reader or of a poor reader. And such tests can allow us to evaluate different processes involved in reading comprehension (see note 47) on similar bases and with a low memory load. See for the use of such tests in the context of L2 learners in developing countries: Wagner, 1993 (pp.88-92).
60. Garcia et al. 2006a (p.596).

2. Description and analyses of different reading assessments

2.1. PIRLS: Progress in International Reading Literacy Study

(a) Aim and theoretical background: It is commissioned by the International Association for the Evaluation of Educational Achievement (IEA) and designed to provide a periodic assessment of literacy in fourth grade children. The last assessment cycle conducted in 2006 took place in 35 countries and the next assessment will be in 2011.

PIRLS is based on the theoretical model developed by Kintsch and Van Dijk (1978; see also, Spiro, Bruce & Brewer, 1980) which deals almost exclusively with written comprehension processes. PIRLS does not consider decoding and word identification operations or the relationship between written and oral language comprehension. It is geared towards measuring reading comprehension, defined in terms of four components abilities:

- Focus and retrieve explicitly stated information;
- Make straightforward inferences;
- Interpret and integrate ideas and information;
- Examine and evaluate content, language, and textual elements.

These four processes are described as accounting both for reading for literary experience and to acquire and use information[61], which are assumed to summarise the type of reading activity experienced by fourth graders across the world.

(b) Target population: Fourth grade was chosen because it represents an important stage of reading acquisition at which students are assumed to have acquired basic decoding skills.

(c) Method of assessment and test content: The PIRLS testing battery is administered collectively over a limited time slot of 80 minutes. The entire set of texts is composed of ten reading passages containing literary and informational passages; however each examinee is only assessed over one of each passage type (one story and one informational text on average). Reading comprehension for each passage is assessed by a series of about 12 questions, half of which provide multiple choice responses and the remaining half requires constructed answers. Students read the passages silently and respond individually and in writing to the questions (See Appendix A). An additional 15–30 minutes is allotted to a student questionnaire.

(d) Design of the material: National Research Coordinators (NRC) constituted of representatives for each participant country submitted passages which were then approved by a Reading Development Group once a "general agreement" was met. Text passages were required to comply with the following guidelines:

- Suitable for fourth-grade students in content, interest, and reading ability;
- Well written in terms of depth and complexity to allow questioning across the processes and strategies defined in the PIRLS 2006 framework;
- Sensitive to cultural groups to avoid specific cultural references, wherever possible;
- Continuous and not exceed 1,200 word.

The PIRLS instruments were prepared in English and then translated into 45 languages following a careful verification process. Each country was allowed some freedom in translating passages when it was necessary to accommodate cultural and linguistic specificity.

---

61. For example, reading for literary experience is often accomplished through reading fiction, while reading to acquire and use information is generally associated with informative articles and instructional texts.

The NRC also created the set of questions bearing on each text passage. In designing the question, NRCs were instructed to pay particular attention to matching closely the purpose of the passage and to covering PIRLS component processes while considering timing, potential sources of bias, and ease of translation.

(e) Pre-PIRLS[62]: An early grade version of PIRLS[63] (for children still in the process of learning to read) is currently being developed. Pre-PIRLS relies on the same principles as PIRLS and employs a similar methodology, however reading passages are shorter than those used in PIRLS (around 400 words), with easier vocabulary and syntax. There is also a greater emphasis on the processes of retrieving information and making straightforward inferences, and less weight placed on integrating ideas and evaluating content in pre-PIRLS than in PIRLS.

The methodology for assessing comprehension relies on questions, however in contrast to PIRLS (in which questions are presented after reading the passage), some of the questions in Pre-PIRLS are also interspersed throughout the text (students have thus less text to recall/skim to find answers, and are able to answer some items even if they do not finish the entire passage). This assessment is still being developed and has not yet been administered.

2.2. PISA: Progress in International Reading Literacy Study

(a) Aim and theoretical background: PISA is a triennial survey of scientific, mathematical and reading abilities commissioned by the OECD. The reading subtest is based on the same theoretical premises as PIRLS and the PISA's stated goal is to go beyond simple decoding and literal interpretation of written information by assessing literacy in real life situations. PISA reading assessment defines five processes associated with achieving a full understanding of a text.

- Retrieving information;
- Forming a broad general understanding;
- Developing an interpretation;
- Reflecting on and evaluating the content of a text;
- Reflecting on and evaluating the form of a text.

(b) The target population: PISA's reading subtest aims at assessing abilities in 15 year old students (irrespective of grade) approaching the end of compulsory education in order to measure how well they are prepared to face challenges of today's society by measuring what they do with what they learn at school.

(c) Method of assessment and test content: In an attempt to represent the kinds of reading that occurs both within and outside the classroom, texts were selected within four types of reading contexts: reading for private or public use, reading for work; reading for education. Reading passages were also composed of continuous (narration, reports, argumentation …) and non continuous texts (charts, maps, advertisements …). Around half the questions measuring written comprehension of the passages were open questions (which required an articulated answer) while the remaining consisted in closed questions (yes/no responses or choice between more than 2 alternative answers). In addition, members of the reading expert group and test developers identified processes having an effect on the difficulty of a reading test and needing to be considered when designing of the test and ranged from:

- Making a simple connexion between pieces of information;

---

62. Personal communication to D.Wagner from K. Trong.
63. Grade not specified.

- Hypothesizing about the text;
- The amount of information to retrieve;
- The number of criteria which the information must satisfy;
- The sequencing of the information to be retrieved;
- The amount of text to be assimilated;
- Specificity (or familiarity) of the knowledge that must be drawn out from the text;
- Prominence of the information (how explicitly the reader is directed towards it).

Each student was assessed over two hours of which testing dedicated to reading occupied from 60 to 90 minutes of the total testing time. Different combinations of passages were grouped in nine different assessments booklets which ensured that each was answered by a representative sample of students.

(d) Design of the material: Participating countries responded to a call for submission for sample texts. They were provided with a guideline outlining the purpose of the project and a number of variables such as text types and formats as well as response formats and context. The selection of authentic material in their original form was encouraged, preferably from the news media or original published texts. Once the development team reviewed the documents submitted, a set of items was validated to take part to the assessment. Several texts and items were also created from scratch by test developers. Following the development team decision, items elected to be included were provided in French or in English to translation teams. Translation notes were provided where possible misinterpretation and difficulty might arise. Summary of intention for items was also provided and suggestions were made to indicate the scope of possible change. Finally test developers ensured of the appropriateness of the translation in each country.

2.3. SACMEQ: Southern and Eastern Africa Consortium for Monitoring Educational Quality[64]

(a) Aim and theoretical background: SACMEQ assessment includes the measurement of <u>both</u> reading and mathematics performance levels for <u>both</u> pupils and teachers. In the SACMEQ II Project "reading literacy" was defined as "the ability to understand and use those written language forms required by society and/or valued by the individual". This was the agreed definition that was used by the 35 countries that participated in the International Reading Literacy Study that was conducted by the International Association for the Evaluation of Educational Achievement (i.e., PIRLS, see Elley, 1992). It was also the general definition accepted by the SACMEQ National Research Coordinators (NRCs). The NRCs found this definition to be general enough to accommodate the diversity of traditions and languages represented in the SACMEQ countries, and yet still sufficiently specific to provide guidance for test construction.

(b) The target population: The target population for both the SACMEQ I and SACMEQ II Projects was focused on the Grade 6 level.

(c) Method of assessment and test content: In both SACMEQ Projects there was an initial detailed curriculum analysis undertaken across all countries in order to define – after exhaustive discussion of the most important skills contained within the reading curricula at Grade 6 level - the reading skills that were considered by each country to be the most important. The NRCs decided to accept the three broad content domains for reading literacy that had been adopted for the International Reading Literacy Study:

---

64. IIEP-UNESCO.

- **Narrative prose:** Continuous texts in which the writer aims to tell a story (fact or fiction).
- **Expository prose:** Continuous text in which the writer aims to describe, explain, or otherwise convey factual information or opinion to the reader.
- **Documents:** Structured information organized by the writer in a manner that requires the reader to search, locate, and process selected facts, rather than to read every word of a continuous text.

Intensive examination of curricula was also conducted in order identify descriptive skill levels that would define a recognizable and meaningful dimension.

- **Level 1:** Pupils at this level should be able to link words and pictures where the pictures depict common objects of a "concrete" nature.
- **Level 2:** Pupils at this level should be able to link words to more abstract concepts such as propositions of place and direction, and, perhaps, ideas and concepts such as comparatives and superlatives (happiest, biggest, below, etc.).
- **Level 3:** Pupils at this level should be able to link words (such as a phrase or short sentence) from one setting to words in another setting where there is a word match between the two settings.
- **Level 4:** Pupils at this level should be able to deal with longer passages of text that contain a sequence of ideas and content, and that require understanding derived from an accumulation of information gathered by reading forward.
- **Level 5:** Pupils at this level should be able to read through a text in order to: confirm understanding, or link new information with a piece of information encountered previously, or link ideas from separate parts of a text, or demonstrate the capacity to infer an author's intention.

These dimensions, taken in combination with the three domains of reading, formed a framework (or blueprint) for the construction of suitable test items. Five reading skill levels were identified as shown below (see Appendix B).

(d) Design of the material: An initial detailed curriculum analysis was undertaken across all participating countries in order to define – after comprehensive discussion of the most important skills contained within the reading curricula at Grade 6 level - the reading skills that were considered by all countries to be the most important.

2.4. PASEC: Programme d'Analyse des Systèmes Educatifs des pays de la CONFEMEN[65]

(a) Aim and theoretical background: PASEC aims at observing the basic educational level in the language of instruction (French) and in Math in children enrolled in primary school in African Francophone countries.

(b) The target population includes second and fifth graders (one pre-test at the beginning of each grade, and one post test at the end of each grade).

(c) Method of assessment and test content: The goals of the main subtests are presented in the Appendix C (C1-2: beginning end of the second grade, C2: beginning of the fifth grade). In contrast with all the other batteries of tests, PASEC is largely focussed on the mastery of grammar, especially for grade 5 children (10 sub-tests). Written comprehension is assessed at the level of words, sentences and texts with cloze tests and with word/sentence-picture

---

65. Personal communications: Jean Bourdon (IREDU) and Jim Stevens (World Bank).

matching tasks (5 subtests at the end of G2). Other tests involved phonemic discrimination (3 subtests at the end of G2)

(d) Design of the material: The test content has been created relative to the school program in the test countries and when possible, on the common aspects of the program (see Appendix C).

2.5. EGRA: Early Grade Reading Assessment

EGRA focuses on the start of primary schooling. That instrument aims at assessing early reading ability by measuring letter recognition, simple word reading and written text comprehension. EGRA has been implemented in 19 developing countries. A toolkit was developed that summarises the basic test outline[66].

(a) Aim and theoretical background: EGRA is based on scientific research outcomes accounting for component reading skills and recommendations from the National Reading Panel (2000; Ehri, 2001a-b). The tests are similar to those included in existing test batteries such as DIBELS[67] which aims at assessing emergent literacy skills known to be correlated with reading and for predicting future reading achievement.

(b) Target population: EGRA focuses on reading assessment at the beginning of reading instruction, mainly in Grade 1 to 4 in developing countries (see Appendix D).

(c) Method of assessment and test content: EGRA is composed of a set of measures of reading abilities that goes beyond the assessment of written comprehension. It assesses decoding skills (accuracy and speed), written and oral comprehension as well as reading related skills such as phonemic awareness. Most subtests require students to read aloud and therefore require the intervention of an enumerator. All the reading aloud tasks involve fluency (that is accuracy and speed) measured by the mean of correct items processed in one minute. The different subtasks are (see also Appendix D):

1) Engagement and relationship to print: Indicate where to begin reading and the direction of reading within a line and a page.
2) Phonemic awareness: phonemic segmentation (pronunciation of the different. phonemes of a word containing from 2 to 5 phonemes), or phonemic counting.
3) Fluency in letter recognition (1 minute test): Provide the name (and sometimes the sound) of upper- and lower-case letters distributed in random order.
4) Fluency in pseudoword reading (1 minute test): Use of grapheme-phoneme correspondences to read simple pseudowords.
5) Fluency in word reading (1 minute test): Read simple and frequent one and two syllable words.
6) Fluency in text reading (1 minute test): Read a short text (a story, 60 words).
7) Reading comprehension: Respond to different type of questions (literal, and inferential) about the text they have read
8) Listening comprehension: Respond to different type of questions (similar to those used to assess reading comprehension) about a story told by an adult enumerator
9) Spelling: Write isolated words dictated by an enumerator to assess orthographic skills in spelling through an isolated word spelling dictation exercise. Words can also be inserted in a sentence to assess use of grammar.

Adaptations of the test may be made following the toolkit's guidelines which suggest

---

66. EGRA Toolkit: English version, 2009a; French version, 2009b.
67. Good, & Kaminski, 2002. see also TOWRE, Torgesen et al., 1999 and 2001.

designing subtests respecting letter, grapheme, and word frequency, syllabic structure and letter position in the language. The toolkit also suggests that comprehension subtests should be designed following examples narratives amongst children's text-books, respecting local culture and questions should be fact based and require inference while avoiding yes/no questions. To summarise, EGRA is not based on a straight translation rather it seeks to account for local linguistic specificity and written language constraints such as the transparency of the writing system and cultural content.

2.6. PRATHAM: READ INDIA

(a) Aim and theoretical background: The objective of READ INDIA[68] PRATHAM is to get all of India's primary age children to a basic level of proficiency in reading and math.

(b) Target population: This program targets children from Grade 1 to Grade 5, whose education is provided in Hindi and English.

(c) Method of assessment and test content: PRATHAM's Read India campaign is active in 350 districts across India. The program involves two components: assessing basic reading and math and assessing higher order skills in reading, writing, and math. It uses an already existing test, the ASER test, whose content is aligned to grade 1 and grade 2 level state textbooks for language. The tests assess: basic reading and arithmetic each year. Every year some new subjects/skills are also assessed like English, comprehension, and problem solving. For reading, it assesses whether students:
o Can correctly identify 4 of any 5 randomly selected letters.
o Can correctly read 4 of any 5 randomly selected common words.
o Can read short 4 sentence passages of approximately 19 words at Grade 1 level that the child reads "like she is reading a sentence, rather than a string of words".
o Can read a 7-10 sentence story of approximately 60 words at Grade 2 level "fluently with ease".
o Can orally answer two questions after reading a text.

In addition to that, PRATHAM uses the following subtests from EGRA battery
o Character recognition naming fluency.
o Fluency in word reading.
o Fluency in pseudoword reading.
o Fluency in text reading.
o Reading Comprehension.

Written language tests were also developed to assess:
o Letter Knowledge: letter dictation.
o Word Knowledge: Match picture with word, Select antonym, Label picture.
o Sentence Comprehension (lexical decision task).
o Cloze Sentence – Select correct word to complete sentence.
o Passage Comprehension (factual & inferential): Read two passages and answer questions.
o Writing Ability: Word dictation (spelling); Label pictures; Construct a sentence; Read passages and answer comprehension questions.

---

68. Evaluating READ INDIA: The Development of Tools for Assessing Hindi reading and writing ability and Math Skills, Personal Communication from D. Wagner.

3. Observations on the previously reviewed assessments

The reading assessments illustrated so far differ in numerous aspects of their underlying principles and methodology. To evaluate the relevance of each assessment to evaluate children's reading level, some issues deserve further consideration and discussion.

- PIRLS and PISA (and SACMEQ) are geared towards measuring reading comprehension with a very clear theoretical background. This is an important point. However, there are some flaws in these assessments, which can be illustrated by an example from PIRLS (see Appendix A)[69]. In section 1.6b, it has been suggested that questions asked after the reading of long texts can be answered by guessing from background knowledge. In the example from PIRLS, it is the case for at least 2 questions (Questions 1 and 12, see Appendix A)[69].

  Also, when a student fails to answer a question, the origin of that failure cannot be determined: Was the student unable to entirely read the text as s/he was slow? Was the student, after having read the text, unable to remember its content once presented with the questions? Or was s/he able to read the text but simply elected not to do?

  In addition, PIRLS and PISA (and SACMEQ) do not include a test aimed to assess the level of oral language comprehension that plays an important role in explaining reading comprehension (section 1.6). The outcomes of such assessments are thus difficult to interpret. They are also difficult to translate into educational practices. Therefore, to fully account for students' reading abilities, a reading assessment battery must measure oral language proficiency and this is important with regard to developing countries, where language familiarity in assessments crucially needs to be considered (for a more in depth discussion, see sections 4.2c and 5.1a).

  Furthermore, PIRLS and PISA do not include a test aimed to account for the level of written word identification, which also plays an important role in explaining reading comprehension (section 1.6). Additionally, PIRLS and PISA (and SACMEQ) do not include measurement of decoding skills which are crucial to the development of written word identification (see sections 1.2 to 1.5). This is problematic because beginning readers and those who have not achieved an expert level of reading (due to limited schooling, for instance), rely mostly on decoding skills to read. Therefore a thorough reading assessment battery must measure word level skills (decoding and written word identification) to fully account for students' reading abilities and must include accuracy as well as speed measures (both measures being crucial to assess the level of mastery of word level skills, see sections 1.2 and 1.3).

  All these problems render the interpretation of the outcomes of PIRLS and PISA assessments difficult to translate into educational practices. Indeed, low level of mastery of written word identification skills and oral language skills can separately or concurrently explain reading comprehension difficulties, and pedagogical interventions will not be the same depending on the underlying explanation to poor reading.

  Finally, PIRLS, PISA and SACMEQ assess reading after basic reading skills are supposed acquired. However, reading skills needs to be assessed as early as possible in order to define and implement policy aiming at improving reading education (see section 1).

- PASEC is geared towards measuring reading skills, plus other language skills (grammar, conjugation, for instance, see Appendix C). However, PASEC is not based on a solid frame of reference. This explains the lack of clarity of the objectives of some tests. In addition, PASEC is the only battery that requires an explicit knowledge of grammar

---

69. See also PIRLS, 2004, Appendix B (pp.81-92) http://www.skolverket.se/content/1/c4/03/71/ramverk_pirls_2006.pdf.

(e.g. children must provide the grammatical categories of the words within a sentence). To our knowledge, there is no result of research showing that an explicit knowledge of the grammatical terminology should help L1 or L2 learning, at least in primary school. Alternatively, none of the tests in PASEC adequately assess the mastery of grapheme-phoneme correspondences. This is a serious shortcoming (see Section 1.2).

In PASEC (and SACMEQ), reading comprehension is assessed at different levels (words, sentences and texts) but listening comprehension is not assessed. Therefore, when a child has a problem in reading comprehension, it is not possible to determine its origin. Nonetheless, if there are gaps in PASEC, some tests from that battery correspond to those often used to assess language proficiency in L2 learners: word/sentence-picture matching tasks and cloze tests used to assess comprehension at the level of words, sentences and texts[70a]. These tests present a double advantage: a collective administration is possible (which saves time), and they can be timed (for example, stop after two-three minutes for each sub-test)[70b].

- EGRA is the sole tool which has explicitly taken into account the results of the current scientific literature. In addition, at the difference of PIRLS and PISA, EGRA is a tool that can be used with beginning readers (as PRATHAM and PASEC), which is also very important. However, some improvements to EGRA are still needed, especially for young L2 learners from developing countries, as suggested by the results presented in Appendix E: scores obtained at the end of Grade 1 by Gambian students having learnt to read in English (L2); by same grade Senegalese students having learnt to read in French (L2) or in Wolof (L1); and by Nigerian students assessed before reading instruction.

  Senegalese and Gambian students read a small number of words and the scores of a high proportion of first graders are at the floor level in the word in context reading task: zero word read by 80%, 50% and 71% of respectively L2-English, L2-French, and L1-Wolof students[71]. These results prevent us from examining reading comprehension skills.

  For spoken comprehension, floor effects were also massive for L2 learners (around 60% in French and in English), not for L1 learners (26% in Wolof) and, as expected, the Senegalese-L1 learners surpassed the Senegalese-L2 learners. In the same task, the Gambian-L2 (English) surpassed Senegalese-L2 (French) learners. To determine the possible origin of that difference, it is necessary to know, at least, the vocabulary level of the children in the language in which they are learning to read, which has not been assessed in the Gambian and Senegalese studies. A test assessing vocabulary level has been developed and used in Niger. As can be seen in Appendix E, this test avoids floor effects (mean percentage of correct responses: 69%, SD=15). It seems thus possible to include this test in the protocol[72].

  The results of the other tests involving spoken language, especially those of the phonemic awareness tasks, are weak and greatly vary between subjects (see the standard deviations). This could be due to the difficulty of the task not only for the student but also for the enumerator who may have problems to code the results because s/he is probably not well

---

70. 70a: For L2 learners, see 1.6b; see also notes 47 and 59; for the use of such tests in the context of developing countries, see Wagner, 1993 (pp.88-92). 70b: Concerning the importance of a short testing duration, see Wagner, 2003.

71. See Sprenger-Charolles, 2008a-b. In a paper published by Abadzi (2008, Table 2), the scores of L1 (Wolof) and L2 (French) Senegalese students are presented together, thus hiding the fact that the scores of L2 learners are not lower than those of L1 learners (see Appendix E). In addition, in the paper by Abadzi, only the scores of the word in context task are presented, not those of the isolated word reading task, which hides another important result: the difference between the reading of isolated word and word in context is very large for the L2 learners of English, not for the L2 learners of French (see also Appendix E). That result once again highlight the fact that there are qualitative differences between the results of students learning to read in an opaque orthography (English) and those learning to read in a quite transparent orthography (French).

72. This test is also very rapid and easy to administer: Children are asked to do something (for instance, to show a part of their body or items from the environment; to yawn or to smile; to put a pen under a book or above a book …. Vocabulary level can also be assessed by word-picture matching tasks: see SACMEQ, PASEC and PRATHAM (sections 2-3); see also Wagner, 1993 (pp.88-92).

aware of the sound pattern of the language in which the testing is done (see here above, 1.2b). As this type of assessment is very important (see 1.2b), a solution has to be found[73]. In addition to the test currently used in EGRA, a test aimed to assess phonological short-term memory has been developed and assessed in the Nigerian study (see Appendix E). The results indicate a high level of accuracy (correct responses: 63%, SD=27). It seems thus possible to include this test in EGRA which can help to discriminate children who struggle to learn to read in L1 or L2 (or to learn spoken vocabulary) because of poor phonological short-term memory.

To summarize, the design of EGRA subtests need to be revised and adapted to the reading level of first graders. Especially, it is not necessary to assess reading comprehension with first graders having a low level of word reading (assessment of spoken comprehension will be enough, but a more in depth assessment than the one implemented currently in EGRA[74]). Floor effects could be avoided by starting the task with very simple and frequent words, and with very short texts (as in SACMEQ, PASEC or PRATHAM, for instance).

Despite the problems here above highlighted for EGRA (that can be solved relatively easily, as we have suggested), that battery of test is currently the only that can help to identify very early and very precisely the origin of the children's reading difficulties, and thus that could guide educational policies.

Other issues about international test batteries are presented in the following sections: Linguistic constraints on assessment methods (Section 4); and other factors affecting test administration (Section 5); and implications for teaching and education (Section 6).

4. Linguistic constraints on assessment methods

4.1. Reading assessments across different languages: Translation

Important efforts in the design and development of international reading assessment are devoted to the adaptation of the test's material to user countries, particularly with respect to the language. However, the methodology employed to adapt assessments across languages differs across the different tools.

For PISA and PIRLS the material is first developed in a single or a few languages, and then translated in the remaining languages. This process is intended to safeguard against differences in text difficulty across countries and languages. It is also motivated by the need for cross country comparison that these assessments aim at fulfilling. However such methodology is not without disadvantages. Implicit in this design, is the fact that the level and age appropriateness is adequately judged and set by the test developers and advisors in the initial phase. As discussed in section 1, the age and difficulty level of a reading passage depends on language specific properties (such as the transparency of the writing system). This cannot be accounted for by subjective judgements from the tests' designers. In addition, translation is a particularly ill suited method to account for the linguistic specificities of each language of administration because translation intends to preserve almost only meaning. The following sections discuss the methodological constraints that should be taken into account to assess the component reading skills, and the necessity to rely on objective methods

---

73. See also 1.2b and note 15c. A solution will be to use a task avoiding the production on the phoneme by the children and the production of the word by the enumerator: for instance, a task with 3 pictures, the name of the object represented on two of these pictures beginning (or ending) with the same phoneme, the children having to designate the "odd" picture, the one that does not have the same beginning (or ending) than the others two (for instance, the pictures of "a bat", "a cat", and "a car"); or a cloze test with a choice between 3 words, 2 beginning by a very similar phoneme as /k/g/, /p/t/, /b/d/… (I knew the weather was cold but I forgot to take my … [coat / goat / shorts]).
74. See for some suggestions: Bianco et al., in press.

sensitive to age, and specificities of the spoken and written languages.

4.2. Oral and written language specificity

The processes involved in learning to read are constrained by the writing system and the language status in which reading instruction is provided (section 1.5). In designing tasks measuring word level reading skills (phonological and orthographic reading routes, see section 1.2 and 1.3) and reading comprehension (section 1.6), consideration needs to be given to the linguistic characteristics of the items. Recommendations are formulated in the following sections, with particular attention to low developing countries.

4.2a. Phonology

EGRA and PRATHAM involve tasks such pseudoword reading. The goal of such tasks is to assess decoding skills, which develop as beginning readers learn how minimal written units (graphemes) match onto minimal oral units (phonemes). The written material used in these assessments should thus be correctly decoded by the students.

On the one hand, many L2 English learners will find it hard to pronounce 'th', which corresponds to two phonemes (as in "the" and "this") that do not exist in French, for instance; French speakers will tend to mispronounce "the" and "this" in a similar manner, a little bit like the 'v' in the French word "velo". When an item (a word or a pseudoword) starting with 'th' is mispronounced, it will not be possible to determine if this is so because the student cannot <u>read</u> and relate the grapheme 'th' to the corresponding phoneme or if s/he is unable to <u>pronounce</u> it. Identifying items including phoneme such as 'th' or 'r' in English (that are likely to be misproduced by L2 speakers such as L1 French students), is central in assessing reading in L2. Similarly, only consonant-vowel syllables exist in some languages, for instance, Bamanankan spoken notably in Mali. Thus, Bamanankan L1 speakers can vary greatly in their pronunciation of consonant-consonant-vowel syllables in L2 (French) either because they find it hard to read, or to pronounce, a word such as "dru".

On the other hand phonological constraints might undergo dialectal variations, especially in developing countries in which the language of instruction (LOI) is not the L1. For instance, the same English word in Ghana or in India will not be pronounced similarly, because L2 speakers have learned a different L1 from birth, which will influence their pronunciation of English differently.

It is thus necessary to adapt the assessment's material in a given LOI according to possible interferences between languages coexisting in a given context which are likely to lead to reading difficulties. The knowledge of the overlap and differences in the phonological systems coexisting in the context of the assessment is necessary to build tests, and to code the oral responses of the students. Taking into account these issues is vital to assess decoding skills, but also phonemic awareness, in optimal condition. Therefore, batteries such as EGRA and PRATHAM need further development at the scale of each country and, for some multilingual low developing countries, at the regional scale.

4.2b. Orthography

Another aspect of the measurement of low level reading skills that is commonly overlooked in international assessments involves the difference between letters and graphemes (see also 6.1b). More precisely, graphemes can have more than one letter (for instance 'oo' in "boot"). In the most studied languages such as French or English, tables providing the frequency of letters, graphemes and syllables are available[75a], as well as frequency tables for words[75b].

---

75. 75a: English, see Fry, 2004; French see Peereman et al., 2007. About syllables, see in note 28 a result that supports the fact that, in order to take into account the linguistic specificities of the different languages, it is better to calculate the number of syllables correctly read in word reading test. 75b: English, see Zeno et al., 1995; French: Lété et al., 2004.

These tables can be used to set-up tests. Other orthographic characteristics of languages should be taken into account: for instance, the fact that letters or group of letters cannot be found at the beginning of a word ('tl' and 'dl', in French), or at its end ('v').

For tests adapted to developing countries, involving languages that are studied less frequently, it is crucial to investigate and document the frequency of these orthographical patterns.

4.2c. Language familiarity

A characteristic of PISA, PIRLS and SACMEQ is their focus on reading comprehension, a capacity that not only depends on processes specific to print (decoding and automatic word identification), but also on oral language skills. Results obtained by L2 students on PISA, PIRLS and SACMEQ may be explained by their spoken language abilities, which are not assessed. This has major implication for developing countries where the language of instruction (LOI) is often not the L1 spoken at home. In this context, lack of familiarity with the LOI will have a negative impact on performance on tests designed in that language.

In addition, some L2 speakers will find it hard to understand a written text and answer questions in L2, although they might well be able to understand a written text in their own native language. And such individuals should fail to understand and answer questions to PISA, PIRLS, and thus be wrongly considered as illiterate.

To account for reading comprehension in developing countries, it is critical that international reading assessment be designed to measure separately processes that are specific to print (decoding and automatic word identification) and oral language comprehension (at the level of words, sentences and texts), if possible in L1 and L2.

4.3. The issue of the age / grade of assessment

An important implication of multilingualism in developing countries concerns the age and grade at which reading in the LOI is introduced. In Kenya for example, some children are taught to read and write in their native language (Kisswahili) for a number of years before being instructed to read in English while others will start reading instruction in English from the start[76]. When tested in English, language competency will differ amongst Kenyan children of the same age, as reading level in L2.

In addition, some international assessments are based on grade (PIRLS, SACMEQ, EGRA), others on age (PISA). This could lead to differences between countries where grade repetition is scarcely used (as in Finland) and countries where grade repetition is a current practice (as in France: a high number of 15-year-old students are in Grade 9, and not in grade 10)[77].

Another important aspect of the age of assessment -- the changes through development in the relations between reading comprehension versus oral language comprehension and decoding skills -- is discussed in section 6.2 and 7.

5. Other factors affecting test administration

5.1. Tasks

5.1a. Silent reading vs. Reading aloud

Methodology for measuring reading skills can be broadly divided up into tests relying on

---

76. Commeyras & Inyega, 2007.
77. At age 15, almost 40% of the French students have repeated a grade and, as often observed, retention has a negative impact on achievement. See the French results from PISA 2000 on the website of the French Ministry of Education (http://educ-eval.education.fr/pdf/ni0152.pdf): Note d'information 01-52. Les élèves de 15 ans: Premiers résulltats d'une evaluation internationale des acquis des élèves (PISA).

silent reading versus reading aloud. This methodological difference is not trivial. In silent assessments, reading skills are mainly inferred from answers to questions at the end of the test (see an example from PIRLS in Appendix A). However, silent reading does not allow knowing what exactly the learner has read and leaves a certain number of important questions open:

- Was the student able to read the text but unable to understand the questions?
- Was the student able to read the text but unable to remember its content once presented with the questions?
- Was the student able to read some of the text but not entirely as s/he was slow?
- Was the student able to read the text but simply elected not to?

Such questions need to be considered separately[78]. In beginning or poor literate students, it is important to have reliable measures of reading skills. In addition, especially (but not only) in developing countries many learners will neither be fluent readers nor completely illiterate. Assessments must thus be sensitive enough to locate examinees along a continuum, for both word and text level reading skills.

As suggested earlier (1.6b), when silent reading tests are used to assess comprehension with young or poor readers, it is better to rely on short texts and questions it is not possible to answer to without having read the test. The use of cloze tests (or of word/sentence-picture matching tasks) allows the assessment of different aspects involved in reading comprehension at different levels (vocabulary, sentence and text) and with similar cognitive and linguistic constraints. These types of test are frequently (and successfully) used with L2 learners[79]. Word-level reading skills should also be assessed, and the measures should take into account both accuracy and speed, which is easier with reading aloud than with silent reading tasks, at least in large scale assessments (see 1.3b and 1.6b).

5.1b. Text Length: memory and attention

Another caveat affecting data collection is linked to the test duration[80]. The assessments reviewed in section 2 report testing times that span from 15 minutes (EGRA) to about 2 hours (PIRLS: 2 texts followed by a dozen of questions, see Appendix A). Therefore, for PIRLS, the final reading score will be a composite of successful ability to decode and understand what is written, and of the ability to remember what has been read. Long testing sessions necessitate also a great deal of attention, which is another factor that can bias the results.

5.2. Training and Learner

Robust correlations between phonemic awareness and decoding skills have been shown in beginning readers (see section 1.2), and phonemic awareness skills are reliable predictors of future reading achievement (see section 1.2b). In contrast, some results based on DIBELS or EGRA report weak correlation between phonemic awareness tasks and tasks assessing decoding skills and, more generally, they emphasize the lack of reliability of these two tasks[81]. These differences could be due to the fact that the enumerators involved in experimental research are most often well trained to deliver these two tasks and encode the results. Alternatively, the enumerators involved in DIBELS or EGRA studies might be less aware of some subtle characteristics of the sound pattern of the language in which the testing is done. This is suggested by the fact that, as noticed in reports, they often complain about the fact that these tasks are difficult to deliver, and their results not easy to code[80]. This constitutes a limitation of EGRA's methodology. As yet explained, as these two types of assessment are

---

78. See for a discussion: Siegel & Heaven, 1986.
79. See section 1.6b, see also Leasch, & Van Kleeck, 1987; for developing countries, see PASEC and PRATHAM (sections 2-3), and Wagner, 1993 (pp.88-92). These tests can be collectively administered, which saves time. This is an important point (see Wagner, 2003).
80. Concerning the importance of a short testing duration, see for instance, Wagner, 2003.
81. For DIBELS (Good, & Kaminski, 2002), see Samuels, 2007; For EGRA, see Sprenger-Charolles, 2008a-b. See also 1.2b and note 15c.

very important, a solution has to be found[82].

Another element that might contribute to poor assessment reliability is the reliance on the manipulation of stop watch by the enumerator. Such measurement is well known for increasing potential errors in turning on and off the device, and reporting the test duration[83].

6. Implications for teaching and educational policies

Reading assessment instruments offer strategic opportunities to address and implement change in teaching and in educational policies. However, their relevance depends on a variety of factors. Efficient reading assessment instruments should examine the key abilities involved in reading in order to ascertain what are those that are mastered and those that are lacking in order to guide teaching and general educational policies.

6.1. Impact of the methodology of testing on teaching practices

6.1a. Does using stopwatch force 'word calling'?

The use of short timed tests in batteries such as EGRA and PRATHAM requires the use of a stopwatch. This could potentially encourage teachers to push students to read quickly but inaccurately, while a slower pace would have avoided mistakes and lead to a more accurate understanding of the text.

This issue has been investigated in two studies in which young students who were asked to read carefully matched passages with instructions insisting on speed or on accuracy[84a-b]. The results of these two studies indicated that students increased their reading fluency significantly under the speed condition. They differed in regard to increase in the number of errors: students making more errors in the speed condition in one study, not in the other. In another study with four groups of adults of different reading levels (see 1.6c), it has been observed that poor and very poor readers benefited from extra time, not average and above average readers. Also, very poor readers increased their scores in the untimed condition, but that condition was not enough to enable them to perform at the same level as average readers, which is the case for poor readers. Providing extra time seems thus to help only poor readers. These findings suggest that students can modify their reading rate, sometimes without change in accuracy; however they do not always do so; and asking students to read for understanding may encourage a slower reading rate, sometimes without improvement in comprehension.

6.1b. Does letter naming lead to shortcomings in learning to read?

Tests involving letter naming might encourage policy makers and teachers to orient teaching practices towards alphabet learning, at the expense of letter sounds. This can lead to great confusion in learners because letter-names differ from letter-sounds (e.g. the name of the letter 'z' is /zed/ or /dzi/, and its sound is /zzzz/) and using the name of the letters, the word "dot" will be pronounced "dioti".

As yet explained, it is letter sounds, or more precisely grapheme-phoneme correspondences, which need to be stressed during instruction, and thus to be assessed[85]. Consequently, as in DIBELS, letter-name knowledge should not be included in reading assessments after the

---

82. See note 73 for some suggestions.
83. Research nowadays relies extensively on computerised and automatic recording/reporting of items response latency (see for instance Olson et al., 1994) and therefore field studies using stop watches might be less sensitive tools to measuring some subtle reading processes reported in research.
84. 81a: Hudson et al., 2009; 81b: Colon, & Kranzler, 2006.
85. See 1.2a; For DIBELS, see Good, & Kaminski, 2002. Nevertheless, it is important to assess whether children are able to recognize letters. A visual task, aimed to assess the ability to discriminate letters, should thus be set-up, as yet suggested (See Sprenger-Charolles, 2008a-b). In addition, in the context of developing countries, an assessment of reading skills should include a rapid screening of the child's vision.

beginning of the first grade (as in EGRA and PRATHAM).

6.2. The need for comprehensive reading assessments

In the very beginning, reading comprehension depends more on written word identification than on the mastery of the spoken language. This is because children are learning to decode and identify words, so word-reading processes can limit comprehension in both L1 and L2 learners. However, as children move beyond the beginnings of learning to read, the correlations between reading and spoken language comprehension increase and oral language mastery (included vocabulary mastery) becomes the most important factor explaining reading comprehension in L2 and L1.

International reading tests need to allow explaining the origin of reading difficulties and, in order to implement efficient reading policies and encourage good educational practice, reading tests must be comprehensive.

It is not the case for test relying on written comprehension only (PIRLS, PISA) as they do not allow us to determine whether reading comprehension difficulties are the results of poor decoding skills, poor oral language skills or cognitive difficulties. Alternatively, reading assessments which mainly measure decoding skills and word identification (EGRA) tend not to take into account the different aspects involved in reading comprehension and to underestimate the importance of controlling for linguistic factors in assessments.

7. Final conclusion

A suitable battery of reading tests must meet the following requirements:
- It should take into account the results of the scientific literature.
- It should take into account the linguistic factors in the development of tests, especially in the context of L2.
- It should allow the assessment of reading skills at different ages and levels of expertise, from the beginning of reading acquisition. With beginning readers or readers with a low level of reading abilities, decoding skills and written word identification need to be primarily assessed (together with phonemic awareness and phonological memory), while oral comprehension and vocabulary level, crucial to reading comprehension, need to be measured too. For students who have reached a sufficient level in word identification skills, reading comprehension should be assessed (with different tests), along with oral comprehension and written word identification skills (especially at the level of speed, not only at the level of accuracy).
- Such assessments should rely on different types of tests: Silent reading tests (especially to assess word, sentence, and text comprehension), because they can be collectively administered (this saves time); and reading aloud tests (especially to assess word level reading skills), because the outcomes of such tests are easier to interpret than those of silent reading tests (we know exactly what has been read) and because it is easier to take into account both accuracy and speed with reading aloud than with silent reading tests, at least in large scale assessments.
- Last, and especially in the context of developing countries, the visual abilities of the students should be examined.

REFERENCES

Abadzi, H. (2008). Efficient learning for the poor: New insight into literacy acquisition for children. *International Review of Education*, 54, 24p.

Adams, M. (1990). Learning to read: Thinking and learning about print. Camridge, MA: MIT Press.

Alegria, J., & Mousty, P. (1996). The development of spelling procedures in French-speaking, normal and reading-disabled children: Effects of frequency and lexicality. *Journal of Experimental Child Psychology, 63*(2), 312-338.

August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20(1),* 50–57

August, D. & Shanahan, T. (Eds.) (2006a). *Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth.* Mahwah (NJ), Lawrence Erlbaum associates. 669p.

August, D. & Shanahan, T. (Eds.) (2006b). *Executive summary: Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth.* Mahwah (NJ), Lawrence Erlbaum associates, 9 pages

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158–173.

Bartlett, F.C. (1967). *Remebering: A study in experimental and social psychology.* Cambridge University Press, 317p (First edition: 1932).

Bianco, M., Bressoux, P., Doyen, A.L., Lambert, E., Lima, L., Pellenq, C., & Zorman, M. (in press). Early training in oral comprehension and phonological skills: Results of a three-year longitudinal study. *Scientific Studies of Reading*.

Bishop, D.W.M. (1997). U*ncommon understanding: Development and disorders of language comprehension in children*. Psychology Press Hove, West Sussex

Bradley, L., & Bryant, P. (1983). Categorizing sounds in learning to read: A causal connection. *Nature, 301*, 419-421.

Bruck, M., Genesee, F., & Caravolas, M. (1997). A cross linguistic study of early literacy acquisition. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 145-162). Mahwah (NJ), Lawrence Erlbaum associates.

Bryson, S. E., & Werker, J. F. (1989). Toward understanding the problem in severely disabled readers. Part I: Vowel errors. *Applied Psycholinguistics*, 10, 1-12.

Cain, K. & Oakhill, J. (Eds.) (2007). *Children comprehension problems in oral and written language: A cognitive perspective*. NY, Guilford Press

Caravolas, M., Volín, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology*, 92, 107-139.

Carlisle, J, Beeman, L. Davis, L. & Spharim, G. (1999). Relationship of metalinguistic capabilities and reading achievement for children who are becomming bilingual. *Applied Psycholinguistics*, 20, 459-478.

Carver, R. P. (1998). Predicting reading level in grades 1 to 6 from listening level and decoding level: Testing theory relevant to the simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 10*, 121-154.

Casalis, S. & Louis-Alexandre, M.F. (2000). Morphological analysis, phonological analysis and learning to read French: A longitudinal study. *Reading and Writing*, 12(3-4), 303-335.

Casalis, S. & Colé, P. (2009). On the relationship between morphological and phonological awareness: Effects of training in kindergarten and in first-grade reading. *First Language*, 29(1), 113-145.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91*, 77-111

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, 49, 278-293.

Chall, J. (1987). Two vocabularies for reading: Recognition and meaning. In M.G. McKeown, & M.E. Curtis (Eds). *The nature of vocabulary acquisition* (pp. 7–17). Hillsdale, NJ: Erlbaum.

Chiappe, P. & Siegel, L.S. (1999). Phonological awareness and reading acquisition in English and Punjabi-speaking Canadian children. *Journal of Educational Psychology*, 91, 20-28;

Chiappe, P., Siegel, L. S., & Gottardo, A. (2002a). Reading-related skills of kindergartners from diverse linguistic backgrounds. *Applied Psycholinguistics*, 23, 95-116

Chiappe, P., Siegel, L., & Wade-Woolley, L. (2002b). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading, 6*(4), 369-400.

Chiappe, P., Stringer, N., Siegel, L. S., & Stanovich, K. (2002c). Why the timing deficit hypothesis does not explain reading disability in adults. *Reading and Writing: An Interdisciplinary Journal, 15*, 73-107.

Colé, P., Magnan, A., & Grainger, J. (1999). Syllable-sized units in visual word recognition: Evidence from skilled and beginning readers. *Applied Psycholinguistics, 20*, 507-32.

Colón, E., & Kranzler, J. H. (2006). Effect of instructions on curriculum-based measurement of reading. *Journal of Psychoeducational Assessment, 24*, 318-328.

Coulmas, F. (2003). *Writing Systems: An introduction to their linguistic analysis.* Cambridge University Press 2003, 292p.

Commeyras, M., & Inyega, H. N. (2007). An integrative review of teaching reading in Kenyan primary schools. *Reading Research Quarterly*, 42(2), 258-281.

Cossu, G., Shankweiler, D., Liberman, I. Y., & Gugliotta, M. (1995). Visual and phonological determinants of misreadings in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal, 7*, 235-256.

Crosson, A.C., Lesaux, N.K., & Martiniello, M. (2008). Factors that influence comprehension of connectives among language minority children from Spanish-speaking backgrounds. *Applied Psycholinguistics*, 29, 603-625.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, 10, 277–299.

Dehaene, S., Le Clec, H. G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *Neuroreport, 13*(3), 321-325.

Démonet, J.-F., Taylor, M. J., & Chaix, Y. (2004). Developmental dyslexia. *The Lancet,* 363, 1451-1460.

Droop, M., & Verhoeven, L. (1998). Background knowledge, linguistic complexity, and second-language reading comprehension. *Journal of Literacy Research, 30,* 253-271

EGRA Toolkit [English version] (2009a). Early Grade Reading Assessment Toolkit https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=149

EGRA Toolkit [French version] (2009b). Manuel pour l'évaluation des compétences fondamentales en lecture [Early Grade Reading Assessment Toolkit, French adaptation] https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&ID=175

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001a). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta- analysis. *Review of Educational Research*, 71(3), 393-447.

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001b). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, 36(3), 250—287

Ehri, L., & Wilce, L. S. (1985). Movement into reading: Is the first stage of printed word learning visual or phonetic? *Reading Research Quarterly*, *20,* 163-179.

Elley, W.B. (1992). How in the world do students read? The Hague, Netherlands, IEA [The International Association for the Evaluation of Educational Achievement]

Everatt, J, Smythe, I, Adams, E, & Ocampo, D (2000). Dyslexia screening measures and bilingualism. *Dyslexia: An International Journal of Research and Practice*, 6, 42-56.

Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society, 31*, 119-122

Fletcher, J.D. (2006). Measuring reading comprehension. *Scientific Studies of Reading*, 10 (3), 327-330

Foorman, B.R., Novy, D.M., Francis, D.J., & Liberman, D. (1991). How letter-sound instruction mediates progress in first grade reading and spelling. *Journal of Educational Psychology*, *83*, 456-469.

Fowler, C. A., Shankweiler, D., & Liberman, I. (1979). Apprehending spelling pattern for vowels: A developmental study. *Language and Speech, 22*(22), 243-251.

Francis, D.J., Snow, C.E., August, D., Carlson, C.D., Miller, J., & Iglesias, A. (2006). Measures of Reading Comprehension: A latent variable analysis of the diagnostic assessment of reading comprehension Spanish children. *Scientific Studies of Reading*, 10 (3), 301-322

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*(1), 71-99.

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading*, 2(1), 31-54

Fry E., *(2004). Phonics: A large phoneme-grapheme frequency count revised. Journal of Literacy Research, 36, 85-98.*

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

Garcia, G.E., McKoon, G., & August, D. (2006a). Synthesis: Language and literacy assessments. In D. August & T. Shanahan (Eds.) *Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 19, pp.583-696). Lawrence Erlbaum associates (LEA)

Garcia, G.E., McKoon, G., & August, D. (2006b). Synthesis: Language and literacy assessments. In D. August & T. Shanahan (Eds.) *Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 20, pp.597-625). Lawrence Erlbaum associates (LEA)

Genesee, F. (2006). Cross linguistic relationships in working memory, phonological processes and oral language. In D. August & T. Shanahan (Eds.) *Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 7, pp.175-184). Lawrence Erlbaum associates (LEA)

Genesee, F. & Geva, E. (2006). Cross linguistic relationships in working memory, phonological processes and oral language. In D. August & T. Shanahan (Eds.) *Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 7, pp.175-184). Lawrence Erlbaum associates.

Genesee, F., Geva, E., Dresler, C. & Kamil, M.L. (2006). Synthesis: Cross linguistic relationships. In D. August & T. Shanahan (Eds.) *Developing literacy in lecond language learners. Report of the National Reading Panel on language minority and youth* (Chapter 6, pp.153-173). Lawrence Erlbaum associates (LEA)

Geva, E. (2006). Second language oral proficiency and second language learners. In D. August & T. Shanahan (Eds.) *Developing Literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 5, pp.123-140). Lawrence Erlbaum associates (LEA)

Geva, E. & Genesee, F. (2006). Oral proficiency and second language literacy. In D. August & T. Shanahan (Eds.) *Developing Literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 8, pp.185-196). Lawrence Erlbaum associates (LEA)

Geva, E., & Siegel, L. S. (2000). Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Reading and Writing: An Interdisciplinary Journal, 12,* 1-30.

Good, R. H., & Kaminski, R. A. (Eds.). (2002). *Dynamic Indicators of Basic Early Literacy Skills / DIBELS* (6th ed.). Eugene, OR: Institute for the Development of Education Achievement.

Goodman, K. (1967). Reading: A psycholinguistic guess game. *Journal of the Reading Specialist, May*, 126-135

Goswami, U., Gombert, J. E., & Barrera, L. F. (1998). Children's orthographic representations and linguistic transparency: Nonsense word reading in English, French and Spanish. *Applied Psycholinguistics, 19*, 19-52.

Goswami, U., Ziegler, J. C., Dalton, L., & Schneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics*, 24, 235-247

Graesser, A.C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 3, 371-395.

Grainger, J., & Ferrand, I. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language, 35*, 623-647.

Hallliday, M.A.K. & Hazan, R. (1976). *Cohesion in English*, London, Longman. 374p.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.

Hudson, R.F., Pullen, P.C., Lane, H.B., & Torgesen, J.K. (2009). The complex nature of reading fluency: A multidimensional view. *Reading & Writing Quarterly,* 25(1), 4-32.

IIALC [International Institute of African Languages and Cultures] (1930). *Practical Orthography of African Languages.* London, Oxford University Press

Jiménez, J.E., Rodríguez, C., & Ramírez, G. (2009). Spanish developmental dyslexia: Prevalence, cognitive profile and home literacy experiences. *Journal of Experimental Child Psychology.*

Joshi, R. M. & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology, 21*, 85-97.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first to fourth grades. *Journal of Educational Psychology*, 80(4), 437-447.

Keenan, J. M., Betjemann, R. S., & Olson, R.K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, 12, 281 - 300.

Kintsch, W. & Van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review,* 85, 363-394.

Laberge, D., & Samuels, S. J. (1974*).* Towards a theory of automatic information processing in reading. *Cognitive Psychology*, 6, 293-323.

Labov, W. (1995). Can reading failure be reversed? A linguistic approach to the question. In V. Gadsden and D. Wagner (Eds.), *Literacy among African-American youth: Issues in learning, teaching and schooling* (pp. 39–68). Cresskill, NJ: Hampton Press.

Leasch, K.B., & Van Kleeck, A. (1987). The cloze test as an alternative measure of language proficiency for children considerered for exit from bilingual education programs. *Language learning*, 37(2), 171-189.

Laing, E., & Hulme, C. (1999). Phonological and semantic processes influence beginning readers? ability to learn to read words. *Journal of Experimental Child Psychology*, 73, 183-207.

Landerl, K. & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology*, 100(1), 150-161.

Lesaux, N.K. & Geva, E. (2006). Synthesis: Development of literacy in language minority students. In D. August & T. Shanahan (Eds.) *Developing Literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 3, pp.53-74). Lawrence Erlbaum associates (LEA)

Lesaux, N.K., Koda, K., Siegel, L., & Shanahan, T. (2006a). Development of literacy in second language learners. In D. August & T. Shanahan (Eds.) *Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 4, pp.75-122). Lawrence Erlbaum associates (LEA)

Lesaux, N.K., Pearson, M.R., & Siegel, L.S. (2006b). The effects of timed and untimed testing conditions on the reading comprehension performance of adults with reading disabilities, *Reading and Writing*, 19(1), 21-48.

Lesaux, N.K., Rupp, A.A., & Siegel, L.S. (2007). Growth in reading skills of children from diverse linguistic backgrounds: Findings from a 5-year longitudinal study. *Journal of Educational Psychology*, 99, 821-834.

Lesaux, N.K. & Siegel, L.S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, 39 (6), 1005-1019

Leseman, P.P.M. & De Jong, P.F. (1998). Home literacy: Opportunity, instruction, cooperation and social-emotional quality predicting early reading achievement. *Reading Research Quarterly*, 33, 294-318

Lété, B., Peereman, R., & Fayol, M. (2008). Consistency and word-frequency effects on spelling among first- to fifth-grade French children: A regression-based study. Journal of Memory and Language, 58, 4, 952-977.

Lété, B., Sprenger-Charolles, L., & Colé, P. (2004). MANULEX: A lexical database from French readers. *Behavioral Research Methods, Instruments and Computers, 36(1),* 156-166.

Lundberg, I., Frost, J., & Peterson, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschooll children. *Reading Research Quarterly, 23*, 263-283.

Mann, W.C. & Thompson, S.A. (1986). Relational propositions in discourse. *Discourse Processes*, 9, 57-90.

Martinet, C., Valdois S. & M. Fayol, M. (2004). Lexical knowledge develops from the beginning of literacy acquisition, *Cognition*, 91, 2, B11-B22

Morais, J. (1987). Segmental analysis of speech and its relation to reading ability. *Annals of dyslexia*, 37, 126-141.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*, 323-333.

Morton, J. (1989). An information-processing account of reading acquisition. In A. M. Galaburda (Ed.), *From reading to neurons* (pp. 43-66). Cambridge, Mass.: Bradford Book, M.I.T. Press.

Nation, K., & Snowling, M.J. (2004). Beyond phonological skills: broader language skills contribute to the development of reading. *Journal of Research in Reading*, 27, 342-356

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups (NIH Publication No. 00-4754).* http://www.nichd.nih.gov/publications/nrp/upload/report_pdf

Olson, R. K., Forsberg, H., Wise, B., & Rack, J. (1994). Mesurement of word recognition, orthographic and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning desabilities: New views on measurement issues* (pp. 243-275). Baltimore/London/Toronto/Sydney: Paul H. Brookes

Parviainen, T., Helenius, P., Poskiparta, E., et al. (2006). Cortical sequence of word perception in beginning readers. *Journal of Neurosciences*, 26-22, 6052-6061.

Paulesu, E., Démonet, J.-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N, Cappa, S. F., Cossu, G., Habib, M., Frith, C. D., & Frith, U. (2001). Dyslexia, Cultural diversity and Biological unity. *Science, 291*, 2165-2167.

Peereman, R., & Content, A. (1998). *Quantitative analysis of orthography to phonology mapping in English and French (on-line).* http://homepages.vub.ac.be/~acontent/OPMapping.html

Peereman, R., Lété, B., & Sprenger-Charolles, L. (2007). Manulex-Infra: Distributional Characteristics of Infra-Lexical and Lexical Units in Child-Directed Written Material. *Behavior Research Methods*, 39, 593-603.

Perfetti, C. A. (1985). *Reading ability*. New-York: Oxford University Press.

Perfetti, C. A., 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.

Perfetti, C. A., Goldman, S. R., & Hogaboam, T. W. (1979). Reading skill and the identification of words in discourse context. *Memory and Cognition, 7*(4), 273-282.

Perfetti, C. A., & Hogaboam, T. (1975). Relationship between single word decoding and comprehension skills. *Journal of Educational Psychology*, 67(4), 461-469

Perfetti, C. A., & Zhang, S. (1995). The universal word identification reflex. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol.33, pp. 159-189). San Diego: Academic Press.

Perea, M., & Gotor, A. (1997). Associative and semantic priming effects occur at very short stimulus-onset asynchronies in lexical decision and naming. *Cognition, 62*, 223-240.

PIRLS-IEA (2004). PIRLS [Progess in International Reading literacy Studies]: 2006: PIRLS 2006 assessment framework and specification. Edited by I.V.S. Mullis, A.M. Kennedy, M.O.Martin, &.M.Sainsbury (100p.). http://www.skolverket.se/content/1/c4/03/71/ramverk_pirls_2006.pdf

PIRLS-IEA (2006). PIRLS [Progess in International Reading literacy Studies]: 2006: Technical report Edited by M.O.Martin, I.V.S. Mullis, & A.M. Kennedy (287p.). http://www.iea.nl/fileadmin/user_upload/graphics/Publications/p06_technical_report.pdf

PISA-OECD (2009). PISA [Program for International Student Assessment] 2006: Technical report. OECD, 419p. http://www.pisa.oecd.org/dataoecd/0/47/42025182.pdf

Plaut, D. C., & Booth J. R. (2000). Individual and developmental differences in semantic priming: Empirical and computational support for a single mechanism account of lexical processing. *Psychological Review, 107*, 786-823.

Proctor, C.P., August, D., Carlo, M.S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology*, 98-1, 159-169.

Proctor, C.P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology*, 97-2, 246-256

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13*(2), 212-21R

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., & Frith, U. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126,* 841-865

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.

Rayner, K., Chace, K.H., Slattery, T.J., & Ashby, J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, 10 (3), 241-255

Rayner, K., Foorman, B. R., Perfetti, C. A., Pesetsky, D. & Seidenberg, M. S. (2001). How psychological science informs the teaching of reading. *Psychological Science in the Public Interest*, 2, 31-74.

Riedel, B. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly.* 42(4), 546-567

Samuels, S.J. (2007). The DIBELS Tests: Is speed of barking at print what we mean by reading fluency? *Reading Research Quarterly*, 42(4), 563-566

Sebba, M. (2007). *Spelling and Society: The Culture and politics of orthography around the world.* Cambirdge University Press, 189p.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143-174.

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology, 72*, 95-129.

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

Share, D.L. (2008). On the Anglocentrism of current reading research and practice: The perils of overreliance on an "outlier orthography". *Psychological Bulletin*, 134, 584-615.

Shanahan , T. & Beck, I.L. (2006). Effective literacy teaching for English language learners. In D. August & T. Shanahan (Eds.) *Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 16, pp.415-488). Lawrence Erlbaum associates (LEA)

Shaywitz, B.A, Skudlarski, P., Holahan J.M; Marchione, K.E, Constable R.T., Fulbright, R.K.; Zelterman D., Lacadie, Ch., Shaywitz, S.E. (2007). Age-related changes in reading systems of dyslexic children, *Annals of neurology*, 61(4):363-70.

Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., Francis, D. J., Pugh, K. R., & Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study at Adolescence. *Pediatrics, 104*(6), 1351-1359.

Siegel, L. S. (1993). Phonological processing deficits in reading as the basis of a reading disability. *Developmental Review, 13*, 246-257.

Siegel, L. S., & Faux, D. (1989). Acquisition of certain grapheme-phoneme correspondences in normally achieving and disabled readers. *Reading and Writing: An Interdisciplinary Journal*, 1, 37-52.

Siegel, L. S., & Heaven, D. (1986). Categorization of learning disabilities. In S.J. Ceci. *Handbook of cognitive, social, and neuropsychological aspects of learning disabilities* (pp. 95-121), Hillsdale, NJ, Erlbaum.

Siegel, L. S., & Ryan, E. B. (1988). Development of grammatical sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Developmental Psychology*, 24, 28-37

Snow, C. (2006). Cross cutting themes and future research directions. In D. August & T. Shanahan (Eds.) *Developing literacy in second language learners. Report of the National Reading Panel on language minority and youth* (Chapter 21, pp.631-651). Lawrence Erlbaum associates (LEA).

Snow, C. E. & Kang, J. Y. (2006). Becoming bilingual, biliterate, and bicultural. In W. Damon, R.M. Lerner, A. Renninger, & I.E. Sigel (Eds.), *Handbook of child psychology, Volume 4, Child psychology in practice* (pp. 75-102). Hoboken, NJ: John Wiley & Sons.

Spiro, R.J., Bruce, B.C., & Brewer, W.F. (Eds.) (1980). Theorical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, articficial intelligence and education. Hillsdale, Erlbaum., 586p.

Sprenger-Charolles, L. (2003). Linguistic processes in reading and spelling. The case of alphabetic writing systems: English, French, German and Spanish. In T. Nunes, & P. Bryant (Eds) *Handbook of children's literacy.* (pp. 43-65). Dordrecht: Kluwer Academic Publishers

Sprenger-Charolles, L. (2008a). *EGRA (Early Grade Reading Assessment): Results from primary school students learning to read in French and in Wolof.* Report for the World Bank. http://pdf.usaid.gov/pdf_docs/PNADL691.pdf

Sprenger-Charolles, L. (2008b). *EGRA (Early Grade Reading Assessment): Results of 1200 Gambian children learning to read in English. Report for the World Bank.* http://pdf.usaid.gov/pdf_docs/PNADL690.pdf

Sprenger-Charolles, L., Colé P., Lacert, P., & Serniclaes, W. (2000). On subtypes of developmental dyslexia: Evidence from processing time and accuracy scores. *Canadian Journal of Experimental Psychology.* 54, 88-104.

Sprenger-Charolles, L., Colé, P., & Serniclaes, W. (2006). *Reading acquisition and developmental dyslexia.* New York, NY: Psychology Press.

Sprenger-Charolles, L., Siegel, L., Béchennec, D., & Serniclaes, W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading and in spelling: A four year longitudinal study. *Journal of Experimental Child Psychology*, 84, 194-217.

Stanovich, K.E. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32-71

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers.* New York: Guilford Press.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86*, 24-53.

Stanovich, K. E. & West, R.F. (1979). Mechanisms of sentence context effects in reading: Automatic activation and conscious attention. *Memory & Cognition*, 1979, 7, 77-85

Stanovich, K.E., West, R.F., & Freeman, D.J. (1981). A longitudinal study of sentence context effects in second-grade children: Tests of an interactive compensatory model. *Journal of Experimental Child Psychology*, 32, 185-199.

Sticht T. G. & James J.H. (1984) Listening and Reading. In P.D Pearson (Ed.). *Handbook of reading research* (pp. 293-317). New York: Longman Press

Stuart, M. & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition*, 30, 139-181

Tapiero, I. (2007). *Situational models and levels of coherence: toward a definition of comprehension.* Mahwah (NJ), Lawrence Erlbaum associates. 219p

Till, R.E., Mross, E.F., & Kintsch, W. (1988). Time course of priming for associate and inference words in a discourse context. *Memory and Cognition*, 16-4, 283-298.

Kintsch, W. & Mross, E.F. (1985). Context effects in word identification. *Journal of Memory and Language.* 24-3, 336-349

Torgeson, J.K., Wagner, R.K., & Rashotte, C. (1999). *Test of Word Reading Fluency (TOWRF).* Austin, Texas: PRO-ED.

Torgesen, J.K., Rashotte, C.A., Alexander, A. (2001). Principles of fluency instruction in reading: Relationships with established empirical outcomes. In M. Wolf (Ed.), *Dyslexia, Fluency, and the Brain.* pp. 333-355. Parkton, MD: York Press.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography, *Journal of Experimental Psychology: General, 124*, 107-136.

Van Orden, G.C & Kloos, H. (2005). The question of phonology and reading. in M. Snowling & C. Hulme (Eds.). *The Science of Reading: A Hanbook* (pp. 61-78). Oxford: Blackwell Publishing

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What we have learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45(1), 2-40.*

Venezky, R. L. (1966). *Automatic spelling-to-sound conversion.* Computation in Linguistics: A Case Book. Indiana University Press, Bloomington.

Verhoeven, L. (1990). Acquisition of reading in a second language. *Reading Research Quarterly*, 25, 90-114

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading, 4,* 313-330

Wagner, D. A. (2003). Smaller, quicker, cheaper: Alternative strategies for literacy assessment in the UN Literacy Decade. *International Journal of Educational Research*, 39.

Wagner, D.A. (1993). *Literacy, Culture, & Development: Becoming literate in Morocco.* USA: Cambridge University Press.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology*, 33(3), 468-479.

Waters, G. S., Bruck, M., & Seidenberg, M. S. (1985). Do children use similar cognitive processes to read and spell words. *Journal of Experimental Child Psychology*, 39, 511-530.

West, R.F., & Stanovich, K.E. (1978). Automatic contextual facilitation in readers of three ages. *Child Development*, 49, 717-727.

Zeno, S. M., Ivenz, S. H., Millard, R. T., & Duvvuri, R. (1995). *The educator's word frequency guide.* Brewster, NY: Touchstone Applied Science Associates.

Ziegler, J. C. (2006). Do differences in brain activation challenge universal theories of dyslexia? *Brain & Language*, 98, 341-343.

Ziegler, J.C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 13(1), 3-29.

Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science*, 9, 429-436.

Ziegler, J. C., Perry, C., Jacobs, A. M., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science*, 12, 379-384

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for -ough and the spelling for /u/? A data base for computing feedforward and feedback consistency in English. *Behavior Research Methods Instruments and Computers, 29*(4), 600-618.

Zwaan, R.A. & Brown, C.M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, 21, 289-327.

Zwaan, R.A. & Radvansky, G.A. (1998). Situational models in language comprehension and memory. *Psychological Bulletin, 12, 162-185.*

APPENDIX A: SAMPLE TEXT FROM PIRLS (IEA)

Ronde de nuit pour les macareux[86]

Chaque année, des oiseaux noirs et blancs au bec orange visitent l'île islandaise de Heimaey. Ces oiseaux sont des macareux. On les appelle les « clowns de la mer » à cause de leur bec de couleur vive et de leurs mouvements maladroits. En effet les macareux ne sont pas très élégants au décollage ni à l'atterrissage étant donné leur corps trapu et leurs ailes courtes.

Halla vit dans l'île de Heimaey. Elle scrute le ciel tous les jours. Perchée sur une haute falaise bien au-dessus de la mer, elle aperçoit son premier macareux de la saison. Elle murmure le mot « lundi », qui signifie « macareux » en islandais.

Bientôt le ciel sera rempli de ces oiseaux: des macareux, des macareux partout. Ils reviennent de leur hivernage en mer, rentrant sur l'île d'Halla et sur les îles désertes des alentours pour pondre et pour élever leurs poussins. Ces « clowns de la mer » reviennent dans les mêmes terriers, année après année. C'est d'ailleurs la seule saison qu'ils passent sur le rivage.

Halla et ses amis montent sur les falaises pour observer les oiseaux. Ils voient les couples qui cognent leur bec en un bruyant « tac-tact-tac ». Bien à l'abri dans les falaises, chaque couple prendra bientôt soin d'un œuf. Quand les œufs des macareux seront éclos, les parents rapporteront du poisson au nid pour nourrir les oisillons. Chaque oisillon deviendra un jeune macareux. La ronde de nuit aura lieu quand tous les jeunes s'élanceront pour leur premier vol. Ces nuits là sont encore bien loin, mais Halla pense déjà à préparer quelques boîtes de carton pour eux.

Pendant tout l'été, les macareux adultes vont à la pêche et prennent soin de leurs petits. En août, les premières fleurs ornent les terriers. Quand les fleurs sont toutes écloses. Halla sait que la ronde de nuit va bientôt commencer.

Les oisillons bien cachés sont devenus de jeunes macareux. Il est temps pour Halla et ses amis de sortir les boîtes et les lampes de poche pour la ronde de nuit de jeunes macareux. A partir de ce soir et pendant les deux semaines à venir, les jeunes macareux s'envolent pour leur hivernage en mer.

Dans la noirceur de la nuit, les jeunes macareux quittent leurs abris pour leur premier vol. Il leur suffit de quelques battements d'ailes pour descendre des hautes falaises. La plupart des oiseaux amerrissent dans un grand éclaboussement et en tout sécurité parce que la mer est en dessous. D'autres sont toutefois désorientés par les lumières du village, pensant peut être que ce sont les reflets de la lune sur la mer. Chaque nuit, des centaines de jeunes macareux atterrissent maladroitement dans le village. Incapables de s'envoler à partir d'un terrain plat, ils courent partout et tentent de se cacher.

Halla et ses amis vont chaque nuit à la recherche des jeunes macareux échoués qui n'ont pas réussi à trouver la mer. Malheureusement, les chats et les chiens cherchent, eux aussi. Même si les chats et les chiens ne les attrapent pas, les jeunes macareux risquent de se faire écraser par les voitures et les camions. Les enfants doivent donc les trouver en premier. Dès vingt-deux heures, les rues de Heimaey sont pleines de vie et d'enfants aux aguets.

Halla et ses amis s'empressent de secourir les jeunes macareux égarés. Armés de lampe de poche, ils sillonnent le village, fouillant les coins sombres. Halla repère un jeune macareux. Elle court, le saisit et le met en sécurité dans une boîte de carton.

Pendant deux semaines, tous les enfants de Heimaey feront la grasse matinée pour rester dehors la nuit. Ils sauveront ainsi des milliers de jeunes oiseaux.

Chaque nuit, Halla et ses amis emportent à la maison les jeunes macareux retrouvés. Le jour suivant, le petit groupe descend sur la plage avec des boîtes pleines de jeunes macareux.

Il est temps de libérer les oiseaux. Halla relâche le premier. Elle le tient bien haut pour qu'il s'habitue à battre des ailes. Puis, en le tenant confortablement dans ses mains elle le soulève d'un mouvement rapide dans les airs et le relâche au-dessus de l'eau, au-delà du ressac. Le jeune

---

86. English version: PIRLS, 2004. Appendix B (pp.81-92) http://www.skolverket.se/content/1/c4/03/71/ramverk_pirls_2006.pdf.

macareux volette un peu avant de se poser sur l'eau dans un éclaboussement.
Jour après jour les protégés d'Halla barbotent et s'éloignent jusqu'à ce que la ronde de nuit des macareux soit terminée pour une autre année. Elle regarde les derniers s'éloigner pour l'hivernage en mer et les salue de la main jusqu'au printemps prochain. Elle leur souhaite un bon voyage et crie « Au revoir! Au revoir! ».

QUESTIONNAIRE
1. Pourquoi les macareux sont-ils maladroits au décollage et à l'atterrissage?
        parce qu'ils vivent sur un sol glacé?
        parce qu'ils viennent rarement sur le rivage?
        parce qu'ils passent du temps sur les hautes falaises?
        parce qu'ils ont un corps trapu et de courtes ailes?
2. Où les macareux passent l'hiver?
        à l'intérieur des falaises?
        sur la plage?
        en mer?
        sur la glace?
3. Pourquoi les macareux viennent-ils sur l'île?
        pour être secourus?
        pour chercher de la nourriture?
        pour pondre des oeufs?
        pour apprendre à voler?
4. Comment Halla sait-elle que les jeunes macareux vont bientôt voler?
        les parents leur apporte du poisson?
        les fleurs sont toutes écloses?
        les poussins sont cachés?
        l'été vient de commencer?
5. Que se passe t'il pendant la nuit de la ronde des macareux?
        les couples de macareux se cognent les becs en bruyant tac-tac-tac?
        les jeunes macareux font leur premier vol?
        les œufs des macareux éclosent et les poussins en sortent?
        les jeunes macareux viennent de la mer se poser sur le rivage?
6. Qu'est ce que les gens du village pourraient faire pour que les jeunes macareux n'atterrissent plus par erreur dans le village?
        Eteindre les lumières?
        préparer des boîtes?
        garder les chiens et les chats dans les maisons?
        illuminer le ciel à l'aide de lampes de poche?
7. Explique comment Halla utilise sa lampe de poche pour secourir les macareux?
        ...
8. Explique comment Halla utilise les boites de carton pour secourir les macareux?
        ...
9. D'après le texte, lequel des dangers suivant menace les jeunes macareux?
        ils risquent de se noyer en se posant sur la mer?
        ils risquent de se perdre dans les buissons?
        ils risquent de ne pas recevoir assez de poisson de leurs parents?
        ils risquent d'être écrasés par les voitures et les camions?
10. A l'aide de l'information tirée du texte, explique pourquoi il faut qu'il fasse jour pour que les enfants libèrent les jeunes macareux? …
11. Que font les jeunes macareux après que Halla et que ses amis les ont libérés?
        ils marchent sur la plage?
        ils s'envolent du haut de la falaise?
        ils se cachent dans le village?
        ils nagent dans la mer?
12. Indique deux sentiments que Halla pourrait éprouver après avoir libéré les jeunes macareux et explique pourquoi elle pourrait éprouver chacun de ces sentiments? …
13. Aimerais-tu aller avec Halla et ses amis secourir de jeunes macareux. A l'aide du texte, explique ta réponse?

APPENDIX B: The test Blueprint for the SACMEQ II Reading test

Skill Level	Narrative	Expository	Document	
Level 1	Word/picture association involving positional or directional prepositions requiring the linkage of a picture to a position or a direction in order to answer the question	Word/picture association involving positional or directional prepositions requiring the linkage of a picture to a position or a direction in order to answer the question	Word/picture association involving positional or directional prepositions requiring the linkage of a picture to a position or a direction in order to answer the question	
Items	2	2	2	6
Level 2	Recognising the meaning of a single word and being able to express it as a synonym in order to answer the question	Recognising the meaning of a single word and being able to express it as a synonym in order to answer the question	Linking simple piece of information to item or instruction	
Items	7	6	9	22
Level 3	Linking information portrayed in sequences of ideas and content, when reading forward	Linking information portrayed in sequences of ideas and content, when reading forward	Systematic search for information when reading forward	
Items	8	10	8	26
Level 4	Seeking and confirming information when reading backwards through text	Seeking and confirming information when reading backwards through text	Linking more than one piece of information in different parts of a document	
Items	9	5	4	18
Level 5	Linking ideas from different parts of text. Making inferences from text or beyond text, to infer author's values and beliefs	Linking ideas from different parts of text. Making inferences from text or beyond text.	Use of embedded lists and even subtle advertisements where the message is not explicitly stated	
Items	6	3	2	11
Total Items	32	26	25	83

APPENDIX C: PASEC MAIN GOALS AND TESTS

C1. Tableau synthétique Début 2ème année

Exercices	Domaines	Objectifs
5	**Compréhension de mots (vocabulaire)**	Identifier parmi 3 mots celui qui correspond à l'image
2	**Compréhension de phrase**	Ecrire une phrase à partir de 4-5 mots donnés dans le désordre
8-9		Identifier la phrase (parmi 3) qui correspond à l'image (2 sous-tests)
1-6	**Lecture / déchiffrement**	Identifier une syllabe dans une série de mots ('pi' dans 'épine, pipe, pilon') Reconnaître un mot identique au mot test parmi 4 mots proches visuellement ou se prononçant de la même façon ('sot': 'saut, seau, pot, sot').
7		Copie: Ecrire le mot qui manque dans une phrase incomplète, la phrase complète étant présentée au dessus de celle qui est à compléter
3-4	**Ecriture**	Ecrire une syllabe (3) ou un mot (4) à partir d'une lettre (2 sous-tests)

C2. Tableau synthétique Fin 2ème année

Exercices	Domaines	Objectifs
1	**Compréhension de mots (vocabulaire)**	Identifier parmi 4 images celle qui correspond au mot écrit présenté
4	**Compréhension de phrases**	Identifier le mot qui donne du sens à la phrase ('Il prend le train à la …' [gare-oiseau-école])
6		Ecrire une phrase à partir de 4-5 mots donnés dans le désordre
9		A l'aide d'une image, identifier la préposition donnant du sens à la phrase ('Sidi est [à-de-dans] la voiture').
10	**Compréhension de texte**	Compléter un texte comportant des mots qui manquent (donnés, mais dans le désordre).
2-3-8	**Lecture – Ecriture: discrimination de sons proches (t-d; f-v;br-pr..)**	Ecrire après écoute la lettre (ou le groupe de lettre) qui manque (par exemple: 't ou d' dans 'maXame' et 'paXate'; 'f ou v' dans 'Xarine' et 'Xie'; 'pr ou br' dans 'XXépare' et 'XXanche, 3 sous-tests)
5	**Grammaire (Conjugaison)**	Identifier le pronom personnel qui va avec le verbe conjugué ('… parles trop' [tu-nous-vous])
7	**Grammaire**	Distinguer le singulier et le pluriel des noms ('Il porte des [cahiers, livre, mètre])

C3. Tableau synthétique Début 5ème année

Exercices	Domaines	Objectifs
1	**Compréhension de mots et de phrases**	Identifier le sens d'un mot dans une phrase: 'la grande soeur a discuté avec son frère' signifie: 'elle a travaillé avec lui', 'elle a joué avec lui', 'elle a parlé avec lui', 'elle a mangé avec lui'
2		Identifier la préposition correcte ('le chavel trotte [contre-sous-dans] la rue')
15-16	**Compréhension de textes**	Répondre à des questions dont la réponse se trouve explicitement dans le texte (lecture d'une notice de médicament) Lire un text à trou et le compéter avec des mots donnés, dont 1 en trop.
3	**Grammaire 1**	Accorder le participe passé: 'Ma mère prépare mon plat… [préféré-préférée-préférés-préférer]'
4		Accorder le verbe avec le sujet: 'Mon père et moi [allons-va-vont] à la foire'
5-6-7	**Grammaire 2 (Conjugaison)**	Identifier le temps d'un verbe (indicatif présent, imparfait, passé composé, et futur simple) Identifier une phrase écrite sans erreur orthographique dans le verbe
8	**Grammaire 3 (Forme de la phrase)**	Transformer une phrase affirmative en une phrase interrogative
9-10-11-13	**Grammaire 4**	Entourer le complement d'objet indirect ou le sujet d'une phrase (2 sous-tests) Entourer le pronom qui peut remplacer le groupe souligné (par exemple, 'la fête aura lieu dimanche' [elles-bous-elle]) Compléter la phrase: 'C'est l'école de Mady et de Kassi: c'est........ école'
12-14	**Orthographe**	Identifier le nom qui se termine par 'x' au pluriel (bleu, chapeau, jupe): - Orthographier correctement des homophones ('il [s'est-ces-ses-c'est] blessé'

APPENDIX D: *ILLUSTRATIVE TABLE OF EGRA FIELD STUDIES PRESENTED BY COUNTRY*[87]

Language(s) of assessment, grade(s) tested, number of children assessed; subtests are numbered accordingly to the list presented above in section 2.5. ("intervention" indicates that EGRA has been implemented to monitor progress in a reading intervention program).

Country	language of assessment	Grade tested (number of students assessed)	Subtests employed	Intervention
Liberia		Grade 2 (429) Grade 3 (407)	1), 2), 3),4), 5), 6), 7), 8)	Intervention
Kenya	English, Kiswahili	Grade 2[88]	2), 3) (in English only), 5), 6), 7)	Intervention
The Gambia	English	Grade 1, 2 and 3 (1200)	1), 2), 3),4), 5), 6), 7), 8), 9)	
Senegal	French Wolof	French: Grade 1 to 3 (502) Wolof: Grade 1 and 3 (186)	1), 2), 3),4), 5), 6), 7), 8), 9)	
Egypt	Arabic	(100)	1), 2), 3),4), 5), 6), 7), 8), 9)	Intervention
Guatemala	Spanish and mother tongue[89] Spanish, Mam, K'iche, Ixil	Grade 2 and 3 Grade 3	2), 3),4), 5), 6), 7), 8), 9) 1), 6), 7)	Intervention
Haiti	Haitian Creole, French	Grade 2 to Grade 4 (3000)	2), 3),4), 5), 6), 7), 8), 9) + an extra vocabulary task	
Honduras	Spanish	Grade 2 to 4 (2226)	1), 6), 7)	
Mali	French, Arabic Bamanankan, Bomu, Songhoi, Fulfulde	French grades 2, 4 and 6; Arabic grades 2 and 4, Grade 1 to 3 in remaining 4 languages[90]	1), 2), 3),4), 5), 6), 7), 8), 9)	
Ethiopia	Ofo Aromo	Grade3	1), 6), 7)	
Guyana	English	Grade 1 to 3 (2699)	2), 3),4), 5), 6), 7), 8), 9)	
Uganda	English, Luganda, Lango		2), 3),4), 5), 6), 7),	

---

87. Other countries where EGRA has been used to date include: Niger, South Africa, Afghanistan, India, Bangladesh, Nepal, Jamaica, Nicaragua, Peru, Vietnam, Timor Leste
88. Data missing on sample size.
89. Information missing on exact language.
90. Data missing on sample size. An extra reading comprehension task has been adde: a cloze test)

APPENDIX E. SOME RESULTS FROM EGRA AND SIMILAR ASSESSMENTS

			Vocabulary*	Spoken comprehension	Written comprehension (reading aloud)	Reading (Aloud) (words in context, isolated words and pseudowords)			Spelling (words and pseudowords)		Phonemic awareness		Syllabic awareness	Phonological Short-term memory**
			No time limit % correct	No time limit Story % correct	No time limit Story % correct	1min Story Word-Min	1min Isolated Word-Min	1min Pseudo word-Min	Word % correct	Pseudoword % correct	Identification % correct	Counting % correct	Counting % correct	% correct
EGRA G1	Senegal (French) n=150	Mean: SD		15 / 23	No calcul: Reading level too low	4 / 7	4 / 4	4 / 6	6 / 16		31 / 33	31 / 35		
	[a]Floor effect % children			63	79	50	27	39	80		38	37		
	[b]Ceiling effect % children			no	no	no	no	no	no		no	no		
EGRA G1	Senegal (Wolof) n=83	Mean: SD		47 / 34	No calcul: Reading level too low	1 / 4	2 / 3	2 / 3	2 / 10		31 / 33	21 / 27		
	[a]Floor effect % children			26	81	71	63	63	95		43	44		
	[b]Ceiling effect % children			no	no	no	no	no	no		no	no		
EGRA G1	Gambia (English) n=419	Mean: SD		25 / 35	No calcul: Reading level too low	2 / 9	1 / 5	1 / 3	3 / 13		23 / 27	21 / 26		
	[a]Floor effect % children			58	75	80	81	91	95		47	51		
	[b]Ceiling effect % children			no	no	no	no	no	no		no	no		
EGRA Age 6 to 17: 547 (6-9: 287)	Niger (Zarma) n=287	Mean 6-9 SD	69 / 15									17 / 20	34 / 32	63 / 27
	[a]Floor effect % children		no									24	24	no
	[b]Ceiling effect % children		no									no	no	no
France G1	Children speaking French + another language n=31	Mean: SD	81 / 13	72 / 28	77 / 30	63 / 42	39 / 16	30 / 13	71 / 23	88 / 21				
	[a]Floor effect % children		no	no	no	no	no	no	no	no				
	[b]Ceiling effect % children		no	26	40	no	no	no	18	41				
France G1	Native and non-native n=125	Mean: SD	89 / 11	79 / 28	82 / 25	65 / 33	44 / 17	32 / 12	76 / 21	91 / 15				
	[a]Floor effect % children		no	no	no	no	no	no	no	no				
	[b]Ceiling effect % children		14	38	50	no	no	no	18	41				

Senegalese and Gambian results: Sprenger-Charolles (2008a-b); French results: Data collected with the help of the staff of the "Inspection Departementale Annecy-Sud": C.Greffe; A.Bats; M.Perret (Unpublished data).
The Nigerian data have been provided by D.Bechennec and the staff "PLAN Niger" (especially D.Alfari and O.Housseini): these assessments have been done before learning to read.
a. Floor effect %: Percentage of children with scores at the floor level (= 0 correct answer).
b. Ceiling effect %: Percentage of children with scores at the ceiling level (100% correct answers).
* Vocabulary: part of the body, items from the environment, actions (to yawn), spatial terms (above, under…). The child has to do something (for instance, to do an action such as put a pen under a book, to show a part of his/her body…).
** Phonological short-term memory: The child should repeat 10 pseudowords (2 to 5 syllables) such as tudoni; gutodira; koguteru; tabaritolu; toduratide.