Available online at www.sciencedirect.com
SCIENCE DIRECT®
Revue européenne de psychologie appliquée 55 (2005) 157–186
Revue européenne de psychologie appliquée
http://france.elsevier.com/direct/ERAP/

Original article

French normative data on reading and related skills from EVALEC, a new computerized battery of tests (end Grade 1, Grade 2, Grade 3, and Grade 4)[1]

Lecture et compétences reliées : données normatives pour la fin de la 1re, 2nde, 3e et 4e année du primaire issues d'une nouvelle batterie de tests, EVALEC

L. Sprenger-Charolles [a,*], P. Colé [b], D. Béchennec [a], A. Kipffer-Piquard [c]

[a] Laboratoire de psychologie expérimentale (UMR 8581), CNRS, université René-Descartes (Paris 5), institut de psychologie, 71, avenue Edouard-Vaillant, 92100 Boulogne-Billancourt, France
[b] LPNC-CNRS-UMR-5105, université de Savoie, Grenoble-Chambery, France
[c] Institut des Jeunes Sourds de La Malgrange, Jarville-Nancy, France

Received 26 April 2004; received in revised form 12 September 2004; accepted 15 November 2004

Abstract

To set-up standardized norms on the development of reading and related skills in French, we have developed a new tool, EVALEC. The data were collected at the end of Grades 1–4 (about 100 children for each level). EVALEC includes four tests focused on written word processing (three reading aloud tests, one silent reading test); both accuracy scores and processing time (time latency and vocal response duration for the reading aloud tests) were recorded. EVALEC also includes tests of phonemic and syllabic awareness, phonological short-term memory, and rapid naming; the latter test is presented in non-reading and in reading contexts in order to compare the impact of the time constraint in both domains. We assessed the effects of regularity (regular vs. irregular words), graphemic length and graphemic complexity (regular words with only single letter graphemes vs. those with digraphs such as "ch" and those with contextual graphemes such as "g") on the reading of high frequency words. We also compared the effect of graphemic length and graphemic complexity on regular word versus pseudoword reading as well as the effect of the length of the items (short vs. long items) on irregular word versus pseudoword reading. According to the ANOVAs, for word reading, the effects of sublexical factors appeared stronger for the youngest children, but regular words were always read better and faster than irregular words. However, these lexical factors did not have the same effect for words and pseudowords. Their impact, when positive, was greater for words (e.g. the graphemic length), while the negative effects were especially marked for pseudowords (e.g. the length of the items). Among the correlations, those between accuracy and latency time were never significant for the youngest children, thus suggesting that some beginning readers favored accuracy to the detriment of speed while others adopted the opposite strategy. Concerning the correlations between the reading and the reading-related tests, only the correlations between phonemic awareness and reading were significant and only with the youngest children. In addition, the correlations between the two rapid naming tests (reading and non-reading tests) were not significant most of the time. These data would allow researchers and speech therapists to assess the reading and reading-related skills of dyslexics as compared to average readers.

[1] The computerized version of this battery of test was set-up by Elisabeth Samain (CNRS-UMR-820) and René Carré (TSI, Traitement du Signal et des Images, ENS des Télécommunications, Paris). We gratefully thank them for their help.

* Corresponding author.
E-mail address: sprenger@linguist.jussieu.fr (L. Sprenger-Charolles).

1162-9088/$ - see front matter © 2005 Elsevier SAS. All rights reserved.
doi:10.1016/j.erap.2004.11.002

Résumé

Il n'existe en français aucune batterie permettant de situer les performances en identification des mots écrits de dyslexiques par rapport à celles de normolecteurs en tenant compte de la précision et du temps de latence des réponses. EVALEC vise à combler cette lacune. Les données ont été recueillies en fin de 1[ère], 2[nd], 3[e] et 4[e] de primaire, auprès d'une centaine d'enfants par niveau. EVALEC comprend quatre épreuves centrées sur les procédures d'identification des mots écrits, une en lecture silencieuse et trois en lecture à haute voix, ces dernières prenant en compte deux indicateurs pour la rapidité : le temps de latence et la durée de la réponse vocale. Cette batterie inclut des épreuves d'analyse phonémique et syllabique, de mémoire à court terme phonologique et de dénomination rapide hors et en lecture, ceci afin d'évaluer l'impact d'une contrainte de rapidité dans ces deux domaines. Dans les épreuves de lecture à haute voix, on a évalué les effets de facteurs sublexicaux sur des mots fréquents: régularité (mots réguliers vs. irréguliers), longueur et complexité graphémique (mots réguliers incluant des graphèmes d'une lettre vs ceux incluant un digraphe, tel que « ch », ou un graphème contextuel, tel que « g »). On a aussi comparé l'effet de la longueur et de la complexité graphémique sur des mots réguliers et des pseudomots, ainsi que celle de la longueur des items sur des mots irréguliers et des pseudomots. D'après les ANOVAs, en lecture de mots, les effets des facteurs sublexicaux sont plus fortement marqués chez les plus jeunes enfants mais, les mots réguliers sont toujours mieux lus et plus rapidement que les irréguliers. De plus, l'incidence de ces facteurs n'est pas la même sur les mots et les pseudomots, leur impact étant plus fort sur les mots lorsqu'ils sont facilitateurs (longueur graphémique) alors que les effets négatifs (longueur des items) sont surtout marqués sur les pseudomots. Parmi les corrélations, celles entre précision et temps de latence ne sont jamais significatives chez les plus jeunes ce qui suggère que certains privilégient la précision au détriment de la rapidité, d'autres adoptant la stratégie inverse, et celles entre les tests de lecture et les autres évaluations ne sont significatives que pour les capacités d'analyse phonémique et ce uniquement chez les plus enfants les plus jeunes. Ces données devraient permettre, dans le cadre d'un bilan de dyslexie, de situer le niveau de lecture des enfants, d'évaluer l'efficience de leurs procédures d'identification des mots et leurs habilités dans les domaines reliés à la lecture.

Keywords: Reading acquisition; Grapheme-phoneme correspondences; Lexicality effect; Length effect; Regularity effect; Graphemic length effect; Digraph; Contextual grapheme; Accuracy scores; Time latency; Vocal response duration; Phonemic awareness; Syllabic awareness; Phonological short-term memory; Rapid naming

Mots clés : Acquisition de la lecture ; Procédure lexicale ; Procédure sublexicale ; Correspondances graphème-phonème ; Lexicalité ; Effet de la longueur ; Régularité ; Traitement des digraphes ; Traitement des graphèmes contextuels ; Temps de latence ; Durée de la réponse vocale ; Analyse phonémique ; Mémoire à court terme phonologique ; Dénomination rapide

1. Introduction

It is currently accepted that dyslexics' reading difficulties are not linked to a comprehension problem, but rather to a deficit in written word processing (Stanovich, 2000; Snowling, 2000). To assess a deficit, it is necessary to rely on tests designed according to an appropriate theoretical framework and based on normative data. Thus, we have created EVALEC, a new battery of tests, and collected normative data at the end of Grades 1–4. EVALEC would allow researchers and speech therapists to assess dyslexics' reading skills (in silent reading and in reading aloud) and their skills in domains linked to reading acquisition (syllabic and phonemic awareness, phonological short-term memory, and rapid naming), as compared to those of average readers of the same chronological age and/or of the same reading level. EVALEC is different from earlier batteries of tests (e.g. the BELEC, Mousty and Leybaert, 1999; Mousty et al., 1994) by the fact that the reading tasks were administered on a computer, so as to take into account both accuracy and response time. This is a crucial point, the efficiency of a skill depending not only on its accuracy but also on its speed, and not taking response time into account often leads to diagnostic errors (see Sprenger-Charolles et al., 2000; Sprenger-Charolles et al., 2005 under revision). EVALEC is also different from other batteries of tests because the data were collected from G1 to G4, each with approximately 100 children, while for the BELEC for example, available norms begin in second grade and the number of children is very different from a grade to the other (200 children in G2 and G4, Mousty and Leybaert, 1999; 12–18 children in G3 and G5, Mousty et al., 1994).

2. Theoretical framework

2.1. Written word processing

Given that the dual-route model (Coltheart et al., 2001) was, and is still, the model of reference for studies in dyslexia, whether for acquired dyslexia or developmental dyslexia, the present study's hypotheses have been developed with regard to this model. According to this model, written words can be read either by a lexical procedure (also called direct route or orthographic procedure) or by a sublexical phonological procedure (also called indirect route or phonological mediation). A frequency effect is reckoned to be the signature of the lexical procedure; high frequency words being assumed to be read better and faster than low frequency words as their "address" could be more easily accessible because more often asked for. A lexicality effect (the superiority of high frequency words compared to pseudowords) is another indicator of use of this procedure since pseudowords cannot have an address in the subject's internal lexicon, unless they are near or analogous to words (for a review, see Ferrand, 2001). If the item to be read does not figure among the words stored in his/her orthographic lexicon (a new word or a proper noun), the reader is unable to use this procedure. He/she therefore has recourse to a sublexical procedure. In alphabetic writing systems, the sublexical units of the written language (graphemes) have to be first translated into the sublexical units of the spoken language (phonemes), these being then "assembled". This procedure, which allows the reader to recognize the words which are part of his/her oral lexicon, as well as to identify new words, can lead to regularization errors on items which have irregular grapho-phonemic correspondences (e.g. the word "sept"/sɛt/will be read as the word "septembre"/sɛptãbr/). Reliance on the sublexical reading procedure is therefore attested by the production of this type of error and by the presence of a regularity effect (the superiority of regular words compared to irregular words) in the absence of any effect of frequency or of lexicality.

2.1.1. Development of written word processing

Most studies on reading acquisition have relied on the dual-route model. As we have emphasized, in this framework the presence of frequency or lexicality effects, in the absence of an effect of regularity, gives evidence that the subjects rely on the lexical procedure; the inverse tendency (a regularity effect in the absence of frequency or lexicality effects) would indicate reliance on the sublexical procedure. For silent reading tasks, the variables most often manipulated are the visual or phonological similarities. Thus, with lexical decision as with semantic categorization tasks, a high rate of erroneous acceptances of pseudohomophones (e.g. "oto" for "auto", /oto/) is interpreted as an indication of reliance on the sublexical procedure.

Apart from a few words that could have been learned by heart[2], beginning readers depend mainly on the sublexical procedure. This was found to be true for English (Backman et al., 1984; Waters et al., 1984), for German (Wimmer and Hummer, 1990), and for French (Leybaert and Content, 1995; Sprenger-Charolles and Bonnet, 1996; Sprenger-Charolles et al., 2003; Sprenger-Charolles et al., 1998b). Other studies indicate that, at least for French, children rapidly take into account graphemes rather than letters as units for the sublexical procedure (Sprenger-Charolles, 1994; Sprenger-Charolles and Casalis, 1995; Sprenger-Charolles et al., 1998b)[3].

In the studies cited above, written word processing is assessed using reading aloud tasks. This compels the reader to produce an oral response, and so to use the phonological representation of the written word, which could possibly affect the results. This is not the case for silent reading. However, in silent reading tasks, the phonological characteristics of the items were also found to have an effect with beginning readers (Booth et al. 1999; Bosman and de Groot, 1996; Sprenger-Charolles et al., 1998a).

Other studies have indicated that the transparency of grapho-phonemic relationships has an effect on reading acquisition (for a review, see Sprenger-Charolles, 2003; Ziegler and Goswami, 2005). Compared to English-speaking children, superior reading scores were found for Spanish-speaking children (Goswami et al., 1998), German-speaking children (Frith et al., 1998; Goswami et al., 2001), and French-speaking children (Bruck et al., 1997; Goswami et al., 1998). Some of these studies suggest that the degree of orthographic transparency has an effect on the nature of the reading processes used by the learners, the German-speaking children relying more on phonologically based processes than English-speaking children (Frith et al., 1998; Goswami et al., 2001). In the study conducted by Goswami et al. (2001), reading aloud and silent reading tasks were both used. Differences were found in the processing of pseudohomophones depending on the task (reading aloud or silent reading) and on the degree of opaqueness of the grapheme-phoneme correspondences (English versus German). In the reading aloud task, the German-speaking children read the pseudohomophones just as well as they read the non-homophone pseudowords, both with a very high level of accuracy. The level of accuracy of the English-speaking children was lower, particularly for the non-homophone pseudowords. In silent reading— a lexical decision task — a pseudohomophone disadvantage was observed only for the German-speaking children. These results suggest that the activation of phonological information is rather automatic and difficult to inhibit for the German-speaking children, but not for the English-speaking children.

All these various studies indicate that phonology plays an important role in learning to read. The sublexical procedure could even play a central role in reading acquisition and particularly in the setting-up of the orthographic lexicon. This procedure allows the reader to process all regular words, whether or not they are known, and works at least partially for reading irregular words which always contain regular grapheme-phoneme correspondences. The reader can therefore read irregular words in part by using the grapheme-phoneme correspondences and then, correct any errors. For example, for the French word "femme" /fam/, reliance on the most frequent grapho-phonemic relationships provides /fɛm/, which is not a French word. But, since there is a frequent word with a very close pronunciation, /fam/, it is possible to infer that the "e" of "femme" should be read as /a/. Strong associations between orthographic and phonological units can thus be made, and memorized, according to the frequency of the grapho-phonemic correspondences and the frequency of the word, both at the sublexical level (between graphemes and phonemes) and the lexical level (between the orthographic and the phonological representations of the word).

The role of the sublexical procedure in the setting-up of the orthographic lexicon was assessed in studies entailing different kinds of training. The results suggest that phonologi-

[2] These pre-reading strategies, called logographics, seem not to play a significant role in the very early stage of reading acquisition, at least not in relatively transparent writing systems (Wimmer and Hummer, 1990, for German; Sprenger-Charolles and Bonnet, 1996, for French).

[3] However, syllabic units may also be used, at least in French (Colé et al., 1999; Colé and Sprenger-Charolles, 1999).

cal training is more efficient than visual training. For instance, young non-readers learnt to read words when they were associated with phonological clues more easily than when the same items were associated with visual clues (Ehri and Wilce, 1985; Laing and Hulme, 1999; Rack et al., 1994). Similar results were obtained with older children (Share, 1999). Further evidence is provided by longitudinal studies showing that the children who, from the beginning of learning to read had accurate phonological reading scores (attested by their scores in pseudoword reading), later obtained the best results in reading, including for irregular words (Byrne et al., 1992; Jorm et al., 1986). Other longitudinal data have highlighted the fact that early pseudoword reading captures the major part of the unique variance of later word reading skills (Manis et al., 1993; Sprenger-Charolles et al., 2003). In one of these studies (Sprenger-Charolles et al., 2003), non-reading kindergartners were tracked during 4 years. Their performances for pseudoword, regular and irregular word reading were examined after 4 months of learning to read, and then at the end of each school year (G1–G4). In the middle of G1, performances for regular words and pseudowords were not different, but were superior to those for irregular words, these being quite low, indicating that reading mainly depended on the sublexical procedure. A few months later, the picture changed considerably. Progress was observed for all three types of items, but was most significant for the regular words which were read better than the pseudowords, which, in turn, were read better than the irregular words. The regular word results may be explained by the double benefit of regularity and frequency of exposure. Neither frequency of exposure alone nor regularity alone is sufficient, as suggested by the lesser progress for irregular words and pseudowords. These data also explain why reading acquisition is accomplished better and faster when the grapho-phonological relationships are relatively transparent, as in Spanish for example.

The sublexical procedure seems thus to be a bootstrapping mechanism in reading acquisition (Share, 1995) and its efficiency is a function of the consistency of the grapho-phonemic correspondences. Nevertheless, if reading acquisition depended solely on the transparency of the orthography, we should not find children with severe reading acquisition difficulties— especially dyslexics — when the orthography is transparent, such as in Spanish or German. This does not seem to be the case (for Spanish see, Jimenez-Gonzalez and Ramirez-Santana, 2002; for German, Wimmer, 1993). However, the level of the reading impairment of the dyslexics seems to be related to the level of opacity of the orthography, the reading impairment of English dyslexics being higher than that of French dyslexics which, in turn, is higher than that of Italian dyslexics, for example (see Paulesu et al., 2001; see also Ziegler et al., 2003, for a review). In addition, whatever the level of opacity of the orthography, the reading impairment of the dyslexics mostly shows up in pseudoword reading, according to accuracy scores and processing time in deep orthographies and primarily according to processing time in shallow orthographies (see Sprenger-Charolles et al., 2000; Jimenez-Gonzalez and Ramirez-Santana, 2002). Thus, if the quality of the lexical procedure depends on that of the sublexical procedure, since the sublexical procedure of the dyslexics is generally impaired it is expected that their lexical procedure will be impaired as well. Therefore, we should not find dissociated profiles in developmental dyslexia.

2.1.2. Written word processing in dyslexia

Studies concerning adults who became dyslexic following brain damage have revealed phonological dyslexics, suffering from a specific deficit in the sublexical procedure (Beauvois and Derouesné, 1979), and surface dyslexics, with a specific deficit in the lexical procedure (Coltheart et al., 1983). As we have already indicated, these dissociated profiles should not be found with developmental dyslexics, i.e. those whose deficit shows up during reading acquisition, and is not the result of a brain damage. Nonetheless, single case studies have highlighted cases of phonological dyslexia (e.g. Campbell and Butterworth, 1985; Snowling et al., 1986; Valdois et al., 2003) and of surface dyslexia (e.g. Hanley et al., 1992; Valdois et al., 2003). However, a close examination of these studies suggests that the so-called phonological dyslexics also suffer from an orthographic deficit while the phonological skills of the so-called surface dyslexics are also impaired (see for a short review Sprenger-Charolles and Colé, 2003). In addition, these single case studies pose two major problems. Firstly, when choosing to work only with dissociated profile dyslexics, those who suffer from a double deficit are excluded de facto and, secondly, the prevalence of the different profiles cannot be determined. To resolve these two difficulties, we should rely on multiple cases studies in which dyslexics not selected to fit a certain profile are included, i.e. all the dyslexics should be incorporated in the study and their profiles then investigated.

In such multiple case studies, two main methods were used to analyze the results: the “classical” method and the “regression” method. More specifically, with the classical method, children are classified as dyslexics when their reading scores are at least 1 standard deviation (S.D.) below the mean of average readers. The dyslexics are called phonological dyslexics when only their phonological route is impaired, as assessed by their pseudoword reading scores; they are called surface dyslexics when only their lexical route is impaired, as assessed by the reading of irregular words. When both routes are impaired, the dyslexics are said to have a mixed profile or a double deficit. With the regression method, the performances of average readers for irregular word reading in comparison with their performances for pseudowords—and vice versa—are taken as the reference which allows the drawing of two regression lines and their confidence intervals (CI). The first line spots the children who have a phonological deficit, i.e. those whose performances for the reading of pseudowords are outside the CI. The second spots those who have an orthographic deficit, in this case, those whose performances are outside the CI for irregular words. The children whose performances are outside the CI in both cases have a mixed profile; those whose performance are outside the CI in only one of the comparisons have a dissociated profile, either a phonological, or a surface profile.

These two methods were used in studies carried out with English dyslexics (Castles and Coltheart, 1993; Manis et al., 1996; Stanovich et al., 1997) and French dyslexics (Génard et al., 1998; Sprenger-Charolles et al., 2000), the indicator of reading efficiency being solely the accuracy of the response except in one study (Sprenger-Charolles et al., 2000), in which the latency time of the correct responses was assessed as well.

With the classical method, when compared to same chronological age average readers, both reading procedures appear to be deficient for most dyslexics, the number of dissociated profiles being fairly small. However, although as many phonological dyslexics as surface dyslexics were found in the English studies which rely on accuracy (Castles and Coltheart, 1993; Manis et al., 1996; Stanovich et al., 1997), this ratio is not the same in the French studies, except for those based on latency time (Sprenger-Charolles et al., 2000). When both indicators are taken into account, as in the study by Sprenger-Charolles et al. (2000), almost all the subjects exhibit a mixed profile.

The same results were re-examined using the regression method, except in the study by Sprenger-Charolles et al. (2000) in which this method could not be used other than with latency time (due to ceiling effects for irregular word accuracy scores by average readers). With this method, as opposed to the results observed with the classical method, in the five studies examined, the proportion of dissociated profiles appeared high and the proportion of mixed profile dyslexics low. However, the percentage of dissociated profiles varied greatly with the studies, the percentage of phonological dyslexics being higher than that of surface dyslexics in two studies (Castles and Coltheart, 1993; Sprenger-Charolles et al., 2000), equal in two others (Manis et al., 1996; Stanovich et al., 1997), and remarkably inferior in another (Génard et al., 1998). In addition, a high number of dyslexics who had no deficit at all were found (over 25% in three studies). The regression method therefore does not seem to be very robust. Consequently, in order to examine the profile of a dyslexic, it seems better to rely on the classical method.

It emerges from the earlier analyses that the dissociated profiles are anything but stable; their number varies according to the analysis methods, the measurements used, and the language. The differences between the English and the French studies may be explained by linguistic factors. Because the grapho-phonemic correspondences are more regular in French than in English (Peereman and Content, 1999), French-speaking dyslexics can overcome their difficulties by making use of the sublexical procedure more easily than English-speaking dyslexics can. This helps to explain why fewer phonological dyslexics were found in French than in English, at least in studies which took only accuracy into account (cf. the results obtained using the classical method and the regression method in the study by Génard et al., 1998, and the results obtained with the classical method in the study by Sprenger-Charolles et al., 2000). However, when the classification of the French-speaking dyslexics was established using the latency time of correct responses (cf. Sprenger-Charolles et al., 2000), whatever the method of analysis, there were as many phonological dyslexics as in the English language studies relying on accuracy (cf. Castles and Coltheart, 1993; Manis et al., 1996; Stanovich et al., 1997). These data suggest that the French-speaking dyslexics would be able to learn to use the grapho-phonemic correspondences reasonably correctly, their phonological deficit expressing itself principally by the slowness of this operation. This accounts for the differences between the French studies which rely on accuracy and those which take processing time into account.

It is also to be noticed that, in the five studies cited, in the comparison between dyslexics and younger average readers of the same reading level, surface profiles—though not phonological profiles—disappeared almost completely. Only phonological dyslexia would therefore correspond to a developmental deviance (see Bryant and Impey, 1986). A similar result was found in the studies involving undifferentiated groups of dyslexics, in which only the phonological reading skills of the group of dyslexics were found to be impaired when compared to those of the group of reading level controls (see the meta-analyses of Rack et al., 1992 and of Van Ijzendoorn and Bus, 1994).

2.2. Skills associated with learning to read

Apart from reading, the greatest deficits found in dyslexics concern the different abilities which imply phonological processing such as phonemic awareness, phonological short-term memory, and rapid naming (cf. Ramus et al., 2003; Snowling, 2000).

2.2.1. Phonemic awareness

To use the sublexical procedure in an alphabetic writing system, it is necessary to be able to make connections between the graphemes (written sublexical units) and the phonemes (the corresponding oral units), the latter being not easily identified in spoken language for reasons of coarticulation; for example, "car" is pronounced in a single articulatory movement. Before seeing this word written, made up of three letters, the child may not be conscious of its phonemic structure. This could explain why phonemic awareness, but not syllabic awareness, develops with reading acquisition (Liberman et al., 1974). This phenomenon cannot be explained simply by maturation because identical results have been obtained when comparing illiterate and literate adults (Morais et al., 1986; Morais et al., 1979). However, longitudinal studies have also highlighted the fact that early phonemic awareness skills, assessed in kindergarten, predicted later reading skills (Kirby et al., 2003; Parrila et al., 2004; Schatschneider et al., 2004). Similarly, it has been found that a deficit in phonemic awareness may be observed in future dyslexics, even before they start learning to read (Lundberg, and Hoien, 1989; Wimmer, 1996; Scarborough, 1990; Sprenger-Charolles et al., 2000).

The relationships between phonemic awareness and reading are therefore bi-directional, which explains why the phonemic awareness deficit of the dyslexics can never be totally compensated (Bruck, 1992; Elbro et al., 1994; Fawcett and Nicolson, 1994).

2.2.2. Short-term phonological memory

Another skill called for by the sublexical procedure is short-term phonological memory (Baddeley, 1986; Baddeley and Wilson, 1993; Baddeley et al., 1998; see also Lecocq, 1991). Indeed, to read words via this procedure, the reader has to blend the phonemic units resulting of the grapho-phonemic decoding process and, thus, to store them in his/her short-term memory. Consequently, short-term memory should have an effect on learning to read. It remains to be known which kind of memory is involved in this reading acquisition. It seems rather clearly established now that it is a deficit in phonological short-term memory, rather than in non-verbal visual memory, which is linked to dyslexia. Researchers from the Haskins Laboratory were the first to establish this fact (Liberman et al., 1982; Mann and Liberman, 1984; Rapala and Brady, 1990; see also McDougall et al., 1994; Sprenger-Charolles et al., 2000). However, the results of several longitudinal studies suggest that the impact of this memory on reading acquisition is less than that of phonemic awareness skills (Lecocq, 1991; Parrila et al., 2004; Wagner et al., 1994; Wagner et al., 1997). Identical tendencies were found when the metaphonemic and short-term memory abilities of adult dyslexics were compared to those of average readers of the same age or same reading level (Pennington et al., 1990). More surprisingly given the long history of the relationships between short-term memory deficits and reading, in other studies no difference in short-term memory was found between average readers and children with reading disabilities (Bowers, 1995; Wimmer, 1993).

2.2.3. Rapid naming

A naming deficit, which would appear particularly when this skill is assessed by a task of serial naming under a temporal constraint, is also assumed to be characteristic of dyslexics (Wolf, 1991). In these tasks, frequent items (images of objects, color patches, letters or numbers) are presented several times on the same page, in a different order (usually five items repeated 10 times), the subjects being required to name the items as rapidly and as accurately as possible. These tasks assess the speed and accuracy of access to the oral lexicon, abilities which could also be required when reading via the sublexical procedure. Indeed, after having operated the grapho-phonemic decoding and blended the result of that operation, the child has to locate the word in his/her oral lexicon, an operation that could be facilitated by a rapid and accurate access to this lexicon. The results indicate that reading disabled subjects are slower than controls (Ackerman and Dykman, 1993; Bowers and Swanson, 1991; Denkla and Rudel, 1976; Felton and Brown, 1990; Wimmer, 1993; Wolf and Obregon, 1992). However, according to some studies, the predictive power of rapid naming on reading was found to be lower than that of phonological awareness (Kirby et al., 2003; Wagner et al., 1994, 1997) while other studies suggest the contrary (Ackerman and Dykman, 1993; Bowers and Swanson, 1991; Felton and Brown, 1990). These inconsistencies may be explained by the fact that rapid naming deficits would be linked to reading level mainly for the poorest readers (Bowers, 1995; McBride-Chang and Manis, 1996), which might be due to the fact that this task involves articulation speed, which increases with age but attains its maximum threshold rapidly in normal subjects, and more slowly in children with language disabilities (Henry and Millar, 1993; Kail and Park, 1994). Therefore, for subjects with no such strong disability, individual differences rapidly disappear. Differences in the tasks used would also explain some discrepancies between the results. For example, naming speed was not found to be a good predictor of the reading level after controlling for decoding skills when the task used involves the naming of letters or of digits (Wagner et al., 1997), but not when the task used involves the naming of colors (Parrila et al., 2004). As explained by Wagner et al. (1997), it is possible that including alphanumeric stimuli makes these tasks "mere proxies for individual differences in early literacy and print exposure" (p. 476), explaining why the predictive power of these skills vanishes when decoding skills are taken into account. This seems not to be the case for the ability to name colors (or objects), probably because this ability is less dependent on the reading level (see Parrila et al., 2004).

3. Overview of the study

The choices which took precedence during the construction of EVALEC were made according to the literature. In particular, we have drafted several reading tests to closely examine how written words are processed. We have emphasized assessments of the sublexical procedure, as the procedure for which dyslexics exhibit the greatest deficit. The goal of the other tests is to assess the skills linked to learning to read: phonemic awareness (using a control task of syllabic awareness), phonological short-term memory, and naming speed. For the last test, we have chosen a rapid serial naming test which is taken under two conditions: using images and in reading. The impact of a time constraint in reading and non-reading contexts can thus be assessed. The other main methodological choices are presented below.

3.1. Two possible comparisons: chronological age and reading level

EVALEC allows users to assess how the reading procedures function in dyslexics compared to average readers either of the same chronological age, or of the same reading level, the latter comparison allowing users to determine if dyslexia corresponds to a developmental delay or to a deviance. To be able to carry out these two assessments, tests adapted to each age group must be available, since the dyslexics must be assessed using a battery of tests corresponding both to their chronological age and their reading age. The solution chosen is the most economical: the same battery of tests was used with all of the children, which implies some constraints. First, to avoid "floor" effects with the youngest children, we were obliged to choose items known by them. Thus, only written words that first graders might have encountered were taken into account. This was verified with the aid of MANULEX (Lété et al., 2004). Second, because the frequency of a word is not the same for 7 and for 10-year-olds, we manipulated lexicality effects rather than frequency effects.

3.2. Assessment of the sublexical procedure and of the lexical procedure

We relied on the reading of items assumed to be the best signature of one of the two reading procedures: pseudowords which are not analogous to real words for the sublexical procedure, and irregular words for the lexical procedure. This is possible in French, as in English, but much more difficult in Spanish and Italian. In the latter two cases, it is only possible to compare the reading of pseudowords to that of frequent regular words. However, regular words can be read using the lexical route as well as the sublexical route. To assess whether words and pseudowords are read using the same procedure, we examined the effects of sublexical factors. If such factors have an effect on pseudowords, and not on words, we could conclude that these two types of items are not read using the same procedure. Our battery of tests therefore includes regular words and pseudowords of different levels of orthographic complexity: items containing only single letter graphemes versus items containing a digraph versus items containing a grapheme whose pronunciation depends on the context, such as 'c' or 'g'.

The preceding task is a reading aloud task, which makes the reader produce an oral response, and therefore use the phonological representation of the written word, which can bias the results. We have therefore constructed a test to assess the efficiency of the lexical procedure in silent reading. The test retained is the one which, according to our earlier results (cf. Sprenger-Charolles et al., 2003; see also Olson et al., 1994), seems to be the best adapted: an orthographic choice test with three possible choices, a correctly written frequent word, and two pseudowords, one which has the same pronunciation but not the same visual form as the target word, and one which is visually close to the target word but is not pronounced the same way (for example, auto, oto, outo).

3.3. Two measurements: accuracy and processing speed

The efficiency of word reading has to be measured not only by its accuracy but also by its speed. However, in statistical analyses, response speed can only be taken into account when the number of correct responses is sufficiently high (over 50% according to Olson et al., 1994). In reading aloud tasks, the indicator most often retained is the latency time of the correct responses, i.e. the delay between presentation of the word on the computer screen and the onset of its pronunciation. For this computation, as in all our earlier studies (Sprenger-Charolles, 1994; Sprenger-Charolles et al., 2003), a voice key was not used because of the limitations of this methodology (see, Kessler et al., 2002; Rastle and Davis, 2002). Indeed, the response must reach a certain threshold in order to be recorded correctly and the voice key reacts the same way to a noise or to speech. Thus, we worked with a sound card which recorded responses. The detection of the onset of the correct response was made using the speech signal. In addition, because the latency time computation may be biased by the nature of the phoneme corresponding to the initial grapheme (cf. Kessler et al., 2002; Rastle and Davis, 2002), the items of the different comparisons were matched for this variable.

Some researchers have also pointed out that latency time can capture a part of the pre-programming of articulatory codes (for a discussion, see Marmurek and Rinaldo, 1992 and, for French data supporting this hypothesis, Sprenger-Charolles, 1994; Sprenger-Charolles et al., 2003). To test this hypothesis, latency time and duration of vocal response production have to be compared. Thus, we have taken into account these two measures, both of which were examined by way of analysis of the speech signal relying on a semi-automatic program to detect the beginning and the end of the responses (see Appendix C).

4. Method

4.1. Participants

Insofar as this battery of tests was meant to be used as a diagnostic assessment of dyslexia, the children were retained for this study according to criteria which characterize the child as dyslexic, and not simply as a poor reader. Thus, the children selected were monolingual French speakers without language or sensori-motor impairments and without psychological problems. They had no cognitive deficit, neither verbal (assessed, according to chronological age, either by the TVAP, the EVIP or by the vocabulary test of the WISC), nor non-verbal (assessed using RAVEN matrices or WISC cubes). The children included in this study were in the grade level corresponding to their age; only those officially diagnosed as dyslexic were excluded. Approximately 100 children per grade level were seen, some of them in Paris and the surrounding region, others in other parts of the country (Brittany, Lorraine, Savoy). In each level, the children came from several schools and several classes (at least 9); this helps to neutralize the variable of teaching as much as possible. Hundred children from the first grade (G1), 120 from the second grade (G2), 105 from the third grade (G3) and 73 from the fourth grade (G4) were examined. The characteristics of the population are presented in Table 1.

Table 1
Characteristics of the population (mean and S.D. in parentheses)

	Number of subjects and gender			Chronological age	Reading age	Non-verbal cognitive level		Vocabulary level (or verbal cognitive level)		
	Total	Boys	Girls	in months	in months	RAVEN	WISC (Khos)	TVAP	EVIP	WISC
Grade 1	100	52	48	83.4 (3.6)	87.4 (7.4)	25.1 (4.5)		47.0 (5.6)		
Grade 2	120	66	54	94.2 (3.2)	97.0 (11.1)	26.2 (3.7)		49.2 (5.0)		
Grade 3	105	57	48	103.1 (3.5)	105.3 (16.5)		112.6 (19.6)			112.4 (24.7)
Grade 4	73	34	39	118.8 (3.4)	116.1 (18.5)	30.6 (3.3)			129.3 (11.0)	

4.2. Tasks

4.2.1. Pre-tests

The verbal IQ was assessed, according to chronological age, either by the TVAP (Deltour and Hupkens, 1980), the EVIP (Dunn et al., 1993), or by the vocabulary test of the WISC (Weschler, 1990). Non-verbal IQ was assessed using RAVEN matrices (Raven, 1947, rééd. 1981) or WISC cubes.

The reading level was assessed via a standardized test: "l'Alouette" (Lefavrais, 1967). The children have a 265 word text to read as rapidly and as accurately as possible. The text includes rare words and some spelling traps: items with silent letters (temps /tã/, nids /ni/), contextual graphemes (gai-geai), and items which are phonologically similar (Annie-amie). This test also tracks contextual anticipation which characterizes the youngest and least skilled readers (Stanovich, 1984; Perfetti et al., 1979). The text contains fixed expressions which are modified ("au clair de lune" instead of the usual "au clair de la lune") or words similar to those that would be predicted by the context: e.g. "poison" rather than "poisson" (fish) after "lac" (lake). We note either the reading time (if under 3 min) or the number of words read in 3 min. Reading level is calculated, taking speed and number of errors into account. This level is then converted into reading age (from 6 to 14 years).

4.2.2. EVALEC: written word processing (WWP)

All of the WWP tests are computerized. In particular, the children's vocal responses are recorded and stored in individual files. This allows users to thoroughly verify the correctness of the responses and, after, to determine the beginning and the end of the production for each correct response by way of the analysis of the speech signal with the semi-automatic program described in Appendix C.

4.2.2.1. LEVORT. The aim of this test is to assess the efficiency of the lexical procedure in reading. It contains 48 frequent words from four levels of orthographic regularity (12 items per category). The items in the first level of regularity are regular words composed of simple graphemes (R1: one letter for one phoneme); those of the second level integrate a digraph (R2: ch, ou, on...), those of the third level contain a grapheme whose pronunciation depends on the context (R3: c, g); and those of the fourth level are irregular (R4). In this last case, as in our earlier studies (in particular, see Sprenger-Charolles et al., 1998b), the irregularity never concerns only the last consonant of a word which is often silent in French orthography, due to the fact that morphological markers are usually not pronounced. This is the case, for most of the flexional marks, for example, those for the plural ("tables" is pronounced in the same way as "table" /tabl/), as well as for most of the derivational marks, as the "p" in "camp" (/kâ/), from which the word "camper" (/kâpe/, "to camp") is derived. Since we would like to track non-phonological reading processing as well as possible via the reading of irregular words, we have avoided words with such irregularities, given that, on the one hand, when the final consonant of a word is a silent letter, it is possible to read this kind of irregular word correctly by way of sequential phonological reading processing which stops just before the last letter, as in "porc" /por/ (pig), "banc" /bâ/ (bench), "tabac" /taba/ (tobacco), etc. On the other hand, phonological processing relying on usual grapheme-phoneme correspondences will generate the correct pronunciation of words ending with non-silent consonants, such as "ours" /urs/ (bear), "cassis" /kasis/ (blackcurrant), "iris" /iris/, "déficit" /deficit/ (deficit), granit" /granit/ (granite), etc.

In each level of orthographic difficulty, the items are matched for length (number of letters, phonemes, and syllables) as well as for orthographic frequency (frequency of the bigrams, Content and Radeau, 1988) and for lexical frequency (calculated using MANULEX, Lété et al., 2004). In order to calculate the latency time, the items in each level were also matched according to the phoneme corresponding to their initial grapheme. The list of the items and the data concerning the matching criteria are found in Appendix A1. The children were instructed to respond as accurately and as quickly as possible. We specified them not to pronounce the word before they had it "right in their minds". Practice items were used to make sure that the child had understood the instructions.

4.2.2.2. LEXORT (lexicality and orthographic complexity) and LEXLENGTH (lexicality and length). The aim of these two tests is to assess the efficiency of the sublexical procedure as compared to the lexical procedure by the reading of regular words and of pseudowords matched in different levels of orthographic complexity (LEXORT) and by the reading of irregular words and pseudowords of different lengths (LEXLENGTH). LEXORT uses the words of the first three levels of regularity of LEVORT and pseudowords designed on the same principle (R1: tomate-pitode; R2: malin-nurin; R3: cinéma-cirate, 12 items per category). The items in each level of orthographic complexity are matched for length (number of letters, phonemes, and syllables), orthographic frequency (bigram frequency, Content and Radeau, 1988), and for their initial grapheme. LEXLENGTH uses 20 irregular words and 20 pseudowords, 10 short and 10 long (a mean of four versus eight letters, for example: écho-opha versus orchestre-orphade), matched on the same criteria as the preceding items. Since there are not many frequent irregular words in French, it was not possible to match up the short and long irregular words for lexical frequency; the long irregular words are less frequent than the short irregular words, according to the MANULEX data for the first grade (Lété et al., 2004).

The list of items used for LEXORT and LEXLENGTH and the data concerning the matching criteria are presented in the appendices A1 (LEXORT) and A2 (LEXLENGTH). The instructions were the same as for LEVORT, except that, for pseudowords, we specified that the task dealt with ″alien″ words.

4.2.2.3. Orthographic choice (trio). The aim of this test is to assess the efficiency of the lexical procedure in a silent reading task while taking into account the accuracy and speed of detection of a correctly spelled frequent word presented at the same time as a pseudohomophone and a visual foil (9 items). The pseudohomophones ("oto", "vélau", "rouje"...) have the same pronunciation as the correct word ("auto" [car], "vélo" [bike], "rouge" [red]) but not the same visual form: they have one letter more or one letter less, except for "rouje", but in this case the letter "j" has a lower number of visual features than the letter used when the target word is correctly written ("g"). The visual foils ("outo", "véla", "rouqe") have the same number of letters as the target word with which they share a strong visual resemblance. A specially designed font was created for this purpose. For example, the letter "a" was represented by an half of circle closed by a vertical line, thus the visual shapes of the visual foils were closer to those of the correct words, than were those of the phonological foils. The mean trigram frequency of the two types of foils is similar (89 for the pseudohomophones and 101 for the visual foils, Content and Radeau, 1988). The items used are listed in the Appendix A3. For each triplet, the correct word was presented at the same time as the two foils, on the same line, in random order. The children were asked to choose the correct word by pressing a key on the computer keyboard (the three keys at the lower right side of the keyboard). Practice items were used to make sure that the children had understood the instructions.

4.2.2.4. Phonological awareness, phonological short-term memory and rapid naming. For the phonological awareness and short-term memory tests, EVALEC uses only pseudowords in order to avoid biases related to differences in vocabulary level. Similarly, to avoid differences due to the quality of the articulation of the experimenter, the items were recorded beforehand and the children heard them via a computer through headphones. The characteristics of the items are presented in Appendix B.

For the three phonological awareness tests, we used tasks of deletion of the first element of an item: either the syllable or the phoneme. The test concerning the syllable includes 10 tri-syllabic items with a simple syllabic structure (Consonant-Vowel). The other two tests include tri-phonemic pseudowords, 12 with a Consonant-Vowel-Consonant structure and 12 with a Consonant-Consonant-Vowel structure. For the first two tests, the initial consonant is either a plosive or a fricative (half/half). For the third test, a plosive or a fricative is followed by a liquid (4 × 2 items, respectively) and a plosive is either followed or preceded by a fricative (2 × 2 items, respectively). The children were instructed to "eat" the beginnings of "alien" words. The syllabic task was followed by the two phonemic tasks, the CVC items being presented before CCV items.

To assess phonological short-term memory, we have designed a list of three to six syllable pseudowords which are composed of six items for each length (three including only CV syllables, three with CVC syllables). The items were presented one by one, according to their length (the six, three-syllable items first, followed by the six, four-, five-, and six-syllable items). The children had to repeat each item as accurately as possible, with no time constraint. The series for which the child gives at least four correct responses out of a possible six are considered to be successful.

Naming speed was assessed using color names, because the ability to name colors is assumed to be less dependent on reading level than the ability to name letters or even digits (see Parrila et al., 2004). The children had to name six colors presented eight times in a different order, as rapidly and as accurately as possible. Three items have a CVC structure: rouge (red), jaune (yellow), vert (green); the other three having a CCV structure: bleu (blue), blanc (white), gris (grey). The stimuli are either color patches or written color names. The order (color patches vs. written color names) is counterbalanced. Before each test, we showed the six color patches and the six names and asked the child to name them (in case of an error, the examiner gave the correct response and verified that the child had understood). As in previous studies, for each of the two tasks, the items were presented on a sheet of paper (six rows of eight items).

4.3. Procedure

The test sessions took place at the end of Grades 1–4. Children were tested individually in a quiet classroom in two (or three) test sessions. In the first test session (lasting from 20 to 30 min), they were presented the pre-tests: Verbal IQ (TVAP or EVIP or the vocabulary test of the WISC); Non-verbal IQ (RAVEN matrices or WISC cubes); Reading level (Alouette). EVALEC was presented in the second test session, in which all the reading aloud tasks were presented together (LEVORT word reading first followed by LEXORT pseudoword reading, and then LEXLENGTH word reading followed by LEXLENGTH pseudoword reading. This session lasted 15–30 min, depending on the reading level of the children. Depending on the length of this session, all the other tasks (Orthographic choice, Phonological awareness, Phonological short-term Memory and Rapid Naming) were presented just after or in a separate test session, lasting from 10 to 15 min.

5. Results

A first analysis presents the results of the ANOVAs, a second examines the correlation patterns between the different tests and a third considers the factors which account for reading level. To facilitate the reading of the text, only significant differences according to the P <0.01 or P <0.05 (two or one stars, respectively) thresholds are reported.

5.1. ANOVAs: tests of written word processing

5.1.1. Words of different levels of regularity (LEVORT)

The main goal of this test was to assess the effect of sublexical factors in word reading: the effect of orthography (by way of comparison between items from four levels of regularity), the effect of regularity (comparison between all of the regular words and the irregular words: R1–R2–R3 versus R4), the effect of graphemic length (comparison between regular words containing a digraph and those with only simple graphemes: R2 versus R1), and the effect of graphemic complexity (comparison between regular words containing a contextual grapheme and simple regular words: R3 versus R1).

Since the words used in LEVORT are frequent words, we should not find sublexical factor effects with the oldest children, who are assumed to have reached a stage near that of expert (cf. for English, Backman et al., 1984; Waters et al., 1984 and, for French, Sprenger-Charolles et al., 2003). These effects, in particular those of regularity and graphemic complexity should, on the other hand, be strong in the youngest children, for accuracy as well as for latency time. For vocal response duration, we expected R2 items to be pronounced more rapidly than R1 items, because, if R1 and R2 items contain the same number of letters, those including a digraph (R2) contain a lesser number of phonemes than those with only simple graphemes (R1). If this effect shows up for accuracy and latency time as well, it would indicate that the children processed graphemes rather than letters. The results of the three indicators studied (correct responses, latency times and duration of vocal productions) are presented in Table 2. A two-factor ANOVA was run (Grade level factor: G1–G4; Orthography factor: R1–R4).

5.1.1.1. For correct response, the two main effects were significant (Grade level, $F(3,394) = 99.18$**; Orthography, $F(3,1182) = 212.25$**). The effect of orthography is explained by the superiority of regular words to irregular words (Regularity effect, $F(1,394) = 325.58$**), and by the inferiority of words containing a contextual grapheme compared to those containing only simple graphemes (Graphemic complexity effect, $F(1,394) = 40.55$**), these being read less well than items containing a digraph (Graphemic length effect, $F(1,394) = 5.72$*). The effect of sublexical factors varied with grade level (Interaction, $F(9,1182) = 47.3$**). As expected, the negative effect of irregularity decreased between G1 and G4 (end of G1–G4, respectively, –28.5%, –12.9%, –3.2% and –3.4% difference between the mean percentages of correct responses for irregular words minus regular words; $F(1,99) = 191.70$**; $F(1,119) = 105.48$**; $F(1,104) = 15.68$**; $F(1,72) = 15.86$**), and the negative effect of graphemic complexity progressively vanished to become nonsignificant for G4 children (G1–G3: –8.8%, –3.3% and –0.95% difference between the mean percentages of correct responses for the regular words with a contextual grapheme minus those without any contextual grapheme; $F(1,99) = 29.60$**; $F(1,119) = 12.99$**; $F(1,104) = 4.98$*). Finally, while the graphemic length effect was positive and significant when all the groups were considered together, it never reached the threshold of significance in the specific comparisons by grade level (G1–G4: +1.6%, +1.4%; +0.2% and + 1.2% difference between the mean percentages of correct responses for the regular words with a digraph minus those without digraphs).

5.1.1.2. For the latency of correct responses, only the results of children who had at least 50% correct responses were ana-

Table 2
LEVORT: correct responses, latency times and vocal response duration for words of different levels of regularity* (mean and S.D.)

	Correct responses (mean percentage)				Latency time (ms)				Vocal response duration (ms)			
	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4
Mean												
Grade 1	88.58	90.17	79.83	57.67	1359	1207	1376	1519	715	606	713	668
Grade 2	93.96	95.35	90.69	80.42	1062	963	1080	1099	710	641	701	675
Grade 3	98.89	99.05	97.94	95.40	847	783	841	934	645	587	630	615
Grade 4	96.99	98.15	97.57	94.21	787	776	793	831	615	566	596	579
Standard deviation												
Grade 1	12.29	10.94	18.85	26.82	540	458	559	726	141	104	163	148
Grade 2	7.79	6.64	9.91	15.83	490	359	504	490	130	105	124	133
Grade 3	3.67	2.66	4.44	8.57	332	228	288	401	104	99	102	98
Grade 4	6.56	4.65	4.70	7.17	222	191	219	244	98	91	96	95

*R1: one letter for one phoneme; R2: items with a digraph (ch, ou...); R3: items with a contextual grapheme (c–g, a + i + ll); R4: items with a grapheme with an exceptional pronunciation (x, sc).

lyzed. Therefore only a part of the G1 children's results could be used (67/100). The main effects were significant (Grade level, *F*(3,361) = 31.18**; Orthography, *F*(3,1083) = 73.44**). The effect of orthography is due to the fact that regular words were read faster than irregular words (regularity effect, *F*(1,361) = 78.71**) as well as words including a digraph as compared to those with only single letter graphemes (Graphemic length effect, *F*(1,361) = 83.66**); the graphemic complexity effect was not significant. The influence of sublexical factors varied according to grade level (Interaction, *F*(9,1083) = 7.86**). The regularity effect always emerged to the detriment of irregular words (for G1–G4, respectively, +206, +64, +110 and +46 ms difference between the mean latency times for irregular words minus regular words; *F*(1,66) = 25.23**, *F*(1,119) = 14.21**; *F*(1,104) = 29.57**; *F*(1,72) = 18.45**) whereas graphemic length has a positive impact on performances only in the first three grade levels (G1–G3, respectively, –152, –99 and –64 ms difference between the mean latency times for the regular words with a digraph minus those without digraphs; *F*(1,66) = 32.52**; *F*(1,119) = 30.40**; *F*(1,104) = 19.43**).

5.1.1.3. For vocal response duration, analyses were carried out on the same G1 children as above. The main effects were significant (Grade level, *F*(3,361) = 15.51**; Orthography, *F*(3,1083) = 123.50**). The effect of orthography is explained by the superiority of words containing a digraph compared to those containing simple graphemes (Graphemic length effect, *F*(1,361) = 420.55**), the latter being, surprisingly, pronounced less rapidly than the items containing contextual graphemes, *F*(1,361) = 9.04**. Surprisingly again, irregular words were pronounced more rapidly than regular words (Regularity effect, *F*(1,361) = 6.46*). The influence of sublexical factors varied with grade level (Interaction, *F*(9,1083) = 5.47**). As expected, graphemic length had a positive impact on the performances of all the children (G1–G4, respectively, –109, –68, –57 and –49 ms difference between the mean vocal response durations for the regular words with a digraph minus those without digraphs; *F*(1,66) = 80.77**; *F*(1,119) = 134.57**; *F*(1,104) = 163.12**; *F*(1,72) = 107.46**). The regularity effect was significant only for the G4 children who pronounced irregular words more rapidly than regular words (–13 ms; *F*(1,72) = 12.89**); the graphemic complexity effect was significant for the same children (G4), but also for the G3 children and was to the benefit of items containing a contextual grapheme (G3 and 4: –15 and –19 ms; *F*(1,104) = 9.04** and *F*(1,72) = 25.70**).

5.1.1.4. To summarize, as expected, the detrimental effect of irregularity appeared for accuracy and latency time, and decreased with age, at least according to the analyses of correct responses. Graphemic complexity had a negative effect on accuracy for the children in the first three grade levels, but not on latency time. The expected positive effect of graphemic length was found for vocal production duration for all grade levels. Graphemic length also had a positive effect on correct responses when all grade levels were considered together, as well as on latency time for the first three grade levels, which suggests that children process graphemes rather than letters. Finally, according to the duration of the vocal response, two unexpected results were found: irregular words were pronounced faster than regular words by G4 children and items containing contextual graphemes were pronounced faster than those containing only simple graphemes by G3 and G4 children.

5.1.2. Words and pseudowords of different levels of regularity (LEXORT)

If the youngest children rely primarily on the sublexical procedure, they should not process words better than pseudowords and the sublexical factors should have the same effect on both types of items. For the oldest readers, we should find a superiority of word reading compared to pseudoword reading and a less marked effect of sublexical factors for words than for pseudowords. The results are presented in Table 3. For the three factors studied (correct responses, latency time and duration of the vocal production), a three-factor ANOVA was carried out (Grade level: G1–G4; Lexicality: Words versus Pseudowords; Orthography: R1–R3).

Table 3
LEXORT: correct responses, latency times and durations of vocal response for words and pseudowords of different levels of regularity* (mean and S.D.)

	Correct responses (mean percentage)						Latency time (ms)						Vocal response duration (ms)					
	Words			Pseudowords			Words			Pseudowords			Words			Pseudowords		
	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3	R1	R2	R3
Mean																		
G1	88.58	90.17	79.83	70.08	73.33	56.58	1359	1207	1376	1689	1646	1727	715	606	713	890	736	881
G2	93.96	95.35	90.69	79.10	79.93	68.26	1062	963	1080	1451	1444	1492	710	641	701	848	765	858
G3	98.89	99.05	97.94	94.76	96.03	87.86	847	783	841	1304	1264	1348	645	587	630	738	666	742
G4	96.99	98.15	97.57	89.38	89.95	83.33	787	776	793	1169	1202	1240	615	566	596	680	613	689
S.D.																		
G1	12.29	10.94	18.85	20.42	19.17	21.79	540	458	559	744	660	760	141	104	163	258	165	265
G2	7.79	6.64	9.91	16.88	16.19	18.05	490	359	504	655	607	696	130	105	124	191	156	208
G3	3.67	2.66	4.44	8.67	8.10	12.07	332	228	288	582	522	587	104	99	102	124	112	138
G4	6.56	4.65	4.70	9.14	10.84	13.89	222	191	219	465	505	543	98	91	96	124	95	120

*R1: one letter for one phoneme, R2: items with a digraph (ch, ou...), R3: items with a contextual grapheme (c-g, a + i + ll), R4: items with a grapheme with an exceptional pronunciation (x, sc).

5.1.2.1. For correct responses, the three main effects were significant (Grade level, $F(3,394) = 90.93^{**}$; Lexicality, $F(1,394) = 727.99^{**}$; Orthography, $F(2,788) = 153.46^{**}$), as were the three two-way interactions (Grade level-Lexicality, $F(3,394) = 42.57^{**}$; Grade level-Orthography, $F(6,788) = 12.09^{**}$; Orthography-Lexicality, $F(2,788) = 36.08^{**}$), though not the three-way interaction. The lexicality effect favored words and was most strongly marked for the youngest children (for G1–G4, respectively, +19.5%, +17.6%, +5.7% and +10% difference between the mean percentages of correct responses for the words minus pseudowords; $F(1,99) = 279.12^{**}$; $F(1,119) = 292.02^{**}$; $F(1,104) = 58.70^{**}$; $F(1,72) = 109.20^{**}$). The interaction between orthography and lexicality is explained by an inferiority of simple regular items compared to those containing a digraph (Graphemic length effect for words and pseudowords, –1.1% and –1.5%; $F(1,394) = 5.72^{*}$; 4.71^{*}) and especially, by the inferiority of items containing a contextual grapheme compared to those containing only simple graphemes, this detrimental graphemic complexity effect being more marked for pseudowords (–9.3% difference between the mean percentages of correct responses for the pseudowords with a contextual grapheme minus those without contextual graphemes as compared to –3.1% for the words; $F(1,394) = 40.55^{**}$; 136.60^{**}).

5.1.2.2. For the latency time of correct responses, in the first grade, analyses were carried out only for the 67 children with more than 50% correct responses. The three main effects were significant (Grade level, $F(3,361) = 17.57^{**}$; Lexicality, $F(1,361) = 482.03^{**}$; Orthography, $F(2,722) = 45.22^{**}$), as were the interactions of Orthography-Lexicality and Grade level-Orthography (respectively, $F(2,722) = 9.52^{**}$; $F(6,722) = 3.65^{**}$). The interaction of Grade level-Lexicality and the three-way interaction were not significant. The lexicality effect always favored words (for G1–G4, respectively, –373, –431, –480 and –418 ms difference between the mean latency times for words minus pseudowords; $F(1,66) = 52.62^{**}$; $F(1,119) = 163.95^{**}$; $F(1,104) = 186.94^{**}$; $F(1,72) = 97.71^{**}$). Length and graphemic complexity had the opposite effect on words and pseudowords. The former, significant only for words, favored words containing a digraph, –82 ms; $F(1,361) = 83.66^{**}$. The latter, significant only for pseudowords, had a negative effect on pseudowords containing a contextual grapheme, +49 ms; $F(1,361) = 11.30^{**}$.

5.1.2.3. For vocal response durations, the three main effects were significant (Grade level, $F(3,361) = 24.29^{**}$; Lexicality, $F(1,361) = 399.89^{**}$; Orthography, $F(2,722) = 263.37^{**}$) as well as the three double interactions (Orthography-Lexicality, $F(2,722) = 17.76^{**}$; Grade level-Lexicality, $F(3,361) = 10.69^{**}$; Grade level-Orthography, $F(6,722) = 9.53^{**}$), but not the three-way interaction. The lexicality effect favored words and was always significant (for G1–G4, respectively, –153, –140, –94 and –68 ms difference between the mean vocal response durations for words minus pseudowords; $F(1,66) = 65.52^{**}$; $F(1,119) = 137.00^{**}$; $F(1,104) = 274.84^{**}$; $F(1,72) = 69.18^{**}$). The interaction between orthography and lexicality may be explained by the fact that the vocal response times were longer for the simple regular items than for those containing a digraph, no matter what the nature of the items (graphemic length effect for words and pseudowords, respectively, +71 and +94 ms; $F(1,361) = 420.54^{**}$; 271.76^{**}) while graphemic complexity had a negative effect only on words, those containing a contextual grapheme being, surprisingly, pronounced more rapidly (–11 ms; $F(1,361) = 9.04^{**}$).

5.1.2.4. To summarize, for the three indicators studied (accuracy, latency and vocal response duration), the lexicality effect always favored words. The effect of the sublexical factors manipulated, whether graphemic complexity or graphemic length, appeared differently according to lexicality and the indicator. The presence of a digraph facilitated the accuracy and duration of response production for both words and pseudowords, but, for latency time, this effect was significant only for words. Graphemic complexity had a negative effect on correct responses for words and even more so for pseudowords. For latency time, this effect was significant only for pseudowords and, for vocal response duration, only for words. However, graphemic complexity had a negative effect on latency time and a positive effect on the duration of vocal response production.

5.1.3. Short and long irregular words and pseudowords (LEXLENGTH)

If irregular words are less well processed than pseudowords and if long pseudowords are less well processed than short pseudowords, we can infer that the subjects rely primarily on the sublexical procedure. This is the result expected for the youngest children. With older children, we should find the opposite tendency: a superiority of words to pseudowords, the effect of length being less marked for words. For the three indicators studied, a three-factor ANOVA was carried out (Grade level, G1–G4; Lexicality, Words versus Pseudowords; Length: Short and long items) (Table 4).

5.1.3.1. For correct responses, the three main effects were significant (Grade level, $F(3,394) = 146.19^{**}$; Lexicality, $F(1,394) = 381.39^{**}$; Length, $F(1,394) = 33.18^{**}$), as were all of the interactions (Lexicality-Length, $F(1,394) = 209.75^{**}$; Lexicality-Grade level, $F(3,394) = 87.24^{**}$; Length-Grade level, $F(3,394) = 10.46^{**}$; Lexicality-Length-Grade level, $F(3,394) = 3.84^{*}$). The effect of lexicality decreased with grade level but, unlike in the preceding analysis, it never favored words (in G1–G4, respectively, –37.1%, –15.1%, –7.5% and –0.6% difference between the mean percentages of correct responses for the words minus pseudowords; $F(1,99) = 391.46^{**}$; $F(1,119) = 93.36^{**}$; $F(1,104) = 26.65^{**}$; $F(1,72) < 1$). The effect of length penalized pseudowords especially, $F(1,394) = 200.88^{**}$ versus

Table 4
LEXLENGTH: correct responses, latency times and durations of vocal response for short and long irregular words and pseudowords (mean and S.D.)

	Correct responses (mean percentage)				Latency time (ms)				Vocal response duration (ms)			
	Irregular words		Pseudowords		Irregular words		Pseudowords		Irregular words		Pseudowords	
	Short	Long	Short	Long	Short	Long	Short	Long	Short	Long	Short	Long
Mean												
Grade 1	33.40	33.20	75.90	64.90	not taken into account							
Grade 2	57.42	60.25	82.25	65.58	1238	1678	1354	1816	558	934	577	1144
Grade 3	76.67	89.05	94.67	86.00	1115	1365	1241	1772	548	818	554	975
Grade 4	81.78	86.58	90.41	79.18	964	1076	1148	1581	492	732	500	839
Standard deviation												
Grade 1	21.00	27.19	19.13	21.25								
Grade 2	17.08	23.24	14.98	18.60	435	1022	529	951	104	250	113	285
Grade 3	15.73	14.38	9.31	16.03	470	786	491	933	131	164	110	238
Grade 4	14.56	13.36	9.64	15.88	289	462	394	726	90	110	75	160

33.07** for words. In fact, length had a negative effect on pseudoword reading at all grade levels (–11.0%, –16.7%, –8.7% and –11.2% difference between the mean percentages of correct responses for long minus short pseudowords; F(1,99) = 36.41**; F(1,119) = 87.24**; F(1,104) = 41.40**; F(1,72) = 39.50**) while, surprisingly, it had a positive effect on word reading for the oldest children (for G3 and G4, respectively, +12.4% and +4.8% difference between the mean percentages of correct responses for long minus short words; F(1,104) = 82.44**; F(1,72) = 9.89**).

5.1.3.2. *For latency time*, because of the low percentage of correct responses by the youngest children for irregular word reading, analyses were carried out only for the three highest grade levels. The three main effects were significant (Grade level, F(2,295) = 6.91**; Lexicality, F(1,295) = 110.51**; Length, F(1,295) = 203.17**), as were all of the interactions (Lexicality-Length, F(1,295) = 26.88**; Lexicality-Grade level, F(2,295) = 8.34**; Length-Grade level, F(2,295) = 3.33*; Lexicality-Length-Grade level, F(2,295) = 7.22**). Unlike in the response accuracy results, the lexicality effect favored words and increased with grade level (–98, –241 and –322 ms difference between the mean latency times for words minus pseudowords in G2–G4, respectively, F(1,119) = 10.60**; F(1,104) = 55.41**; F(1,72) = 101.63**). The negative effect of length penalized pseudowords more than words (respectively, F(1,295) = 233.36** and 76.21**) but was always significant, both for words and for pseudowords. However, the negative impact of length decreased between sessions for words (+441, +251 and +111 ms difference between the mean latency times for long minus short words in G2–G4, respectively; F(1,119) = 39.52**; F(1,104) = 27.56**; F(1,72) = 16.29**), not for pseudowords (+462, +531 and +433 ms difference between the mean latency times for long minus short pseudowords; F(1,119) = 74.80**; F(1,104) = 92.78**; F(1,72) = 78.32**).

5.1.3.3. For vocal response duration, the three main effects were significant (Grade level, F(2,295) = 32.73**; Lexicality, F(1,295) = 179.91**; Length, F(1,295) = 2100.4**), as were all of the interactions (Lexicality-Length, Lexicality-Grade level, Length-Grade level and Lexicality-Length-Grade level, respectively, F(1,295) = 163.92**, F(2,295) = 5.99**; F(2,295) = 41.68**; F(2,295) = 4.52*). The impact of lexicality diminished with grade level but it always favored words (–75, –57 and –35 ms difference between the mean vocal response durations for words minus pseudowords in G2–G4, respectively; F(1,119) = 75.59**; F(1,104) = 80.68**; F(1,72) = 35.70**). The negative effect of length was more marked for pseudowords than for words (F(1,295) = 1618.63** vs. 1140.46**), particularly for the G2 children (+567 ms difference between long and short pseudowords vs. +376 ms between long and short words, F(1,119) = 678.54** vs. 365.18**) but was always significant (for pseudowords versus words in G3 and G4, respectively: +421 vs. +269 and +339 vs. +240 ms; F(1,104) = 583.71** vs. 676.10** and F(1,72) = 433.31** vs. 633.58**).

5.1.3.4. *To summarize*, the lexicality effect was detrimental to irregular words for accuracy, but favored them for latency and vocal response times. However this effect varied according to the length of the items and to the children's grade level. For pseudowords, whatever the indicator, long items were always penalized. On the other hand, long irregular words were read better than short irregular words, especially by the older children (G3 and G4), whereas they were systematically processed more slowly, according to latency time and vocal response duration. Finally, length had a negative effect on vocal response duration and this effect was more marked for pseudowords, particularly for the youngest children.

5.2. ANOVAs: phonological awareness and rapid serial naming

The number of children was the same as in the preceding analyses (for G1–G4, respectively, 100, 120, 105 and 73), except in G1 for the syllabic and phonemic deletion tests (64 children) and, in G2, for the same tests as well as for the rapid naming tests (52 children), because of differences in

Table 5
Reading level, syllabic and phonemic deletion, rapid naming of colors, orthographic choice (mean and S.D.)

	Grade level	Reading level (in months)	Phonological awareness and short-term memory (STM)				Rapid naming (RAN)		Orthographic choice (trio)	
			Syllable deletion (CVCVCV)	Phoneme deletions (CVC)	Phoneme deletion (CCV)	Phonological STM	Color patches	Reading	Time (ms)	Correct responses
Mean										
	Grade 1	87.41	68.60	84.67	61.58	66.00	57.88	38.42	7613	85.11
	Grade 2	96.96	81.96	87.74	63.70	72.32	46.93	26.73	6131	85.15
	Grade 3	105.30	81.43	93.33	69.84	78.57	43.62	24.92	3747	96.51
	Grade 4	116.10	86.30	94.63	79.22	72.37	40.90	23.63	3111	96.65
Standard deviation										
	Grade 1	7.38	31.37	19.77	26.85	15.34	20.23	16.08	3024	15.46
	Grade 2	11.04	21.63	17.20	26.99	15.75	10.48	5.474	2495	18.03
	Grade 3	16.46	21.81	13.60	28.41	14.19	13.85	4.448	1192	7.657
	Grade 4	18.48	15.41	7.53	15.78	15.52	11.62	3.474	1048	6.596

the tasks used[4]. The means and S.D.s for these tests are presented in Table 5. ANOVAs were carried out only for the tests which evaluate the same skill in different modalities i.e. deletion of a syllable vs. a phoneme and rapid naming of color patches vs. written color names. For the other tests, we examined only the correlation patterns. Regression analyses were also carried out in order to determine the variables which predict reading level.

Two ANOVAs were carried out, with a factor of Grade level (G1–G4) and a factor of Test (three levels for the first test: syllabic deletion, CVC phonemic deletion and CCV phonemic deletion; two levels for the second: naming of color patches and reading of color names). As for the previous results, to facilitate the reading of the text, only significant differences according to the $P < 0.01$ or $P < 0.05$ (two or one stars, respectively) thresholds are reported.

5.2.1. Syllabic and phonemic segmentation tests

An effect of the type of test was predicted: the syllabic deletion test is assumed to be easier than the two phonemic deletion tests, and the phonemic test with CVC syllables is assumed to be easier than that with CCV syllables. The main effects were significant (effect of Grade level, $F(3,326) = 14.25$**; effect of Test, $F(2,652) = 108.09$**), but the interaction was not ($F(6,652) = 1.72$). In fact, the difference between the syllabic deletion task and the two phonemic deletion tasks never reached the significance threshold (in G1–G4, respectively, $F(1,99) = 2.52$; $F(1,51) = 1.92$; the two other Fs >1). On the other hand, as expected, it was less difficult to delete a consonant followed by a vowel than one that was followed by another consonant (+23.1%, +24.0%, +23.5%, +15.4% difference between the mean percentages of correct responses for the CVC minus the CCV task for G1–G4, respectively; $F(1,99) = 66.77$**; $F(1,51) = 38.26$**; $F(1,104) = 72.95$**; $F(1,72) = 70.40$**).

[4] Instead of a rapid color naming test, the other second graders had to perform tests involving rapid naming of letters, digits and objects. Similarly, the other first graders were not exposed to the same items for the syllabic and phonemic deletion tests.

5.2.2. Rapid naming tests

The naming of the color patches was assumed to be more difficult than the reading of color names and the difference between the two tasks was expected to be less for younger children as their written word processing is not yet automatic. The main effects were significant (Effect of Grade level, $F(3,290) = 38.80$**; Effect of Test, $F(1,290) = 469.34$**), but the interaction was not ($F < 1$).

The reading test was more rapid, without clear cut changes between sessions (–19.5, –20.2, –18.7 and –17.3 s difference between mean response times for the naming of written words minus the naming of color patches for G1–G4, $F(1,63) = 48.66$**; $F(1,51) = 142.58$**; $F(1,104) = 258.16$**; $F(1,72) = 175.34$**).

5.2.3. Conclusion

The results of the phonological awareness tests corroborated our hypotheses only partially. On the one hand, even for the youngest children, the results for syllabic deletion were not higher than those found for phonemic deletion. On the other hand, as expected, the test requiring the deletion of a consonant followed by a vowel was easier than that requiring the deletion of a consonant followed by another consonant. For the rapid serial naming tests, the reading test was, as expected, more rapid than the other. However, the difference between the two tasks did not allow us to differentiate the youngest from the oldest children. These results will be re-examined after the analysis of the patterns of correlations between the tests.

5.3. Correlations and reading level predictors

One of the tests of Written Word Processing (the one involving the regular words of LEXORT) was not taken into account because the items partially cover those of LEVORT. The indicated significance threshold is that obtained after the Bonferroni correction, i.e. taking into account the number of correlations tested (190 in G1 and 276 in each of the three other grades). The correlations between the different tests are presented in Tables 6–9, one Table per grade level.

Table 6
First grade. Correlations (** *P* <.01; * *P* <.05 after Bonferroni correction)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Chronological age																			
2. Non-verbal IQ (RAVEN)	-.06																		
3. Verbal IQ (TVAP)	+.07	+.35																	
4. Reading level (alouette)	+.07	+.35	+.22																
5. Phonological STM	+.01	+.25	+.40**	+.28															
6. Syllable deletion	–.09	+.14	+.27	+.32	+.39*														
7. Phoneme deletion (CVC)	+.03	+.15	+.23	+.39*	+.20	+.33													
8. Phoneme deletion (CCV)	+.15	+.27	+.30	+.40**	+.22	+.38*	+.30												
9. RAN (color patches)	+.01	–.17	–.14	–.17	+.08	+.07	–.20	+.05											
10. RAN (written color names)	–.13	+.05	–.30	–.60**	–.15	–.10	–.41	–.24	+.26										
11. Orthog. choice (time)	–.15	+.04	+.00	–.47**	–.19	–.06	–.15	–.15	–.02	+.52**									
12. Orthog. choice (accuracy)	+.04	+.25	+.23	+.49**	+.15	+.29	+.38*	+.16	–.07	–.36	–.24								
13. Accuracy word list 1	–.04	.29	+.27	+.66**	+.24	+.41**	+.52**	+.48**	–.08	–.71**	–.48**	+.54**							
14. Pseudoword list 2	–.04	+.15	+.28	+.64**	+.19	+.45**	+.46**	+.45**	+.07	–.58**	–.32	+.49**	+.79**						
15. Irregular word list 3	+.11	+.34	+.23	+.73**	+.30	+.33	+.37*	+.43**	–.04	–.48**	–.45**	+.38*	+.72**	+.68**					
16. Pseudoword list 4	+.01	+.18	+.31	+.51**	+.22	+.47**	+.51**	+.43**	+.00	–.57**	–.18	+.45**	+.75**	+.76**	+.58**				
17. Latency time word list 1	–.09	–.15	–.06	–.59**	–.04	+.06	–.13	–.04	+.46*	+.58*	+.51**	–.24	–.31	–.16	–.45*	–.13			
18. Pseudoword list 2	–.02	–.07	+.05	–.52**	–.03	+.08	+.10	+.07	+.25	+.26	+.43	–.26	–.13	–.10	–.25	+.05	+.79**		
19. Response time word list 1	–.12	–.21	–.26	–.28	–.37	+.05	–.36	–.21	+.19	+.23	+.32	–.19	–.26	–.01	–.26	–.03	+.16	+.04	
20. Pseudoword list 2	–.16	–.14	–.23	–.37	–.25	+.10	–.39	–.25	+.14	+.33	+.24	+.01	–.34	–.13	–.43	–.22	+.29	+.06	+.63**

N = 100 except for RAN (64) and for time for words and pseudowords (64).
Note. 5: Phonological short-term memory (span); 9 and 10: rapid naming (processing speed); 13–17–19: List 1 (LEVORT); 14–18–20: List 2 (LEXORT); 15–16: Lists 3 and 4 (LEXLENGTH).

Table 7
Second grade. Correlations (** P < .01; * P < .05 after Bonferroni correction)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. Chronological age																							
2. Non-verbal IQ (Raven)	–,02																						
3. Verbal IQ (TVAp)	+.00	+.34*																					
4. Reading level (alouette)	–.05	+.19	+.34*																				
5. Phonological STM	+.07	+.08	+.36*	+.38*																			
6. Syllable deletion	+.05	+.11	+.23	+.30	+.24																		
7. Phoneme deletion (CVC)	+.18	–.02	+.22	+.23	+.18	+.20																	
8. Phoneme deletion (CCV)	–.02	+.50*	+.12	+.31	+.28	+.42	+.26																
9. RAN (color patches)	–.01	–.10	+.09	–.18	–.13	+.07	–.10	–.16															
10. RAN (written color names)	–.26	+.03	+.08	–.22	–.05	+.05	+.02	–.05	+.37														
11. Orthog choice (time)	+.01	+.02	+.22	–.10	+.18	–.06	–.23	–.17	+.34	+.23													
12. Orthog choice (accuracy)	+.06	+.18	+.09	+.30	–.01	+.21	–.04	+.05	–.18	–.41	–.38**												
13. Accuracy word list 1	+.01	+.28	+.28	+.60**	+.37*	+.33	+.36	+.48	–.27	–.22	–.19	+.27											
14. Pseudoword list 2	–.03	+.18	+.18	+.49**	+.21	+.26	+.30	+.41	–.25	–.19	–.23	+.43	+.66**										
15. Irregular word list 3	+.05	+.23	+.25	+.67**	+.33	+.32	+.21	+.38	–.23	–.25	–.24	+.37*	+.69**	+.62**									
16. Pseudoword list 4	+.01	+.20	+.18	+.32	+.23	+.27	+.20	+.43	–.10	–.14	–.26	+.24	+.44**	+.59**	+.44**								
17.Latency time word list 1	+.00	–.20	–.25	–.59**	–.32	–.32	–.08	–.41	+.16	+.22	+.07	–.22	–.47**	–.39**	–.51**	–.31							
18. Pseudoword list 2	–.01	–.16	–.23	–.62**	–.37*	–.25	–.07	–.33	+.07	+.17	–.01	–.08	–.40**	–.39**	–.44**	–.29	+.84**						
19. Irregular word list 3	+.05	–.18	–.30	–.58**	–.33	–.14	–.04	–.28	+.09	+.18	+.08	–.13	–.39**	–.36*	–.46**	–.23	+.81**	+.79**					
20. Pseudoword List 4	+.10	–.09	–.31	–.60**	–.31	–.21	–.09	–.23	–.01	+.27	+.03	–.07	.35*	–.30	–.36*	–.21	+.73**	+.80**	+.81**				
21.Response time word list 1	–.11	+.02	–.07	–.24	–.08	–.11	+.12	+.13	–.04	+.31	–.02	–.01	–.20	–.12	–.11	–.10	+.18	+.08	+.14	+.11			
22. Pseudoword list 2	–.04	–.02	+.00	–.34*	–.08	–.18	–.05	–.01	+.14	+.28	+.19	–.18	–.38**	–.34	–.33	–.22	+.13	+.07	+.08	+.01	+.64**		
23. Irregular word list 3	–.13	+.02	+.09	–.19	–.13	–.05	+.14	+.06	+.05	+.27	+.14	–.06	–.11	–.01	–.13	–.04	+.10	–.02	+.15	+.01	+.54**	+.55**	
24. Pseudoword list 4	–.01	+.08	+.11	–.33	–.01	–.05	–.10	+.04	+.11	+.31	+.26	–.19	–.34*	–.21	–.24	–.07	+.10	+.01	+.10	+.06	+.54**	+.71**	+.55**

N = 120 except for the phonemic deletion tasks and RAN in reading (52).
Note. 5: Phonological short-term memory (span); 9 and 10: rapid naming (processing speed); 13–17–21: List 1 (LEVORT); 14–18–22: List 2 (LEXORT); 15–16–19–20–23–24: Lists 3 and 4 (LEXLENGTH).

Table 8
Third Grade. Correlations (** *P* <.01; * *P* <.05 after Bonferroni correction)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. Chronological age																							
2. Non-verbal IQ (WISC-Khos)	–.13																						
3. Verbal IQ (WISC)	+.09	+.21																					
4. Reading level (Alouette)	–.00	+.21	+.03																				
5. Phonological STM	–.03	+.26	+.14	+.36*																			
6. Syllable deletion (CVCVCV)	+.06	+.17	+.22	+.14	+.23																		
7. Phoneme deletion (CVC)	–.06	+.23	+.20	+.25	+.26	+.24																	
8. Phoneme deletion (CCV)	–.14	+.07	+.10	+.29	+.18	+.21	+.26																
9. RAN (color patches)	+.01	–.01	+.07	–.35	–.17	–.07	–.07	–.04															
10. RAN (written color names)	–.03	–.07	–.01	–.60**	–.19	–.06	–.13	–.13	+.56**														
11. Orthographic choice (time)	+.11	–.24	–.16	–.54**	–.14	–.07	–.06	–.11	+.05	+.25													
12. Orthogr choice (accuracy)	–.02	–.07	–.07	+.14	–.21	–.17	–.12	–.10	–.12	–.13	–.18												
13. Accuracy word list 1	–.19	+.03	+.15	+.31	+.14	+.09	+.02	+.09	+.05	–.26	–.35	+.10											
14. Pseudoword list 2	–.11	+.18	+.14	+.42**	+.16	+.25	+.42**	+.28	–.03	–.24	–.31	–.11	+.38*										
15. Irregular word list 3	–.09	+.21	+.16	+.57**	+.26	+.21	+.23	+.31	–.19	–.37*	–.45**	+.04	+.51**	+.48**									
16. Pseudoword list 4	–.01	+.22	+.13	+.24	+.25	+.26	+.30	+.25	+.03	–.02	–.22	–.10	+.22	+.62**	+.28								
17.Latency time word list 1	+.14	–.11	–.10	–.60**	–.20	+.03	–.21	–.16	+.08	+.45**	+.54**	–.15	–.50**	–.27	–.56**	–.01							
18. Pseudoword list 2	+.12	–.05	–.02	–.59**	–.18	–.03	–.26	–.11	+.17	+.43**	+.43**	–.26	–.37*	–.32	–.43**	–.10	+.82**						
19. Irregular word list 3	+.16	–.07	–.11	–.58**	–.22	+.02	–.21	–.14	+.14	+.47**	+.52**	–.18	–.38*	–.29	–.55**	–.03	+.89**	+.83**					
20. Pseudoword list 4	+.12	–.12	–.06	–.58**	–.23	+.02	–.26	–.05	+.21	+.47**	+.46**	–.20	–.32	–.25	–.41**	–.10	+.79**	+.88**	+.87**				
21.Response time word list 1	+.01	–.03	–.06	–.16	–.10	–.14	–.25	–.13	–.05	+.11	+.03	+.24	–.26	–.25	–.17	–.14	+.24	+.17	+.12	+.10			
22. Pseudoword list 2	–.08	–.04	–.08	–.27	–.16	–.18	–.23	–.15	+.09	+.22	+.03	+.19	–.24	–.29	–.26	–.23	+.22	+.21	+.10	+.10	+.88**		
23. Irregular word list 3	+.06	–.07	–.08	–.31	–.18	–.16	–.18	–.23	+.09	+.20	+.17	+.18	–.36*	–.38*	–.47**	–.21	+.28	+.19	+.19	+.10	+.76**	+.78**	
24. Pseudoword list 4	–.02	+.03	–.03	–.28	–.05	–.13	–.06	–.23	+.07	+.25	+.14	+.10	–.30	–.30	–.45	–.20	+.18	+.12	+.10	+.05	+.58**	+.70**	+.82**

N = 105.
Note. 5: Phonological short-term memory (span); 9 and 10: rapid naming (processing speed); 13–17–21: List 1 (LEVORT); 14–18–22: List 2 (LEXORT); 15–16–19–20–23–24: Lists 3 and 4 (LEXLENGTH).

Table 9
Fourth Grade. Correlations (** *P* <.01; * *P* <.05 after Bonferroni correction)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1. Chronological age																							
2. Non-verbal IQ (RAVEN)	+.15																						
3. Verbal IQ (EVIP)	+.01	+.26																					
4. Reading level (Alouette)	–.11	+.03	+.29																				
5. Phonological STM	–.18	+.02	+.25	+.29																			
6. Syllable deletion (CVCVCV)	–.22	–.10	+.13	+.19	+.17																		
7. Phoneme deletion (CVC)	–.05	+.13	+.11	+.17	+.23	–.05																	
8. Phoneme deletion (CCV)	–.08	+.08	+.03	+.22	+.24	+.16	+.25																
9. RAN (color patches)	–.09	+.01	–.06	–.05	+.02	–.06	–.05	+.16															
10. RAN (written color names)	–.09	–.04	–.11	–.42**	–.01	–.10	–.13	+.05	+.28														
11. Orthographic choice (time)	+.07	–.02	–.11	–.47**	–.24	–.15	–.18	–.19	–.08	+.21													
12. Orthog choice (accuracy)	–.12	–.08	–.25	–.12	–.21	+.12	–.06	–.06	+.08	+.04	–.13												
13. Accuracy word list 1	–.18	+.08	–.02	+.21	+.09	–.03	+.00	+.36	+.04	+.03	–.30	+.02											
14. Pseudoword list 2	+.03	+.04	–.22	+.28	–.05	–.01	+.23	+.24	–.04	–.14	–.23	–.10	+.33										
15. Irregular word list 3	–.13	+.07	+.14	+.51**	+.26	+.15	–.05	+.30	+.11	–.17	–.51**	+.15	+.42*	+.27									
16. Pseudoword list 4	–.04	–.02	+.03	+.23	+.18	+.11	+.25	+.28	–.15	–.03	–.29	–.09	+.20	+.35	+.37								
17. Latency time word list 1	+.11	+.11	–.04	–.49**	–.06	–.30	+.05	–.06	+.02	+.39	+.61**	–.11	–.24	–.24	–.42**	–.06							
18. Pseudoword list 2	+.03	+.07	–.05	–.56**	–.08	–.40	–.04	–.11	–.04	+.36	+.44**	+.07	–.21	–.32	–.29	–.08	+.77**						
19. Irregular word list 3	+.10	+.05	–.09	–.60**	–.13	–.27	–.03	–.14	–.05	+.27	+.71**	–.06	–.37	–.35	–.56**	–.13	+.81**	+.80**					
20. Pseudoword list 4	+.02	+.09	–.06	–.53**	–.12	–.41	–.06	–.14	–.13	+.31	+.61**	–.07	–.27	–.32	–.39	–.04	+.72**	+.84**	+.87**				
21.Response time word list 1	–.21	+.15	+.06	–.05	–.14	+.21	+.09	+.06	–.00	+.12	+.18	+.01	–.04	–.03	–.03	–.18	+.16	+.10	+.09	+.04			
22. Pseudoword list 2	–.16	+.07	–.03	–.12	–.10	+.18	–.17	–.09	+.01	+.16	+.13	+.13	–.07	–.04	–.12	–.20	+.11	+.15	+.02	+.01	+.75**		
23. Irregular word list 3	–.06	+.07	+.05	–.06	–.07	+.24	–.00	+.09	+.03	+.12	+.15	+.05	–.07	–.17	–.09	–.09	+.12	–.01	–.04	–.09	+.72**	+.75**	
24. Pseudoword list 4	–.15	–.13	+.03	+.05	+.05	+.32	–.04	+.05	+.01	+.04	–.04	+.20	+.02	–.27	–.26	+.04	–.01	–.05	–.20	–.21	+.49**	+.67**	+0.67**

N = 73.
Note. 5: Phonological short-term memory (span); 9 and 10: rapid naming (processing speed); 13–17–21: List 1 (LEVORT); 14–18–22: List 2 (LEXORT); 15–16–19–20–23–24: Lists 3 and 4 (LEXLENGTH).

5.3.1. Correlations between reading tests

For the different tests focused on written word processing (WWP), we reported the correlations within a specific measurement (accuracy or latency time or vocal response duration) between the different tests (LEVORT words, LEXORT pseudowords, LEXLENGTH irregular words and LEXLENGTH pseudowords) and the correlations within a specific test (LEVORT words or LEXORT pseudowords or LEXLENGTH irregular words or LEXLENGTH pseudowords) between the different measurements (accuracy and latency time on one hand, latency time and response duration on the other).

5.3.1.1. WWP: correlations within a specific measurement. The correlations between the different tests were significant for latency times and for vocal response durations (two out of two in G1 and six out of six for each of the other three grade levels). For accuracy scores, the number of significant correlations decreased from G1 to G4. They were all significant in G1 and G2 but only four out of six were in G3, the non-significant correlations being those between the LEXLENGTH pseudowords and the LEVORT or LEXLENGTH words. In G4, only one significant correlation was found: between the LEVORT and LEXLENGTH words.

5.3.1.2. WWP: correlations between the different measures. The correlations accuracy and latency times were not significant in G1, three out of four were significant in G2, two out of four in G3, and only one in G4, the strongest correlations being those for irregular words. Finally, the correlations between latency time and vocal response duration were never significant, which suggests that the two measurements capture relatively independent phenomena. It should be noted that the correlations between vocal response duration and any of the other variables were only rarely significant: seven out of 1018. Vocal response duration will therefore not be a part of subsequent examinations of correlation patterns.

5.3.1.3. WWP and rapid naming. All the correlations between the rapid reading of color names test and WWP tests were significant in G1, at least for accuracy, and in G3, at least for response time whereas no significant correlations were found in G2 and G4. Similarly, significant correlations were found between the orthographic choice test and WWP, according to accuracy for the youngest children (G1) and according to processing times for the oldest children (G3 and G4).

5.3.1.4. WWP and Alouette. All 14 correlations with response latency time were significant while the number of significant correlations with accuracy decreased from G1 to G4 (four out of four then 3, 2 and 1 in G1–G4, respectively), the strongest correlations being those which included irregular word reading. In all grade levels except G2, the correlations between "l'Alouette" and the orthographic choice test were significant, for accuracy and response time for the youngest children (G1) and only for response time for the oldest children (G3 and G4). Likewise, except in G2, "l'Alouette" was correlated significantly with the rapid reading of color names test.

5.3.2. Correlations between reading and other variables

The correlations between the three phonological awareness tests and the reading tests were only significant for G1 children: 11 out of 12 with accuracy for the WWP tests and two out of three with "l'Alouette", the two non-significant correlations being for syllable deletion tests. There were very few significant correlations between the reading tests and the phonological short-term memory test (four out of 46), with no clear tendency standing out (two with WWP variables in G2; two with "l'Alouette": one in G2 and one in G3). Correlations with rapid naming of color patches were quite rare (two out of 46, one in G1 with latency times for the words of LEVORT and one in G3 with the rapid naming of written color names). Finally, out of 46 possible correlations between chronological age or non-verbal cognitive level and the other variables, none reached the significance threshold and only one correlation was found between reading and verbal cognitive level (with "l'Alouette" in G2).

5.3.3. Reading level predictors

The predictors considered were those that are the most strongly correlated with the reading level assessed by the test of "l'Alouette". We have, however, systematically integrated the scores for the two irregular word reading tests (LEVORT and LEXLENGTH) and for the two pseudoword reading tests (LEXORT and LEXLENGTH), for accuracy only for G1 (because response times were available only for some of the children) and for accuracy and response time for G2–G4. We used the forward entry method combined with the backward removal method. The results are presented in Table 10.

5.3.3.1. Grade 1. Nine variables were retained for the level G1: the scores for the two phonemic deletion tests, response

Table 10
Prediction of reading level (Alouette)

	R	R-two adjusted	R	R-two adjusted	R	R-two adjusted	R	R-two adjusted
	First grade		Second grade		Third grade		Fourth grade	
Backward removal method	0.793	**0.603**[a1]	0.789	**0.612**[b1]	0.797	**0.617**[c1]	0.699	**0.467**[d1]
Forward entry Method	0.762	**0.567**[a2]	0.781	**0.603**[b2]	0.797	**0.617**[c2]	0.699	**0.467**[d2]

Grade 1: a1. Irregular words, accuracy (LEVORT); pseudowords, accuracy (LEXORT); pseudowords, accuracy; (LEXLENGTH); rapid naming in reading (time); a2. Irregular words, accuracy (LEVORT); pseudowords, accuracy (LEXORT); Grade 2, b1. Irregular words, accuracy (LEXLENGTH); pseudowords, latency time (LEXORT); pseudowords, latency time (LEXLENGTH); b2. Irregular words, accuracy (LEXLENGTH); pseudowords, latency time (LEXLENGTH); Grade 3, c1. Irregular words, accuracy (LEXLENGTH); pseudowords, latency time (LEXORT); rapid naming in reading; phonological short-term memory; orthographic choice, time; c2. idem c1; Grade 4, d1. Irregular words, accuracy (LEXLENGTH); pseudowords, latency time (LEXORT); rapid naming in reading; d2. idem d1.

time for the rapid naming task involving the reading of color names, response accuracy and response time for the orthographic choice test, accuracy for the LEVORT plus the LEXLENGTH irregular words and for the LEXORT plus the LEXLENGTH pseudowords.

The backward removal method kept four of these variables which explained 60% of the variance in the test of "l'Alouette" (accuracy for LEVORT irregular words and for LEXORT plus LEXLENGTH pseudowords, response time for the rapid naming in reading). The forward entry method selected only two (accuracy for LEVORT irregular words and for LEXORT pseudowords) which alone accounted for 57% of the variance.

5.3.3.2. Grade 2. Apart from the eight variables for irregular word and pseudoword reading (accuracy and latency time for the LEVORT plus LEXLENGTH irregular words and for the LEXORT plus LEXLENGTH pseudowords), two other variables were significantly correlated to "L'Alouette" in G2: vocabulary and short-term memory.

The backward removal method retained only accuracy for LEXLENGTH irregular words and latency time for the LEXORT and LEXLENGTH pseudowords, which accounted for more than 61% of the variance, compared to 60% for the forward entry method which kept only accuracy for LEXLENGTH irregular words and latency time for the LEXLENGTH pseudowords.

5.3.3.3. Grade 3. Eleven variables were introduced for the level G3: the eight irregular word and pseudoword scores, time for the orthographic choice test and for the test of rapid naming in reading, and short-term memory scores.

The same five variables, which accounted for almost 62% of the variance, were selected by both methods: accuracy in irregular word reading (LEXLENGTH), latency time in pseudoword reading (LEXORT), phonological short-term memory scores and response time for orthographic choice and for rapid naming in reading.

5.3.3.4. Grade 4. Because the number of children was reduced in G4, only seven variables were entered in the model: response time for rapid naming in reading, for orthographic choice and latency time for the four irregular word and pseudoword reading tests, plus accuracy for the LEXLENGTH irregular words.

Both models kept three variables, which explained almost 47% of the variance in the test of "l'Alouette": accuracy for the LEXLENGTH irregular words and latency time for the LEXORT pseudowords, plus response time for rapid naming in reading.

5.3.3.5. To summarize, whatever the grade level, the reading level of the children assessed by "l'Alouette" is explained mainly by accuracy for irregular word reading and by latency time for pseudoword reading, at least when this indicator can be used (from the end of G2). Another test emerges relatively systematically: that involving the rapid naming of written color names (in G1, G3 and G4). No other clear tendency stands out, even when certain abilities assumed to be linked to reading (phonemic awareness in G1, phonological memory in G2 and G3) or other skills (vocabulary level in G2) can be taken into account. The only exception is phonological memory which explains a part of the additional variance in G3.

6. General discussion

6.1. Relationships between reading and associated skills

The correlation analyses suggest that there is a great amount of independence between non-verbal or verbal cognitive skills and reading level. This reproduces earlier results (Siegel, 1989; Stanovich, 1992). The absence of significant correlations with chronological age may be explained by the fact that only children who were not held back in school were integrated into the study. However, within each grade level, the youngest children were separated from the oldest children by 11 months. This gap could have had an effect on the youngest children's performances. This was not the case; even in Grade 1, reading level was never correlated with chronological age.

Correlations between phonological awareness tests and reading tests were significant only for the youngest children (those of G1), especially for the phonemic tests. However, these skills added no significant part to the variance explained in reading, at least according to the test that was used. This result conforms to what is found in the literature, particularly the data from studies carried out in languages with a less opaque orthography than that of English (in Finnish or in German, for instance, e.g.: Holopainen et al., 2001; Landerl and Wimmer, 2000; Mann and Wimmer, 2002). The ANOVA brought out a surprising result. The children generally are supposed to be more successful with syllabic segmentation tests than with phonemic segmentation tests (Anthony et al., 2003; Liberman et al., 1974; and for results in French, Courcy et al., 2000). No significant difference between syllabic and phonemic awareness was found in our study, which might be due to the fact that the phonological awareness tasks were pre-recorded, thus eliminating the possibility of relying on lip movements which may facilitate the recognition of syllable borders (see Colé and Sprenger-Charolles, 1999).

The correlations between reading and phonological short-term memory were few and no clear tendency stands out, except in G3 in which phonological memory explained an additional part of the variance in reading. These results are along the same lines as the data in the literature which indicate that the weight of short-term memory as predictor of future reading level is not strong (Bowers, 1995; Lecocq, 1991; Parrila et al., 2004; Wagner et al., 1994, 1997; Wimmer, 1993).

The correlations between reading tests and test of rapid naming of color patches were rarely significant. This result is consistent with the data indicating that the naming speed does not seem to be the best predictor of reading (Wagner et al., 1994, 1997). However, other studies suggest that this is the case (Ackerman and Dykman, 1993; Bowers and Swanson, 1991; Felton and Brown, 1990; Parrila et al., 2004). As we have already indicated, these inconsistencies may be explained by differences in the tasks used. For instance, the predictive power of letter or digit naming abilities was found to vanish after controlling for the reading level (Wagner et al., 1997), which was not the case for the rapid naming of colors (Parrila et al., 2004). These differences might be due to the fact that the former abilities are more dependent on the reading level than the latter. This interpretation does not allow us to explain the results of the present study in which a task involving color names was used. Another explanation is therefore necessary. It is possible that the inconsistencies between the studies may be explained by the reading level of the children being evaluated: correlations between rapid naming and reading level are only significant for poor readers (McBride-Chang and Manis, 1996; Bowers, 1995). This result could be explained by the fact that the naming speed is linked to speed of articulation which increases with age and attains its maximum level rapidly for average readers but more slowly for children with specific language impairments (Henry and Millar, 1993; Kail and Park, 1994). Consequently, for subjects without language impairments, individual differences are neutralized rapidly. This also permits us to understand why, in our study, the duration of vocal response production (cf. the WWP results) was only exceptionally correlated with other reading tests. Similar results were reported by Parrila et al. (2004) concerning the fact that the correlations between articulation rate and different reading variables were fairly small.

Finally, out of the four possible correlations between the two rapid naming tests (color patches and written color names), only one was significant. This result indicates that these two tests capture relatively independent skills. Consequently, if in an assessment of dyslexia, a processing time deficit is found in reading, though not in non-reading, we may be led to believe that the speed deficit is specifically linked to reading.

6.2. Relationships between the different reading tests

The correlations between “l’Alouette” and the different tests focused on WWP were strong, the strongest being those involving irregular word reading. In addition, the regression analyses indicated that, whatever the grade level of the children, their reading levels assessed by the “Alouette” were explained principally by accuracy in irregular word reading and by latency time for pseudoword reading, at least when the indicator could be used, i.e. starting in the second grade. As we have already emphasized, this indicates that “l’Alouette” does not simply evaluate the quality of the sublexical procedure.

Concerning the correlations between irregular words and pseudowords, they were systematically significant for vocal production duration and for latency time, whatever the grade level of the children. For accuracy, they were significant only for the first two grade levels. These results suggest that, for the youngest children, these different types of items are processed at least partially by the same procedure. Less than half of the correlations between accuracy and latency time for items of the same category were significant, and none for the youngest children. There seem to be therefore, at least for beginning readers, an inverse relationship between accuracy and time, some children privileging accuracy to the detriment of time, and others adopting the opposite strategy (see also Olson et al., 1994). Finally, the correlations between latency time and vocal response duration were never significant, which suggests that these two measurements capture relatively independent phenomena. This result may also be due to the fact that, as noted earlier, articulation speed increases with age and, for children without language impairments, rapidly reaches its maximum level. However, insofar it has been shown that this is not the case for children with language impairments (Henry and Millar, 1993; Kail and Park, 1994), it is important to take this indicator into account in an assessment of dyslexia.

6.3. Reading procedures

The LEVORT test highlighted the effect of sublexical factors on word reading, which indicated that these items, though frequent, were not processed solely by the lexical procedure. However, these effects varied as a function of grade level: the sublexical factors, taken together, had a stronger effect on the youngest children’s performances, whatever the measurement. The effect of regularity was, however, always significant. These results reproduce those obtained in earlier studies with French-speaking children. In particular, in a longitudinal study, Mousty and Leybaert (1999) found a regularity effect for response accuracy as well as for response time in the second and third grade. This was also the case in another longitudinal study (from the middle of G1 to the end of G4), except for latency time from the third grade (for a discussion, see Sprenger-Charolles et al., 2003). These results suggest that the regularity effect is strong in French.

The effect of graphemic complexity (comparison between words containing a contextual grapheme and those with only simple graphemes) was only observed for accuracy, not for latency time, and decreased with grade level. In fact, it only had a negative effect on the performances of children in the first three grade levels. This effect, therefore, seems less robust than that of regularity (cf. also Mousty and Leybaert, 1999), which may be explained by the fact that contextual graphemes follow contextual regularities: ″c″ and ″g″ are pronounced, /s/, /Z/ or /k/, /g/ according to the following vowel.

The effect of graphemic length was never detrimental to items containing a digraph. We found the expected facilitating effect on the duration of the vocal production at all grade levels. The presence of a digraph also had a positive effect on accuracy for all grade levels taken together, as well as on the latency time in the first three grade levels. These results reproduce those obtained for French-speaking beginning readers (Sprenger-Charolles, 1994; Sprenger-Charolles et al., 1998b). Those found for children in the fourth grade for latency time are in line with those of French-speaking adults, at least for frequent words (cf. Rey et al., 1998). This may be explained by the fact that, if the basic unit of the sublexical procedure is the grapheme rather than the letter, on the one hand, readers have fewer units to decode and then assemble when the items contain a digraph than when they are composed only of simple graphemes. On the other hand, they also have fewer phonemic units to program in order to provide an oral response. Thus, the presence of a digraph may have a facilitative effect on reading overall when the reader strongly relies on grapheme-phoneme correspondences and when this processing is not yet automatized, that is for the youngest children.

For the vocal production durations, two unexpected results were found. Irregular words were pronounced more rapidly than regular words (in G4), as were words containing contextual graphemes compared to words containing only simple graphemes (in G3 and G4). These results may be due to the fact that, in view of our hypotheses on the effect of graphemic length, we have matched up the items in the first two categories according to all of the factors which can bias the results, in particular the number of consonant clusters. These controls were less strict for the two other categories of items. In fact, there was no consonant cluster in the irregular words, and only two in the items containing a contextual grapheme, compared to 5 and 4 for the items containing only simple graphemes and for those containing a digraph (for an analysis of the effect of the presence of consonant clusters on response accuracy in reading, see among others Sprenger-Charolles and Siegel, 1997).

For the LEXORT and LEXLENGTH tests, the most remarkable result was that the lexicality effect differed according to the type of items (regular words or irregular words vs. pseudowords) and the measurement (response accuracy or latency time). On one hand, for accuracy, according to LEXORT, regular words were read better than pseudowords while, according to LEXLENGTH, pseudowords were read better than irregular words. For latency time, as for vocal response duration, the lexicality effect was always detrimental to pseudowords. However, the effect of the sublexical factors manipulated was not the same on words and pseudowords, suggesting that these two types of items were not processed the same way.

In the LEXLENGTH test, the lexicality effect varied according to item length and grade level. For pseudowords, no matter what the indicator, long items were always penalized. On the other hand, long irregular words were read better than short ones by the oldest children (G3 and G4), but more slowly, according to both latency time and vocal response duration. Length having a positive effect on accuracy in irregular word reading can be understood if we admit that this is primed by regularity. Indeed, irregular words are never totally irregular, the portion of regularity being stronger for longer words ("sculpture" /skyltyr/) than for short ones ("sept" /sɛt/). The fact that length always had a negative effect on latency time, though less marked for irregular words, and especially for the oldest children, might be explained by the fact that this time captures a part of the pre-programming of the articulatory codes (for a discussion, see Marmurek and Rinaldo, 1992; Kessler et al., 2002). These codes exist for words and not for pseudowords (unless they are very similar to words), and are better encapsulated for the oldest children (Henry and Millar, 1993; Kail and Park, 1994).

In the LEXORT test, whatever the grade level, orthography did not have the same effect on regular words and on pseudowords. Graphemic length, whatever the type of item and measurement, never had a negative effect. The presence of a digraph even had a systematically positive effect on words and pseudowords, except for pseudoword latency time, pseudowords containing a digraph being processed no more rapidly than those containing only simple graphemes. These results suggest that children process graphemes rather than letters. In this case, as we have already emphasized, there are indeed fewer units to decode, then assemble, when the items contain a digraph than when they are composed only of simple graphemes; there are also fewer phonemic units to program in order to provide an oral response. Graphemic length having a facilitator effect on latency times for words, and not for pseudowords, may be explained by the fact that lexical information can help in the case of words, but not for pseudowords. Conversely to the effect of graphemic length, the impact of graphemic complexity was negative. For accuracy, this effect was more strongly marked for pseudowords, and, for latency time, it was significant only for pseudowords. It can be understood that this negative effect is especially detrimental to pseudowords if we admit that there is no lexical recuperation possible for these items.

Taken together, these results indicate that the two reading procedures do not function independently. Above all, all of the sublexical factors manipulated (length of items or of graphemes, regularity of grapho-phonemic correspondences and graphemic complexity) virtually always have an effect on reading which can be observed in the accuracy of the response and/or latency time, or even vocal production duration. When this effect is positive, its impact seems stronger for words (the effect of graphemic length); conversely, when it is negative (item length, graphemic complexity), it seems more noticeable for pseudowords.

6.4. One example of an assessment of dyslexic children relying on EVALEC

If dyslexics suffer mainly from phonological deficits (Casalis, 2003; Ramus et al., 2003; Snowling, 2000), we should expected weaker performances for them as compared to average readers in all the tasks which imply phonological processing. Thus, in reading-related tasks, the dyslexics' diffi-

Table 11
Mean scores, scores in the norms (less than ±1 S.D. from the mean) and scores 1 or 1.65 S.D. [a] below (for accuracy scores) or above (for processing times) the mean for reading-related skills (phonological awareness, phonological short-term memory, rapid naming) and for the two silent reading tasks (rapid naming of color names and orthographic choice)

	Syllable deletion (CVCVCV items) accuracy (percentage)				Phoneme deletion (CVC items) accuracy (percentage)				Phoneme deletion (CCV items) accuracy (percentage)				Phonological short-term memory (Span) (maximum = 6)			
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
+1 S.D.	100	100	100	100	100	100	100	100	88.4	90.7	98.3	95.0	4.9	5.3	5.6	5.3
Mean	68.6	82.0	81.4	86.3	84.7	87.7	93.3	94.6	61.6	63.7	69.8	79.2	4.0	4.3	4.7	4.3
–1 S.D.	37.2	60.3	59.6	70.9	64.9	70.5	79.7	87.1	34.7	36.7	41.4	63.4	3.0	3.4	3.9	3.4
–1.65 S.D.	16.8	46.3	45.4	60.9	52.0	59.4	70.9	82.2	17.3	19.2	23.0	53.2	2.4	2.8	3.3	2.8
	RAN (color patches) response time (s)				RAN (written color names) response time (s)				Orthographic choice response time (ms)				Orthographic choice accuracy (%age)			
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
+1 S.D.	37.6	36.5	29.8	29.3	22.2	21.3	20.5	20.2	4589	3635	2555	2063	100	100	100	100
Mean	57.9	46.9	43.6	40.9	38.4	26.7	24.9	23.6	7613	6131	3747	3111	85.1	85.2	96.5	96.7
–1 S.D.	78.1	57.4	57.4	52.5	54.6	32.2	29.4	27.1	10637	8626	4939	4159	69.7	67.1	88.9	90.1
–1.65 S.D.	91.3	64.2	66.5	60.1	65.2	35.8	32.3	29.4	12603	10248	5713	4840	59.6	55.4	83.9	85.8

[a] The threshold of 1.65 S.D. corresponds to the lowest 5% of the distribution.

culties should show up in the phonemic awareness tasks, particularly for the most difficult task, that involving the deletion of a consonant followed by another consonant. Phonological short-term memory and rapid naming deficits should also show up, especially for the latter task when it involves reading. For the reading tests, a deficit in pseudoword reading should emerge, even in the comparison with average readers of the same reading level. In addition, the sublexical factors are assumed to have a more detrimental effect on pseudoword reading than on word reading, especially for dyslexics. Thus, for example, a strong detrimental effect of item length should be found when dyslexics have to read pseudowords because their reading deficit is assumed to be particularly noticeable for the reading of long pseudowords.

These hypotheses were partially corroborated by a study in which all the children who attended a special school for dyslexics (mean chronological age in months, 110.7, S.D. 8.1) and who could be matched to 7 year-old controls of the same reading level were enrolled (Sprenger-Charolles et al., under revision). In that study, to state that a dyslexic was only suffering from a developmental delay, his/her scores had to be less than ± 1 S.D. from the mean of the same reading level controls and to state that he/she had a developmental deviant trajectory, his/her scores had to be 1.65 S.D. below (for accuracy scores) or above (for processing times) the mean of the same controls, the threshold of 1.65 S.D. being chosen because it corresponds to the lowest 5% of the distribution, thus to a score assumed to characterize dyslexics' performances. The normative data from EVALEC which we used (those from the first graders) are presented in Tables 11 and 12. In these Tables, the data for second, third and fourth graders are also presented in order to help users to rapidly determine whether the performance of a child is in the norms of the same chronological age and/or of the same reading level controls, or out of these norms.

As regards these normative data, in Sprenger-Charolles et al.'s study (under revision), the scores obtained by most of the dyslexics in the phonological tasks that did not involve reading (phonological awareness, phonological short-term memory and rapid naming) were similar to those of the reading level controls, thus suggesting that, as regards these skills, most of them were only suffering from a developmental delay. This result may be explained by the intensive phonological training to which these dyslexics were submitted. This training does not seem, however, to have a strong effect on their phonological reading abilities since, according to the most difficult phonological reading task—the one involving long pseudowords—all these dyslexics showed a deviant profile as their scores were found to lag more than 1.65 S.D. behind those of the same reading level controls, for most of them according to accuracy and, for the others, according to latency time. These results once again point out that poor phonological skills, and especially poor phonological reading skills, are the main signature of developmental dyslexia (see for a review, Sprenger-Charolles and Colé, 2003) and that it is difficult to compensate for deficits in the domain of phonology (Vellutino et al., 1996). They also highlight the importance of taking into account both accuracy and processing time in the assessment of reading disabilities, all the more so because the results obtained for accuracy scores reach a ceiling level very soon (from the end of third grade; see Tables 11 and 12, and especially the results for the words and the pseudowords from LEXORT). Therefore, it is difficult to assess the efficiency of a child's reading skills based only on his/her accuracy scores, at least for older children. For that reason, the fact that latency times were taken into account in EVALEC is an important step forward.

In conclusion, EVALEC should allow researchers and speech therapists to accurately assess dyslexics' reading and related skills and consequently to establish specific remediation programs according to the skills that are the most severely

Table 12
Mean scores, scores in the norms (less than ± 1 S.D. from the mean) and scores 1 or 1.65 S.D. [a] below (for accuracy scores) or above (for processing times) the mean for the reading aloud tasks (accuracy scores and latency times)

	LEVORT all words (R1–R3, R4) [b]				LEVORT irregular words (R4) [b]				LEXORT regular words (R1–R3) [b]				LEXORT pseudowords (R1–R3) [b]			
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
+1 S.D.	93.3	97.4	100	100	84.5	96.2	100	100	97.9	99.3	100	100	84.1	89.5	100	96.1
Mean	79.1	90.1	97.8	96.7	57.7	80.4	95.4	94.2	86.2	93.3	98.6	97.6	66.7	75.8	92.9	87.6
–1 S.D.	64.8	82.8	94.5	93.1	30.8	64.6	86.8	87.0	74.5	87.4	95.8	93.9	49.3	62.1	84.9	79.0
–1.65 S.D.	55.6	78.0	92.3	90.7	13.4	54.3	81.3	82.4	66.9	83.5	93.9	91.5	37.9	53.1	79.6	73.4
	LEXLENGTH: short irregular words				LEXLENGTH: long irregular words				LEXLENGTH: short pseudowords				LEXLENGTH: long pseudowords			
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
+1 ET	54.4	74.5	92.4	96.3	60.4	83.5	100	99.9	95.0	97.2	100	100	86.1	84.2	100	95.1
Mean	33.4	57.4	76.7	81.8	33.2	60.3	89.0	86.6	75.9	82.3	94.7	90.4	64.9	65.6	86.0	79.2
–1 ET	12.4	40.3	60.9	67.2	6.0	37.0	74.7	73.2	56.8	67.3	85.4	80.8	43.7	47.0	70.0	63.3
–1.65 ET	0	29.2	50.7	57.8	0	21.9	65.3	64.5	44.3	57.5	79.3	74.5	29.8	34.9	59.6	53.0
Latency times (ms)																
	LEVORT all words (R1–R3, R4) [b]				LEVORT irregular words (R4) [b]				LEXORT regular words (R1–R3) [b]				LEXORT pseudowords (R1–R3) [b]			
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
+1 S.D.	818	605	556	583	794	609	533	587	807	594	549	578	989	830	758	709
Mean	1351	1047	848	796	1519	1099	934	831	1311	1033	824	785	1684	1464	1304	1203
–1 S.D.	1884	1488	1140	1009	2245	1590	1335	1075	1814	1472	1099	992	2380	2098	1850	1697
–1.65 S.D.	2230	1775	1330	1147	2717	1908	1596	1233	2141	1757	1278	1127	2832	2511	2205	2018
	LEXLENGTH: short irregular words				LEXLENGTH: long irregular words				LEXLENGTH: short pseudowords				LEXLENGTH: long pseudowords			
	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4	G1	G2	G3	G4
+1 S.D.		802	645	676		656	580	614		825	751	755		864	839	855
Mean		1238	1115	964		1678	1365	1076		1354	1241	1148		1816	1772	1581
–1 S.D.		1673	1585	1253		2701	2151	1538		1883	1732	1542		2767	2705	2308
–1.65 S.D.		1956	1890	1441		3365	2661	1838		2227	2051	1798		3386	3311	2780

[a] The threshold of 1.65 S.D. corresponds to the lowest 5% of the distribution.
[b] R1: one letter for one phoneme, R2: items with a digraph (ch, ou...), R3: items with a contextual grapheme (c–g, a + i + ll), R4: irregular items.

impaired and the compensatory strategies that may have been developed[5].

Notes

We gratefully thank two anonymous reviewers for their very helpful comments on an earlier version of this paper.

Appendix A. Tests focused on WWP

A1. LEVORT (words of four levels of regularity) and LEXORT (words of the first three levels of regularity of LEVORT and matched pseudowords)

-R1: each letter corresponds to one phoneme.

-R2: items containing a digraph (ch, ou, on-an-in).

-R3: items containing a contextual grapheme (c–g, a + i + ll).

-R4: items containing at least one grapheme with an exceptional pronunciation (x, sc).

(In each sublist, for the initial grapheme: two vowels, four plosives of which one being voiced, four fricatives, two nasals).

	Length in number of				Frequency	
	Letters	Phonemes	Syllables	Consonant clusters	Bigrams	Lexical
Words						
R1	5.58	4.83	2.00	5	1121.60	25.50
R2	5.42	4.00	1.67	4	1180.06	29.36
R3	5.83	4.67	2.00	2	1118.60	28.25
R4	5.75	4.25	1.58	0	861.98	32.00
Mean R1–R4	5.65	4.44	1.81	3.67	1070.56	28.78
Mean R1–R3	5.61	4.50	1.89	3.67	1140.08	27.7
Pseudowords						
R1	5.58	4.67	2.08	5	1061.68	
R2	5.58	3.92	1.50	3	1221.81	
R3	5.67	4.50	1.83	3	1050.00	
Mean R1–R3	5.61	4.36	1.81	3.67	1111.16	

[5] Contact the first author to obtain specific information concerning when and where this battery of tests will be available.

List of items for LEVORT and LEXORT (order of presentation in parentheses)

Words (Practice: porte, école, pile)				Pseudowords (practice: truche, lapre)		
R1	R2	R3	R4	R1	R2	R3
avril (1)	écharpe (15)	acide (14)	aiguille (12)	atrul (7)	énoure (8)	égibe (23)
abri (29)	oncle (5)	agité (33)	album (41)	adrile (2)	onfre (34)	écine (3)
pilote (2)	poudre (17)	paille (34)	pied (8)	pidre (35)	pirche (17)	pogide (27)
plume (22)	tante (23)	caillou (32)	piscine (43)	pitode (26)	cande (5)	traillou (30)
tulipe (4)	tache (19)	cuisine (38)	compte (44)	tople (4)	durche (6)	cuifle (24)
tomate (42)	danse (20)	guide (47)	deuxième (39)	tanepi (29)	toupre (10)	guive (13)
marmite (7)	montre (21)	magie (6)	maximum (46)	matore (14)	moube (1)	mocile (15)
minute (10)	malin (24)	merci (30)	monsieur (36)	mopade (19)	nurin (21)	nacide (9)
farine (11)	fourmi (25)	fragile (18)	femme (48)	furpe (12)	frante (16)	fogir (33)
fable (45)	cheval (27)	facile (40)	scie (26)	fudre (18)	chaful (32)	saille (25)
sable (13)	soupe (28)	cinéma (9)	six (16)	simade (28)	sintar (20)	cipre (22)
samedi (35)	sapin (31)	citron (3)	sept (37)	sinope (11)	supon (31)	cirate (36)

A2. LEXLENGTH. Ten short and 10 long items for each sub-list (words and pseudowords)

	Length in number of			Frequency	
	Letters	Phonemes	Syllables	Bigrams	Lexical
Words					
Short	4.30	3.10	1.30	713.45	18.90
Long	8.00	6.00	2.40	988.72	10.20
Short and long	6.15	4.55	1.85	851.08	14.55
Pseudowords					
Short	4.40	3.10	1.20	768.64	
Long	7.80	6.10	2.50	1053.96	
Short and long	6.10	4.60	1.85	911.30	

List of items for LEXLENGTH (order of presentation in parentheses)

Irregular words		Pseudowords	
Practice: orgueil, ailleurs	Test	Practice: aspouche, tarmine	Test
Short	dix (1)	Short	dul (1)
	août (3)		oume (2)
	oeil (4)		oude (5)
	faon (5)		funve (8)
	pays (8)		plou (10)
	écho (11)		opha (12)
	ennui (12)		altin (14)
	poêle (13)		poibe (16)
	clown (15)		conde (18)
	short (20)		chile (20)
Long	automne (2)	Long	opaurir (15)
	baptême (6)		bartome (9)
	seconde (14)		silième (6)
	condamné (10)		couciron (4)
	septième (18)		sartopin (7)
	accident (16)		évaloupe (11)
	aquarium (7)		acribion (13)
	technique (19)		tainouque (19)
	sculpture (9)		scaltoure (17)
	orchestre (17)		orphade (3)

A3. TRIO. Orthographic choice between a correct item, a visual foil and a phonological foil

(order of presentation in parentheses),
auto-outo-oto (8), vélo-véla-vélau (9), train-troin-trin (7),
loup-louq-lou (1), pigeon-pigean-pijon (3),
pomme-pomne-pome (6), fraise-froise-frèze (5),
rouge-rouqe-rouje (2), blanc-blauc-blan (4).

NB. Within each series, the order of the presentation of the items is randomized.

Training: tulipe, rose, chaise,

Appendix B

B1. Syllabic and phonemic deletion

Syllabic deletion: First syllable (pseudowords with a consonant–vowel structure: CVCVCV, order of presentation in parentheses).

-First syllable starting with a plosive:

povidu (1), buliva (7), tokali (2), tipango (3), banidé (5), kossila (6),

-First syllable starting with a fricative or a liquid:

zofitu (5), retouda (8), valoté (9), soguté (10),

-Training: pajomi, cobuna, parotu.

Phonemic deletion: First phoneme (pseudowords with a consonant-vowel–consonant structure, order of presentation in parentheses).

-Syllable starting with a plosive: puf (1), bir (6), tal (7), dour (5), kip (4), gof (8),

-Syllable starting with a fricative: fek (9), vaf (12), sat (10), zil (3), chol (11), jor (2),

-Training: fur, voul, tof.

Phonemic deletion: First phoneme (pseudowords with a consonant–consonant–vowel structure, order of presentation in parentheses).

-Plosive followed by a liquid: klo (1), blo (5), pra (2), grou (12);

-Fricative followed by a liquid: flin (8), sla (9), vri (10), sri (3);

-Plosive followed or preceded by a fricative: tsé (4), sti (6), pso (7), spa (11)

-Training: tru, gron, blo.

B2. Phonological short-term memory: repetition of three to six syllable pseudowords (between parentheses: order of presentation)

	Items with only consonant–vowel syllables			Items containing a consonant–vowel–consonant syllable		
3S	moukola (1)	favéli (3)	varéla (4)	bartino (2)	linourac (5)	Chadurlé (6)
4S	gontadiro (7)	nuronlado (9)	todonkino (12)	rikalpéta (8)	sazidulor (10)	Farvikéru (11)
5S	tabaritolu (13)	Munolivoura (7)	takorétidou (18)	mandurlanoti (14)	rutadilérac (15)	tirsatabito (16)
6S	pédonuratilé (22)	vafitaludéro (23)	pubagoritélu (24)	vardotivaruté (19)	toziltéfavilo (20)	munigamessotir (21)

Appendix C

Automatic detection of latency time and duration of vocal response production.

Program created by Elisabeth Samain* in collaboration with René Carré**.

*CNRS-UPS-2259, Villejuif; **CNRS-UMR-820, TSI (Traitement du Signal et des Images), ENS des Télécommunications, Paris.

Detection of latency time.

Latency time detection from a visual or auditory cue during speech production is essential for our studies, yet rather complicated to set-up reliably. The "voice key" systems usually proposed (E Prime, DMDX) are activated from one adjustable amplitude threshold, which means that various parasitic sounds (due to the microphone and ambient noise) may disturb the detection of the actual onset of the speech signal. In order to avoid counting these parasitic sounds as part of the pertinent speech signal, we have taken advantage of the energy characteristics of the speech signal.

This energy is calculated, and then smoothed, with a time constant of 10 ms. The amplitude obtained is then converted into dB. Another smoothing process is carried out with a time constant of 50 ms. This second smoothing process allows us to eliminate the short duration parasitic sounds and to obtain a preliminary assessment of the position of the signal onset. From this preliminary position, we work again backwards on the first energy signal (smoothed with a time constant of 10 ms) to detect the onset with more precision. This double process allows us to take the beginning of the word well into account.

The following example shows how the algorithm functions with the signal corresponding to the word, 'danse' (dance).

The Fig. 1 shows the speech signal. The true beginning of the word, taking into account the pre-voicing (negative voice onset time, VOT) of the plosive /d/ is located, as indicated by the vertical dotted line, at 635 ms. Note also, the two parasitic impulses which precede the dotted line and have to be eliminated.

Fig. 2 shows the energy after the two successive smoothing processes. The two parasitic impulses have almost disappeared, but the energy threshold is reached at 662 ms (see the dotted line), after pre-voicing onset.

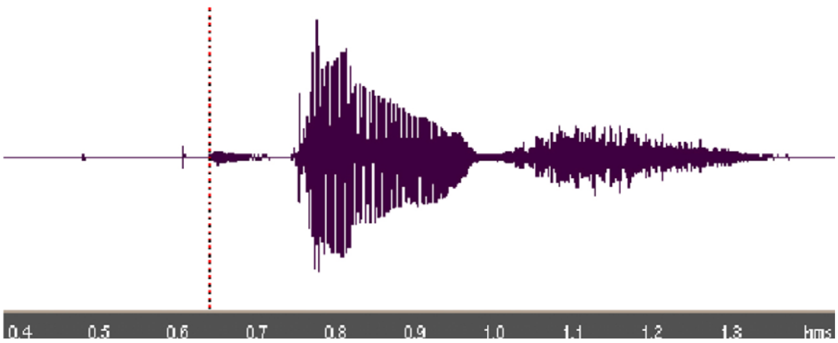

Fig. 1. Speech signal for the word 'danse' (dance) and for the two parasitic sounds which have to be eliminated.

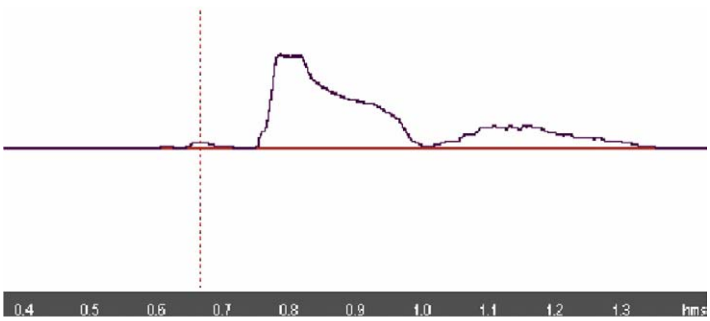

Fig. 2. Energy after the two successive smoothing processes.

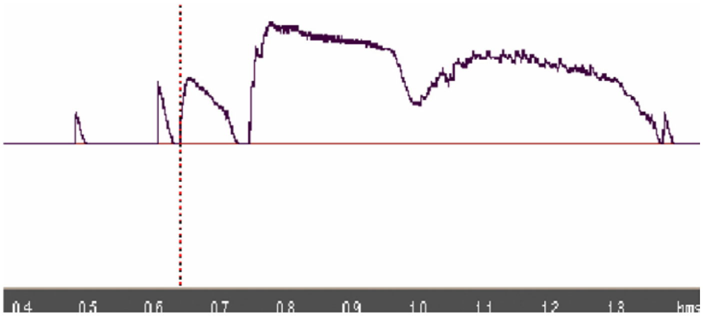

Fig. 3. Energy, in dB, smoothed with a time constant of 10 ms.

Fig. 3 shows the energy, in dB, smoothed with the first time constant of 10 ms. On this curve, by looking backwards from the preceding dotted line (see Fig. 2), we find the beginning of the energy rise corresponding to the beginning of the VOT. We thus obtain a starting point at 635 ms, which corresponds to the effective onset of the speech signal shown in Fig. 1. The two peaks preceding this point correspond to the two parasitic impulses and are not taken into account.

After using this program, we noticed that if the pertinent signal is framed by a sufficiently wide temporal window, the detection of the beginning of the signal is very reliable (less than 1/100 of the effective time, according to verification by listening). Parasitic impulses with very short durations are not taken into account. Unfortunately, parasitic noises such as coughing or hesitations are not eliminated. We have therefore created a complementary program which allows us to rectify the errors by visualizing and listening to the speech signal.

Detection of the duration of the vocal productions

The program developed also allows us to localize the end of the signal, and therefore to calculate the duration of the vocal production, which we cannot do with a voice key. To detect the end of the signal, we reverse the recorded speech signal and then use the same algorithm as we did for the detection of the signal onset. This detection is, however, less accurate than that carried out by the program for the calculation of the start of the signal; the gap between the values calculated with the algorithm and by listening may reach 1/10 of the actual time.

Seidenberg and McClelland, 1989

References

Ackerman, P.T., Dykman, R.A., 1993. Phonological processes, confrontational naming, and immediate memory in dyslexia. Journal of Learning Disabilities 26, 597–609.

Anthony, J.L., Lonighan, C.J., Driscoll, K., Phillips, B., Burges, S.R., 2003. Phonological sensitivity: a quasi parallel progression of word structure units and cognitive operations. Reading Research Quarterly 38, 470–487.

Backman, J., Bruck, M., Hebert, M., Seidenberg, M.S., 1984. Acquisition and use of spelling sound correspondences in reading. Journal of Experimental Child Psychology 38, 114–133.

Baddeley, A.D., 1986. Working Memory. Oxford University Press, Oxford.

Baddeley, A.D., Gathercole, S., Papagno, C., 1998. The Phonological loop as a language learning device. Psychological Review 105, 168–173.

Baddeley, A.D., Wilson, B.A., 1993. A developmental deficit in short-term phonological memory: implications for language and reading. Mem. 1, 65–78.

Beauvois, M.F., Derouesné, J., 1979. Phonological alexia: three dissociations. Journal of Neurology, Neurosurgery and Psychiatry 42, 1115–1124.

Booth, J.R., Perfetti, C.A., MacWhinney, B., 1999. Quick, automatic and general activation of silent word reading and phonological representations in young readers. Developmental Psychology 35, 3–19.

Bosman, A.M.T., de Groot, A.M.B., 1996. Phonological mediation is fundamental to reading: Evidence from beginning readers. The Quarterly Journal of Experimental Psychology 49A, 715–744.

Bowers, P.G., 1995. Tracing symbol naming speed's unique contributions to reading disabilities over time. Reading and writing: An Interdisciplinary Journal 7, 189–216.

Bowers, P.G., Swanson, L.B., 1991. Naming speed deficits in reading disability: multiple measures of a singular process. Journal of Experimental Child Psychology 51, 195–219.

Bruck, M., 1992. Persistence of dyslexics' phonological awareness deficits. Developmental Psychology 28, 874–886.

Bruck, M., Genesee, F., Caravolas, M., 1997. A cross linguistic study of early literacy acquisition. In: Blachman, B. (Ed.), Foundations of Reading Acquisition and Dyslexia: Implications for Early Intervention. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 145–162.

Bryant, P., Impey, L., 1986. The similarities between normal readers and developmental and acquired dyslexics. Cognition 24, 121–137.

Byrne, B., Freebody, P., Gates, A., 1992. Longitudinal data on the relations of word-reading strategies to comprehension, reading time and phonemic awareness. Reading Research Quarterly 27, 141–151.

Campbell, A., Butterworth, B., 1985. Phonological dyslexia and dysgraphia in a highly literate subject: a developmental case with associated deficits of phonemic processing and awareness. Quarterly Journal of Experimental Psychology 37a, 435–475.

Casalis, S., 2003. The concept of dyslexia. In: Nunes, T., Bryant, P. (Eds.), Handbook of Children's Literacy. Kluwer Academic Publishers, Dordrecht, pp. 257–273.

Castles, A., Coltheart, M., 1993. Varieties of developmental dyslexia. Cognition 47, 149–180.

Colé, P., Magnan, A., Grainger, J., 1999. Syllable-sized units in visual word recognition: evidence from skilled and beginning readers of French. Applied Psycholinguistic 20, 507–532.

Colé, P., Sprenger-Charolles, L., 1999. Traitement syllabique au cours de la reconnaissance de mots écrits chez des enfants dyslexiques, lecteurs en retard et normo-lecteurs de 11 ans. Revue de Neuropsychologie 4, 323–360.

Coltheart, M., Masterson, J., Byng, S., Prior, M., Riddoch, J., 1983. Surface dyslexia. Q. J. Exp. Psychol. 35, 469–595.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., Ziegler, J., 2001. DRC: a dual route cascaded model of visual word recognition and reading aloud. Psychological Review 108, 204–256.

Content, A., Radeau, M., 1988. Données statistiques sur la structure orthographique du Français. Cahiers de Psychologie Cognitive/ European Bulletin of Cognitive Psychology (Special issue), 87.

Courcy, A., Beland, R., Pitchford, N.J., 2000. Phonological awareness in French-speaking children at risk for reading disabilities. Brain and Cognition 43, 124–130.

Deltour, J.J., Hupkens, D., 1980. Test de vocabulaire actif et passif pour enfants (5 à 8 ans). E.A.P, Issy-les-Moulineaux.

Denkla, M.B., Rudel, R.G., 1976. Rapid automatized naming (R.A.N.): dyslexia differenciated from other learning disabilities. Neuropsychologia 14, 471–476.

Dunn, L.M., Thériault-Whalen, C.M., Dunn, L.M., 1993. Echelles de vocabulaire en image Peabody. Adaptation française du Peabody Picture Vocabulary Test-Revised. Psycan, Toronto.

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.

Elbro, C., Nielsen, I., Petersen, D.K., 1994. Dyslexia in adults: evidence for deficits in non-word reading and in the phonological representation of lexical items. Annals of Dyslexia 44, 205–226.

Fawcett, A.J., Nicolson, R.I., 1994. Persistence of phonological awareness deficit in older children with dyslexia. Reading and Writing: An Interdisciplinary Journal 7, 361–376.

Felton, R.H., Brown, L.S., 1990. Phonological processes as predictors of specific skills in children at risk for reading failure. Reading and Writing: An Interdisciplinary Journal 2, 39–59.

Ferrand, L., 2001. Cognition et lecture : processus de base de la reconnaissance des mots écrits chez l'adulte. De Boeck & Larcier, Bruxelles.

Frith, U., Wimmer, H., Landerl, K., 1998. Differences in phonological processing in German- and English-speaking children. Scientific Study of Reading 2, 31–54.

Génard, N., Mousty, P., Content, A., Alegria, J., Leybaert, J., Morais, J., 1998. Methods to establish subtypes of developmental dyslexia. In: Reitsma, P., Verhoeven, L. (Eds.), Problems and Interventions in Literacy Development. Kluwer, Dordrecht, The Netherlands, pp. 163–176.

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., Dalton, L., Schneider, W., 2001. Pseudohomophone effects and phonological recoding procedures in reading development in English and German. Journal of Memory and Language 45, 648–664.

Hanley, J.R., Hastie, K., Kay, J., 1992. Developmental surface dyslexia and dysgraphia: an orthographic processing impairment. Quarterly Journal of Experimental Psychology 44a (2), 285–319.

Henry, L., Millar, S., 1993. Why does memory span improve with age? A review of the evidence for two current hypothesis. European Journal of Cognitive Psychology 5, 241–287.

Holopainen, L., Ahonen, T., Lyytinen, H., 2001. Predicting delay in reading achievement in a highly transparent language. Journal of Learning Disabilities 34 (5), 401–413.

Jimenez-Gonzalez, J.E., Ramirez-Santana, G., 2002. Identifying subtypes of reading disability in the Spanish language. The Spanish Journal of Psychology 5 (1), 3–19.

Jorm, A., Share, D.L., Matthews, R., Maclean, R., 1986. Behaviour problems in specific reading retarded and general reading backward children: A longitudinal study. Journal of Child Psychology & Psychiatry 27, 33–43.

Kail, R., Park, Y.S., 1994. Processing time, articulatory time and memory span. Journal of Child Psychology & Psychiatry 57, 281–291.

Kessler, B., Treiman, R., Mullennix, J., 2002. Phonetic biases in voice key response time measurements. Journal of Memory and Language 47, 145–171.

Kirby, J.R., Parrila, R.K., Pfeiffer, S.L., 2003. Naming Speed and Phonological Awareness as Predictors of Reading Development. Journal of Educational Psychology 95, 453–464.

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., 2000. Deficits in phoneme segmentation are not the core problem of dyslexia: evidence from German and English children. Applied Psycholinguistics 21, 243–262.

Lecocq, P., 1991. Apprentissage de la lecture et dyslexie. Mardaga, Liège.

Lefavrais, P., 1967. Test de l'Alouette: Manuel. Les Editions du Centre de Psychologie Appliquée, Paris.

Lété, B., Sprenger-Charolles, L., Colé, P., 2004. MANULEX: a lexical database from French readers. Behav. Res. Methods Instrum. Comput. 36, 156–166.

Leybaert, J., Content, A., 1995. Reading and spelling acquisition in two different teaching methods: a test of the independence hypothesis. Reading and Writing: An Interdisciplinary Journal 7, 65–88.

Liberman, I.Y., Mann, V.A., Werfelman, M., 1982. Children's memory for recurring linguistic and non-linguistic material in relation to reading ability. Cortex 18, 367–375.

Liberman, I.Y., Shankweiler, D., Fisher, W.F., Carter, B., 1974. Explicit syllable and phoneme segmentation in the young child. Journal of Experimental Child Psychology 18, 201–212.

Lundberg, I., Hoien, T., 1989. Phonemic deficits: a core symptom of developmental dyslexia. The Irish Journal of Psychology 10 (4), 579–592.

Manis, F.R., Custodio, R., Szeszulski, P.A., 1993. Development of phonologic and orthographic skills: a 2-year longitudinal study of dyslexic children. Journal of Experimental Child Psychology 56, 64–86.

Manis, F.R., Seidenberg, M.S., Doi, L.M., McBride-Chang, C., Peterson, A., 1996. On the basis of two subtypes of developmental dyslexia. Cognition 58, 157–195.

Mann, V.A., Liberman, I.Y., 1984. Phonological awareness and verbal short-term memory: can they presage early reading problems? Journal of Learning Disabilities 17, 592–599.

Mann, V.A., Wimmer, H., 2002. Phoneme awareness and pathway to literacy: a comparison of German and American children. Reading and Writing: An Interdisciplinary Journal 15, 653–682.

Marmurek, H.H.C., Rinaldo, R., 1992. The development of letter and syllable effects in categorization, reading aloud and picture naming. Journal of Experimental Child Psychology 53, 277–299.

McBride-Chang, C., Manis, F.R., 1996. Structural invariance in the associations of naming speed, phonological awareness, and verbal reasoning in good and poor readers: a test of the double deficit hypothesis. Reading and Writing: An Interdisciplinary Journal 8, 323–339.

McDougall, S., Hulme, C., Ellis, A., Monk, A., 1994. Learning to read: the role of short-term memory and phonological skills. Journal of Experimental Child Psychology 58, 112–133.

Morais, J., Bertelson, P., Cary, L., Alegria, J., 1986. Literacy training and speech segmentation. Cognition 24, 45–64.

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

Mousty, P., Leybaert, J., 1999. Evaluation des habiletés de lecture et d'orthographe au moyen de la BELEC: données longitudinales auprès d'enfants francophones testés en 2° et 4° années. Revue Européenne de Psychologie Appliquée 49, 325–342.

Mousty, P., Leybaert, J., Alegria, J., Content, A., Morais, J., 1994. BELEC: une batterie d'évaluation du langage écrit et de ses troubles (1994). In: Grégoire, J., Pierart, B. (Eds.), Evaluer les troubles de la lecture: les nouveaux modèles théoriques et leurs implications diagnostiques. De Boeck, Bruxelles, pp. 127–1145.

Olson, R., Forsberg, H., Wise, B., Rack, J., 1994. Measurement of word recognition, orthographic and phonological skills. In: Lyon, G.R. (Ed.), Frames of Reference for the Assessment of Learning Disabilities: New Views on Measurement Issues. Brookes, Baltimore, pp. 243–275.

Paulesu, E., Demonet, J.F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., et al., 2001. Dyslexia, cultural diversity and biological unity. Science 291, 2165–2167.

Parrila, R., Kirby, J.R., McQuarrie, L., 2004. Articulation rate, naming speed, verbal short-term memory and phonological awareness: longitudinal predictors of early reading development. Scientific Study of Reading 8, 3–26.

Peereman, R., Content, A., 1999. LEXOP: a lexical database providing orthography-phonology statistics for French monosyllabic words. Behavioral Methods, Instruments and Computers 31, 376–379.

Pennington, B.F., Van Orden, G.C., Smith, S.D., Green, P.A., Haith, M.M., 1990. Phonological processing skills and deficits in adult dyslexics. Child development 61, 1753–1778.

Perfetti, C.A., Goldman, S.R., Hogaboam, T.W., 1979. Reading skills and the identification of words in discourse context. Memory and Cognition 7, 273–282.

Rack, J., Hulme, C., Snowling, M.J., Wightman, J., 1994. The role of phonology in young children learning to read words: the direct mapping hypothesis. Journal of Experimental Child Psychology 57, 42–71.

Rack, J.P., Snowling, M.J., Olson, R.K., 1992. The nonword reading deficit in developmental dyslexia: a review. Reading Research Quarterly 27, 29–53.

Ramus, F., Rosen, S., Dakin, S.C., Day, B.L., Castellote, J.M., White, S., et al., 2003. Theories of developmental dyslexia: insights from a multiple case study of dyslexic adults. Brain 126, 841–865.

Rapala, M.M., Brady, S.A., 1990. Reading ability and short-term memory: the role of phonological processing. Reading and Writing: An Interdisciplinary Journal 2, 1–25.

Rastle, K., Davis, M.H., 2002. On the complexities of measuring naming. Journal of Experimental Psychology Human Perception and Performance 28, 307–314.

Raven, 1947. Progressives Matrices Standard. E.A.P., Issy-les-Moulineaux (rééd. 1981).

Rey, A., Jacobs, A.M., Schmidt-Weigand, F., Ziegler, J., 1998. A phoneme effect in visual word recognition. Cognition 68, B71–B80.

Scarborough, H.S., 1990. Very early language deficits in dyslexic children. Child Development 61, 1728–1743.

Schatschneider, C., Fletcher, J.M., Francis, D.J., Carlson, C.D., Foorman, B.R., 2004. Kindergarten prediction of reading skills: a longitudinal comparative analysis. Journal of Educational Psychology 96, 265–282.

Seidenberg, M.S., McClelland, J.L., 1989. A distributed developmental model of word recognition and naming. Psychological Review 96, 523–568.

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

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.

Siegel, L.S., 1989. IQ is irrelevant to the definition of learning disabilities. Journal of learning disabilities 22, 469–478.

Snowling, M.J., 2000. Dyslexia. Blackwell, Oxford.

Snowling, M.J., Stackouse, J., Rack, J., 1986. Phonological dyslexia and dysgraphia: a developmental analysis. Cognitive Neuropsychology 3 (3), 309–339.

Sprenger-Charolles, L., 1994. L'acquisition de la lecture en français: étude longitudinale de la première à la seconde année du primaire. Année Psychologique 94, 553–574.

Sprenger-Charolles, L., 2003. Reading acquisition: cross linguistic data. In: Nunes, T., Bryant, P. (Eds.), Handbook of Children's Literacy. Dordrecht, Boston, New York, London. Kluwer Academic Publishers b.v, pp. 43–66.

Sprenger-Charolles, L., Bonnet, P., 1996. New doubts on the importance of the logographic stage. Current Psychology of Cognition 15, 173–208.

Sprenger-Charolles, L., Casalis, S., 1995. Reading and spelling acquisition in French first graders: longitudinal evidence. Reading and Writing: An Interdisciplinary Journal 7, 1–25.

Sprenger-Charolles, L., Colé, P., 2003. Lecture et Dyslexie: Approches Cognitives. Dunod, Paris (292 pages).

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., Pinton, F., Touzin, M., Béchennec, D., Billard, C., 2005. Significance and prevalence of dyslexics' phonological deficits in reading and reading-related tasks: comparisons with reading level controls (under revision).

Sprenger-Charolles, L., Siegel, L.S., 1997. A longitudinal study of the effects of syllabic structure on the development of reading and spelling skills in French. Applied psycholinguistics 18, 485–505.

Sprenger-Charolles, L., Siegel, L.S., Béchennec, D., 1998a. Phonological mediation and semantic and orthographic factors in silent reading in French. Scientific Studies of Reading 2, 3–29.

Sprenger-Charolles, L., Siegel, L.S., Béchennec, D., Serniclaes, W., 2003. Development of phonological and orthographic processing in reading aloud, in silent reading and in spelling: a 4 year longitudinal study. Journal of Experimental Child Psychology 84, 194–217.

Sprenger-Charolles, L., Siegel, L.S., Bonnet, P., 1998b. Phonological mediation and orthographic factors in reading and spelling. Journal of Experimental Child Psychology 68, 134–155.

Stanovich, K.E., 1984. The interactive compensatory model of reading: a confluence of developmental, experimental and educational psychology. Remedial and Special Education 5, 11–19.

Stanovich, K.E., 1992. Speculations on the causes and consequences of individual differences in early reading acquisition. In: Gough, P., Ehri, L., Treiman, R. (Eds.), Reading Acquisition. Erlbaum, Hillsdale, NJ, pp. 307–342.

Stanovich, K.E., 2000. Progress in understanding reading: scientific foundations and new frontiers. Guilford Press, New York.

Stanovich, K.E., Siegel, L.S., Gottardo, A., 1997. Converging evidence for phonological and surface subtypes of reading disability. Journal of Educational Psychology 89, 114–127.

Valdois, S., Bosse, M.L., Ans, B., Carbonnel, S., Zorman, M., David, D., et al., 2003. Phonological and visual processing deficits can dissociate in developmental dyslexia: evidence from two case studies. Reading and Writing: An Interdisciplinary Journal 16, 541–572.

Van Ijzendoorn, M.H., Bus, A.G., 1994. Meta-analytic confirmation of the non-word reading deficit in developmental dyslexia. Reading Research Quarterly 29, 266–275.

Vellutino, F.R., Scanlon, D.M., Sipay, E.R., Small, S.G., Pratt, A., Chen, R., et al., 1996. Cognitive profiles of difficult-to-remediate and readily remediated poor-readers: early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. Journal of Educational Psychology 88, 601–638.

Wagner, R.K., Torgesen, J.K., Rashotte, C.A., 1994. Development of reading related phonological processing abilities: new evidence of bi-directional causality from a latent variable longitudinal study. Developmental Psychology 30, 73–87.

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, 468–479.

Waters, G.S., Seidenberg, G.M.S., Bruck, M., 1984. Children's and adults' use of spelling sound information in three reading tasks. Memory & Cognition 12, 293–305.

Weschler, D., 1990. Echelle d'intelligence de Weschler pour enfants, forme révisée. Editions du Centre de Psychologie Appliquée, Paris.

Wimmer, H., 1993. Characteristics of developmental dyslexia in a regular writing system. Applied Psycholinguistics 14, 1–33.

Wimmer, H., 1996. The early manifestation of developmental dyslexia: evidence from German children. Reading and Writing: An Interdisciplary Journal 8, 171–188.

Wimmer, H., Hummer, P., 1990. How German speaking first graders read and spell: doubts on the importance of the logographic stage. Applied Psycholinguistics 11, 349–368.

Wolf, M., 1991. Naming speed and reading: The contribution of cognitive neurosciences. Reading Research Quarterly 26, 123–141.

Wolf, M., Obregon, M., 1992. Early naming deficits, developmental dyslexia and a specific deficit hypothesis. Brain and Language 42, 219–247.

Ziegler, J., Goswami, U., 2005. Reading Acquisition, Developmental Dyslexia and Skilled Reading Across Languages: A Psycholinguistic Grain Size Theory. Psychological Bulletin.

Ziegler, A., Perry, C., Ma-Wyatt, A., Ladner, D., Schule-Körne, D., 2003. Developmental dyslexia in different languages: language specific or universal? Journal of Experimantal Child Psychology 86, 169–193.