

Reading development in an orthographically regular language: effects of length, frequency, lexicality and global processing ability

Pierluigi Zoccolotti · Maria De Luca ·
Gloria Di Filippo · Anna Judica · Marialuisa Martelli

© Springer Science+Business Media B.V. 2008

Abstract The acquisition of reading skill was studied in 503 Italian children in first to eighth grade using a task that required reading of lists of words and non-words. Analysis of the metric characteristics of the measures indicated that reading speed but not accuracy was normally distributed across all ages considered. The role of specific effects (length, word frequency, and lexicality) versus global factors in reading speed was examined using the Rate–Amount Model (RAM). A global processing factor accounted for a large portion of the variance. Specific influences of length, frequency, and lexicality were detected in different periods of development over and above the global processing factor. Length modulated performance at early stages of learning and progressively less later on; in the case of non-words, the effect of length was large but did not change as a function of grade. The *lexicality* effect, present at all ages for high frequency words and by third grade for low frequency words increased with reading practice indicating a progressive differentiation in the ability to read words and non-words. Finally, the effect of word frequency was highest in third grade and then decreased. These findings are discussed in terms of their relevance for reading acquisition in a language with transparent orthography and their implications for evaluating developmental reading deficits. Overall, it is proposed that RAM is a useful tool for disentangling the role of specific versus global factors in reading development.

Keywords Frequency effect · Length effect · Lexicality effect ·
Rate-and-Amount Model · Reading

P. Zoccolotti (✉) · G. Di Filippo · M. Martelli
Psychology Department, Sapienza University of Rome, via dei Marsi 78, 00176 Rome, Italy
e-mail: pierluigi.zoccolotti@uniroma1.it

P. Zoccolotti · M. De Luca · A. Judica · M. Martelli
Neuropsychology Unit, IRCCS Fondazione Santa Lucia, Rome, Italy

Introduction

The acquisition of reading skills varies considerably in languages with opaque orthography, such as English and French, and in languages with more transparent orthography, such as Finnish, German or Italian. It is well documented that learning to read in English is characterized by a slow increase in accuracy; although this is quite apparent for irregular words, many errors in reading regular words can still be expected even after several years of schooling (e.g., Coltheart & Leahy, 1996).

In languages with more transparent orthography, high levels of reading accuracy are reached quite rapidly (e.g., German: Wimmer & Hummer, 1990; Italian: Cossu, Gugliotta, & Marshall, 1995). This is certainly true for Italian. Elementary school teachers in Italy claim that first graders are able to read most short words in the language by Christmas. This was confirmed in a study that investigated reading in children in the fourth month of first grade (Orsolini, Fanari, Tosi, De Nigris, & Carrieri, 2006). About half of the children were able to read approximately 80% of the words correctly. Similarly, Cossu (1999) reported high performance for reading short and long words and non-words in children tested between the end of January and late March of first grade. In a study that compared learning to read in several European languages, Seymour, Aro, and Erskine (2003) found that by the end of first grade Italian children read correctly approximately 95% of short familiar words in a list. Thus, compared to the orthography of several other European languages, that of Italian seems to be one of the easiest to learn (Seymour et al., 2003).

Only a few studies have evaluated Italian children's reading development in elementary and middle school. In one study, Tressoldi (1996) studied children's reading of lists of words and non-words and observed low error rates even at the lowest age tested (second grade). Some improvement in performance was detected at higher grades, however, by third grade the children's error rate was less than 5% (Tressoldi, 1996).

Although high levels of accuracy are present early in development, in transparent orthographies *reading speed* is expected to improve more slowly. This point was made most clearly by Wimmer and colleagues (Landerl, 2001; Wimmer, 1993; Wimmer & Mayringer, 2002; Wimmer, Mayringer, & Landerl, 1998). In a systematic series of investigations of Austrian children, these authors found that reading speed was the most critical parameter for documenting reading acquisition and breakdown in German, a language with shallow orthography. Data on reading speed are available for Italian children from second to eighth grade (Tressoldi, 1996). However, the material used in the study (from Sartori, Job, & Tressoldi's, 1995 test battery) contained words that varied for both frequency and length; thus, it was impossible to evaluate the two effects separately and to determine their modulating role on reading performance.

The general aim of the present study was to examine the reading performance of readers of Italian through a large span of schooling experience (from first through eighth grade); we focused on segregating the role of specific and global factors that affect word recognition. Research on experienced adult readers demonstrated that several factors exert a specific influence on the reading of Italian words, i.e., frequency, number of orthographic neighbors, and word length (see Barca, Burani,

& Arduino, 2002). However, only a few studies have dealt with the impact of these factors during reading acquisition and none has examined them as a function of the full range of elementary and middle school grades. In this study, we focus on three factors: length, lexicality, and frequency. We also wished to establish the possible role of global information processing factors in reading acquisition. Throughout childhood and adolescence, there are consistent age differences in speed of processing; these appear to reflect some general (i.e., non-task specific) component of performance that influence the quality of performance across a variety of cognitive domains (Kail, 1991; Kail & Salthouse, 1994). It seemed important to partial out the role of global components in evaluating the specific role of length, frequency, and lexicality. It should be noted that, on cross-sectional studies of proficient readers, age and reading experience strictly co-vary during development and their possible relative role cannot be determined; therefore, we will refer to the different groups in terms of their grade level. To carry out the investigation, we chose the reading of lists of words that varied for frequency and length and lists of non-words that varied for length. We expected this procedure to be sufficiently sensitive to capture the effects of the manipulated variables and, at the same time, simple enough to represent a potential instrument for clinical assessment. Below we summarize the available evidence on the role of specific (length, lexicality and frequency) and global factors in reading with special emphasis on studies on Italian children.

In early reading phases, a critical factor that affects decoding is the *length* of the orthographic string. The effect of length should manifest in two ways. First, greater latencies are expected for longer words prior to the onset of pronunciation. Second, more time is needed to pronounce longer words than shorter ones. Clearly, in reading lists of words and non-words, reading time is a combined measure of these two components. The latency before pronunciation should reflect the first stages of word decoding; it is commonly examined in terms of vocal reaction times (RTs) for naming words of different length. Very clear developmental differences have been reported using this paradigm. In first grade, children showed a mean RT increase of 173 milliseconds per letter, which decreased progressively in second and third grade (Zoccolotti et al., 2005). Burani, Marcolini, and Stella (2002) reported a progressive decrease in the influence of length from third through fifth grade. Finally, small, but detectable, word length effects were found in adult readers (Barca et al., 2002; Bates, Burani, D'Amico, & Barca, 2001). In contrast, no data are available to date as to the effect of length in the case of non-words. Studies on English subjects generally indicate a stronger effect of length in the case of non-words than words (e.g., Weekes, 1997). In research on reading, pronunciation time has received less attention than vocal RTs, presumably indicating greater interest in the decoding and planning processes than in the articulatory components of reading. Furthermore, individual differences in speech rate do not seem to contribute appreciably to reading speed in either opaque (Ackerman & Dykman, 1993; Ellis, 1985) or transparent (Wimmer, 1993) orthographies. In general, more time is needed to pronounce longer than shorter words. Reading time is linearly related to the number of phonemes in the word plus a fixed additional time (intercept of the linear regression; for a discussion see Appendix in Trueswell, Tanenhaus, & Garnsey, 1994).

Developmental data on the effect of word length on the reading of lists of words and non-words are not available for Italian children. Such data may prove useful in various ways. First, they will allow evaluating the role of length across a wider spectrum of grades than what done so far. Second, they will permit a novel test of the effect of length in the case of non-words and a comparison with that of words. Finally, this instrument may prove useful in the evaluation of dyslexic children. In fact, eye movements (De Luca, Borrelli, Judica, Spinelli, & Zoccolotti, 2002; De Luca, Di Pace, Judica, Spinelli, & Zoccolotti, 1999), vocal RTs (Judica, De Luca, Spinelli, & Zoccolotti, 2002; Zoccolotti et al., 1999), and lexical decision data (Di Filippo, De Luca, Judica, Spinelli, & Zoccolotti, 2006) indicate that Italian dyslexics are particularly sensitive to the effect of word length.

The *lexicality* effect (i.e., advantage of words over non-words) refers to the difference in reading strings of letters that represent items from the lexicon, i.e., words, and strings that can be pronounced but do not have entries in the lexicon, hereafter non-words. One well-known theoretical framework proposes that since non-words cannot be read by means of the lexicon they constitute a specific way to evaluate efficiency in the use of the grapheme-to-phoneme rules, or the non-lexical routine in reading (e.g., Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). However, a facilitatory effect due to the presence of lexical morphemes in the non-word has been shown in Italian (e.g., Arduino & Burani, 2004; Job, Perressotti, & Cusinato, 1998). Burani et al. (2002) provided information on vocal RTs for naming non-words in third- to fifth-grade children. These authors detected an effect due to morphological structure. Tressoldi (1996) reported data on the reading of lists of words and non-words in second to eighth graders. He found a clear effect of lexicality for both reading speed and reading accuracy in all groups of children. The lexicality effect was extensively referred to in studies on dyslexic children. Several authors (e.g., Rack, Snowling, & Olson, 1992) reported that dyslexics selectively fail in reading non-words. A marked lexicality effect would indicate a phonological disturbance in dyslexia (for meta-analyses of this effect see Rack et al., 1992; Van IJzendoorn & Bus, 1994).

Word frequency is one of the factors most studied in word recognition, at least in opaque orthographies such as English. In Italian, a frequency effect was reported using vocal RTs in naming in adults (Barca et al., 2002; Bates et al., 2001) and in elementary school children (Burani et al., 2002). Burani et al. (2002) found an influence of word frequency already in third graders. The effect remained constant in fourth and fifth grade and did not interact with word length, pointing to the independence of the two effects (Burani et al., 2002). To date, no developmental data on the effect of word frequency on reading lists of words and non-words are available for Italian. High frequency words are contrasted with low frequency words to measure a frequency effect. This provides information about the structure of the orthographic lexicon as a function of reading acquisition. Additionally, extending the analysis to first and second grade enables seeing whether a time can be detected in early stages when frequency does not yet influence reading. This would suggest that only a non-lexical procedure is used before that time. Also, it might be interesting to evaluate whether prolonged reading practice (as in the case of upper middle school children) modifies the size of the frequency effect.

Along with the effect of specific factors, such as those already described, it should not be overlooked that *global ability in information processing* is expected to change considerably during the elementary and middle school years (De Brauwer, Verguts, & Fias, 2006; Kail, 1991). Although this observation may appear intuitively clear, only a few studies (e.g., Kail & Hall, 1994) have attempted to evaluate the contribution of the global processing factor in the acquisition of written language. Using a path analysis paradigm, Kail and Hall (1994) investigated the connection between domain-specific and global processes in reading. In particular, they examined Spring and Davis's proposal (1988) that automatization in digit naming is a prerequisite of reading skills and, in turn, automaticity is a function of age-related experience. Note that in this study (as well as in the recent one by Bonifacci & Snowling, 2008) the authors identified a measure that was intended to capture global processing independent of reading. A different approach is that proposed within models such as the Rate-and-Amount Model (RAM), proposed by Faust, Balota, Spieler, and Ferraro (1999), or the State Trace Model (Bamber, 1979). In this perspective, it is assumed that (most) measures of performance will capture the influence of both global and specific influences. Accordingly, the aim of these models is to segregate the global and specific components of performance within each measure.

In the present study, we follow this latter perspective and analyze our data according to RAM (Faust et al., 1999), a model that makes explicit predictions about the general and specific components of individual variations in information processing. Two characteristics of RAM are important for the present study. First, RAM makes explicit predictions about the global factor that contributes to individual performance on different tasks; e.g., it predicts a linear relationship between the condition means for a group and those of a different group that varies for global processing ability (e.g., children at different grade levels). Second, it allows disentangling the contribution of general and task-specific factors to individual performance. With development, we should expect general changes in performance as well as variations in the influence of the parameters manipulated (i.e., frequency and length) on reading rate. In standard parametric analyses (such as ANOVAs), the first changes are expressed as main effects and the second as interactions between the parameter investigated and grade. However, overall changes in performance can directly influence the size of the interaction (the so-called over-additivity effect). This point is frequently raised in RT measurements (Salthouse & Hedden, 2002). An effect due to any experimental manipulation should be smaller for a reader with relatively fast RTs and larger for one with slower RTs. Therefore, accepting *prima facie* the results of ANOVAs on raw data creates a bias toward interpreting the effect as due to the specific manipulation, while it could in fact be entirely (or in part) explained by changes in global information processing.

Concerning reading time measurements, the effect of a variable (e.g., word length) might be smaller at later ages when reading times generally shorten because of a general performance change not because of a genuine change in the influence of that specific variable. As a consequence, a "spurious" interaction may be produced. Considering the direction of this effect, we will refer to it as the "under-additivity effect". Note that under-additivity can also reduce a "real" effect. Thus, if the

modulating role of a variable increases with age, the actual difference observed between two critical conditions will be attenuated at later ages because of the general improvement in performance (i.e., reduced reading times). To analyze reading performance as a function of grade, appropriate score transformations will be used to control for this under-additivity effect (see “Data analysis” section). Such transformations are appropriate in the case of time measurements but are not applicable to closed scale measures, i.e., the number of correct responses (accuracy). Therefore, to examine the specific effects of length, lexicality, and frequency we will focus mainly on reading speed measures.

We were also interested in taking into account the metric characteristics of our measures in capturing reading ability at different ages. To this aim, we examined the score distributions of accuracy and speed as a function of grade. If accuracy saturates rapidly in development, the error distribution will likely deviate from normality in the first few grades. By contrast, it is likely that reading speed will be less sensitive to ceiling effects and will capture individual differences in performance more systematically across development. In order to assess reading ability, it is important to establish the metric characteristics of reading speed and accuracy. In fact, individual performance is commonly evaluated on a scale relative to the performance of the sample tested for standardization purposes (i.e., by means of normalized scores). This inferential procedure is based on the assumption that the performances of the sample used for standardization distribute normally and that, consequently, means and standard deviations can be reliably estimated. While the reviewed literature suggests that reading speed is a more sensitive indicator of performance than accuracy in languages with transparent orthographies, a systematic evaluation of the characteristics of these measures is lacking for Italian.

Operationally, in the present study we examined the performances of Italian children from first through eighth grade in reading lists of words (which varied for frequency and length) and non-words (which varied for length) to:

- define the metric properties of speed and accuracy measures across this range of schooling;
- establish the contribution of a global factor in information processing to reading acquisition;
- evaluate the specific role of stimulus length, lexicality, and word frequency in reading as a function of grade.

Method

Participants

The study was part of a research agreement between the IRCCS Fondazione Santa Lucia and four public schools in middle class areas of Rome. The research project involved testing reading and related abilities in Italian elementary and middle schools. As part of the research protocol, the parents received a description of the study and had to approve their child’s participation. Subsequently, all information

Table 1 Sample characteristics

Grade	N	Males	Females	Age	
				M	SD
1st	52	31	21	6.8	0.3
2nd	46	32	14	7.6	0.3
3rd	56	31	25	8.7	0.4
4th	51	31	20	9.8	0.3
5th	42	22	20	10.8	0.3
6th	109	61	48	11.8	0.4
7th	83	30	53	12.8	0.5
8th	64	31	33	13.9	0.6

concerning individual performance was considered private and was analyzed strictly for research purposes. We tested 503 first-to-eighth graders from a total of 25 classes. A few children did not participate because they were absent on the testing days. Table 1 shows the demographic characteristics of the sample as a function of grade.

Materials

Stimuli were words that varied for frequency and length and non-words that varied for length. Four- to five-letter words were considered “short” and eight- to ten-letter words “long”. The same held for non-words. Short words and non-words were two-to-three syllables long; long words and non-words were three-to-five syllables long. Frequency was based on the *Vocabolario Elettronico della Lingua Italiana* (“Electronic Vocabulary of the Italian Language”; IBM Italia, 1989). Frequency between 10 and 100 (on 10,000,000 occurrences) was considered “low”; frequency between 1500 and 10000 was considered “high”. The list of stimuli is presented in Appendix. Original stimuli can be downloaded at <http://www.hsantalucia.it/modules.php?name=Content&pa=showpage&pid=1032>.

Stimuli were presented in four 30 word lists and two 30 non-word lists. There was one list of high frequency short words, one with high frequency long words, one with low frequency short words and, finally, one with low frequency long words. Median frequency was 9,091, 6,795, 74 and 65 in the four lists, respectively. Non-words are pronounceable strings of letters matched for length with the short and long words. Two practice lists of 20 words or non-words were also used.

Stimuli were laser printed in small case, Palatino font, size 12, and arranged in two vertical columns.

The participant’s task was to read each list of stimuli as quickly and accurately as possible.

Procedure

Children from second grade on were examined during the months of March and April; children in first grade were examined in May. Thus, the time difference

between grades was circa 12 months in all cases except first and second grade where it was between 10 and 11 months.

Participants were tested individually in a quiet room. Stimuli were placed on a flat horizontal surface at a comfortable distance. A short 20 non-word practice list was administered and then the short and the long non-word lists were presented. Similarly, a 20-word practice was given before the word lists; these were ordered as follows: high frequency short words, high frequency long words, low frequency short words and low frequency long words.

The time needed to complete the task was measured separately for each list with a stopwatch. The dependent measure was the mean time in seconds per list. Thus, we measured the time needed by the reader independently from accuracy. However, errors were also measured during the test administration. One error score was given to an item (either word or non-word) independent of the number of pronunciation errors made on that item. Omissions, insertions, reversals or substitutions of letter, and wrong stress assignment were considered as pronunciation errors. **Also self-corrections (but not hesitations) were scored as errors.** For off-line checks of reading times and errors the participant's vocal output was also tape-recorded.

Data analysis

Firstly, we examined the metric characteristics of accuracy and speed measures in evaluating reading at different grades. The score distributions for reading speed and reading errors were evaluated as a function of grade. For this purpose we controlled whether a Gaussian model was suitable for representing the distribution of speed and accuracy measures in each age range. The frequency of occurrence of each case (speed range or number of errors) was computed for each condition in each age group, separately for reading speed and accuracy. In the case of reading speed, data were grouped using a fixed range of 4 s. The range was chosen to capture the score distributions from first- to eighth-graders in all conditions. **To control for the difference in sample size across grade level** the frequency was normalized by the total number of cases. We applied a Gaussian model to the data:

$$f = a + b(\exp(-(x - \mu)^2/\sigma^2)) \quad (1)$$

as a function of reading time and accuracy, where a is the value that the tails of the curve approach for large values of x , b is the peak of the curve, μ is the mode of the distribution and σ is the bell width. To exemplify these findings, we graphically present the data obtained in the long high-frequency word condition (distributions for all other conditions can be requested from the authors). Kolmogorov–Smirnov tests (χ^2) were also carried out on the distributions of errors, separately for each condition and subgroup of children; due to the large number of tests ($N = 48$), based on the Bonferroni correction, an alpha level of .001 was adopted.

In a bell curve, mean and mode are coincident. In the case of departure from normality, one might expect the mean value of the raw data to differ from the mode of the distribution. We graphically compared the mean and the mode in expressing

the central tendency of the data obtained when speed and accuracy were measured. The mean was calculated on the raw data; the mode is the peak of the Gaussian distribution fitted to the data. In the case of departure from normality at small x , the parameter μ obtained by the fit was negative. In these cases the mode of the distribution was set to zero.

To describe the rate of reading acquisition during development, the form of the developmental function relating speed or accuracy to grade was examined (to permit a comparison between speed and accuracy, mode data were used in this analysis). A power fit

$$y = ax^{-b} \quad (2)$$

and an exponential fit

$$y = ae^{-bx} \quad (3)$$

was performed on the data, where a is the intercept and b the slope of the function. By calculating the coefficients of determination we estimated how well this family of curves fit the original data. Due to the similarity between the models all curves give good fits by normal standards. Newell and Rosenbloom (1981) recommend examining the shape distortion of each family of curves in order to choose the best model representing the data. Therefore, to compare the power and the exponential models we also looked at the distributions of the residuals.

Another aim of the study was to investigate the effects of length, frequency, and lexicality during reading acquisition. As it will be apparent in “Results” section, results of the fit of the Gaussian model indicate that only speed measures are suitable for inferential statistics. Consequently, parametric analyses to test the effect of grade, length, frequency, and lexicality on reading were carried out on speed measures only, while their effects in modulating measures of accuracy were only presented graphically (see below).

As noted in the Introduction, in standard parametric analyses the effect of specific manipulations on different groups of subjects may be altered if they vary in terms of general ability in information processing. RAM makes several explicit predictions about this general factor (Faust et al., 1999). Therefore, before presenting data on ANOVAs we reported data on the fitness of the reading time data to the predictions of RAM. In particular, we tested the predictions referring to group comparisons. Namely, we examined the prediction of a linear relationship between the condition means of two groups varying for global processing ability. To this aim, we compared the performances of children with the least amount of reading experience (first graders) and the longest experience (eighth graders). However, we will also mention the results obtained comparing intermediate grades. Further, RAM predicts a linear relationship between the condition means and the standard deviation in the same conditions. The presence of these relationships is a condition supporting the usefulness of adopting data transformations that partial out the effect of global processing ability. Faust et al. (1999) proposed various methods for accomplishing this, including the use of z score transformations. **Z scores are based on the individual means in all conditions** and indicate an individual participant’s

performance in a given condition relative to all others. This transformation controls for changes in general performance that take place with age. Consequently, the interactions between reading practice and experimental manipulations tested here can be obtained independent of the spurious effect of general performance.

Based on the results of these tests (see below), we decided to run both ANOVAs on raw and z score data. First, we carried out an ANOVA on the raw data with grade (eight levels) as unrepeated factor and length (short, long) and condition (high frequency words, low frequency words, non-words) as repeated measures. All main effects and interactions were highly significant (at least $p < .001$). Therefore, we performed separate analyses for each effect (e.g., length in the case of high frequency words, length in the case of low frequency words, etc.). Only the results based on these latter ANOVAs will be presented. When appropriate, the *a posteriori* Tukey HSD test was used to de-compose interactions; considering the large number of comparisons, a conservative $p < .01$ significance level was adopted. Second, the same analyses were carried out as above using the z scores.

We present the results of the ANOVAs on both raw data and z scores. As stated by Faust et al. (1999), comparing raw data and z score results provides information on the under-additivity influence in modifying differences between conditions as a function of reading practice. We will describe each of the three effects (length, lexicality, and frequency) separately by pooling together qualitative considerations based on the inspection of figures and major statistical results.

To investigate the impact of length, lexicality, and frequency on accuracy we examined the modes of error distributions as a function of grade for each condition. A comparison of the effects of our variables on accuracy and the effects obtained on speed could reveal the presence of trade-offs between these measures.

Finally, we looked at the pattern of correlations between speed and accuracy in each grade, separately for each condition. Since the accuracy distributions deviated considerably from normality, Spearman non-parametric correlations were carried out. Due to the large number of tests ($N = 48$), based on Bonferroni correction, we adopted an alpha level of .001.

Results

Metric characteristics of accuracy and speed

Figure 1 shows the speed (Fig. 1a) and accuracy (Fig. 1b) distributions obtained by each of the eight age groups while reading long high frequency words.

An inspection of the figure indicates that both speed and accuracy are higher in later grades. However, reading time is normally distributed in every age group, but accuracy is not. In second (and in higher) grade errors are not normally distributed around the mean; in third grade almost half of the children make no errors in reading the list of words; and in sixth (and higher) grade most children read without making any errors.

Kolmogorov–Smirnov tests generally confirmed these observations also for all of the other conditions. In the first grades no significant deviations were detected for

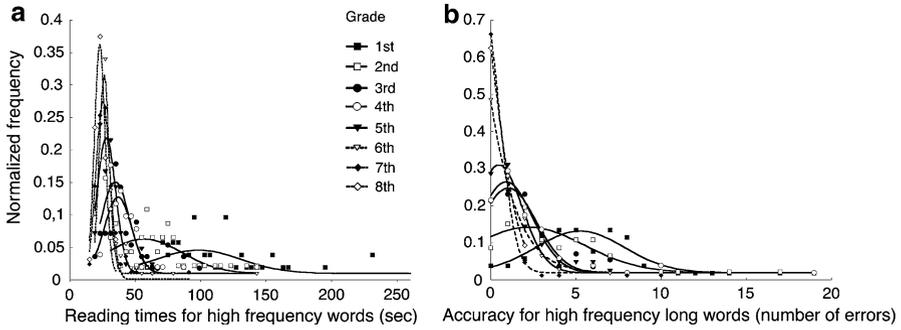


Fig. 1 Normalized frequency for each age group as a function of speed (a) and accuracy (b). The figure reports the results obtained by the children while reading long, high frequency words

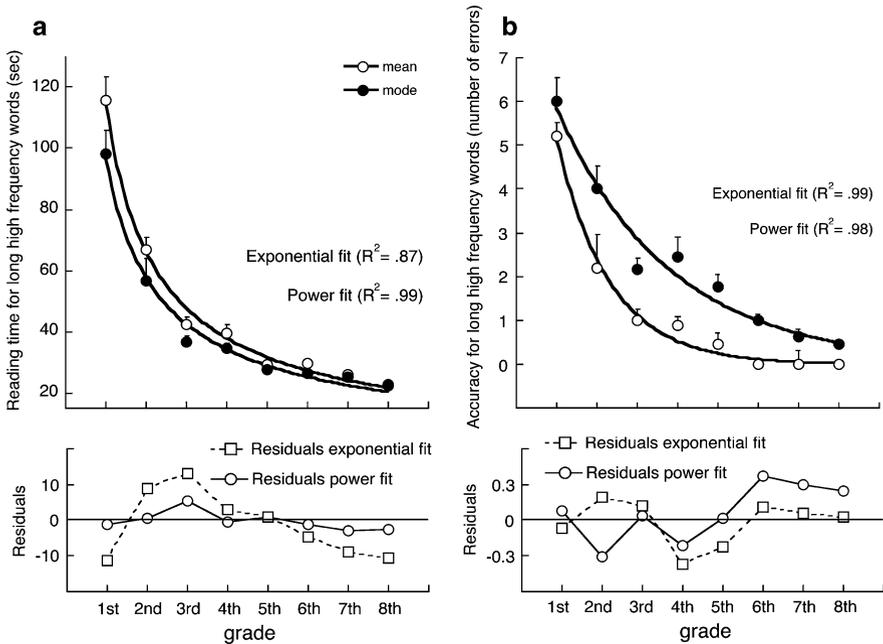


Fig. 2 Reading time (a) and accuracy (b) as a function of grade. The figure reports the data obtained while reading long, high frequency words. The mode (solid circles) and the mean (empty circles) are compared in the top part of the figure. Speed data are fit by a power function (a top), accuracy data by an exponential function (b top). The lower panels of the figure compare the residuals of the power and the exponential fit to the mode data for speed (a bottom) and accuracy (b bottom)

accuracy for any of the six conditions; from sixth grade on significant deviations from normality were observed in a large proportion of cases (12/18, i.e., six conditions by three grade levels). None of the total 48 tests carried out indicated a significant deviation from normality for speed.

Figure 2 compares the mode and the mean as a function of grade for speed (Fig. 2a) and accuracy (Fig. 2b) measures. Both the mean and the mode describe

well the performance as a function of grade when reading time is considered, as indicated by the lack of difference between the two measures (Fig. 2a). The mean is consistently higher than the mode for accuracy in all grades except for the first one (Fig. 2b).

To assess speed and accuracy as a function of grade a power and an exponential model were fit to the data. For speed, the mode distribution is better described by a power ($R^2 = .99$) than by an exponential ($R^2 = .87$) curve. There are systematic distortions from the exponential fit, while the power curve represents the data equally well in all grades considered, as shown by the trend of the residuals (lower panel of Fig. 2a). Raw data are higher (negative residuals), lower, and then higher again relative to the exponential model (These are the expected distortions for an exponential model in explaining a power relationship; see Newell & Rosenbloom, 1981). For accuracy, the relationship with grade is similarly represented by the exponential ($R^2 = .99$) and the power ($R^2 = .98$) models. However, the residuals of the exponential fit do not show the typical pattern of distortion evident in the case of speed measures, while the power fit fails to capture the tail of the distribution (lower panel of Fig. 2b). Therefore, in Fig. 2b we chose to present the exponential fit.

The same analyses performed on the other five conditions produced similar results. The proportion of explained variance for speed was consistently higher in the case of the power function (median $R^2 = .98$) than the exponential function (median $R^2 = .87$). A mixed pattern emerged for accuracy: high frequency short words ($R^2 = .99$ for both power and exponential function), low frequency short words ($R^2 = 1$ and $.99$ for power and exponential function, respectively), low frequency long words ($R^2 = .80$ and $.93$), short non-words ($R^2 = .93$ and $.87$), long non-words ($R^2 = .49$ and $.55$).

Comments

The results showed that reading speed but not accuracy was normally distributed across all ages considered. Consistently, the mode and the mean similarly estimated the change in reading speed as a function of grade. When accuracy was considered, the two indexes were similar in first grade but diverged considerably thereafter. The mode of the distribution dropped rapidly reaching floor in sixth grade, while the mean was not yet at floor for the older children tested. This suggests that, in the case of errors, the mode of the error distribution may be more representative in describing how performance improves with age. In clinical settings, pathological performance is usually expressed in terms of a scaled distance from the mean of the reference group; based on the present findings, this is likely to misrepresent the degree of reading impairment of a given child.

The rate of development in reading speed and accuracy was non-linear. Reading speed improved in a power relation with grade. A power law is often used to describe the effect of practice (Logan, 1992; Newell & Rosenbloom, 1981) and it indicates an improvement in performance that is progressively smaller over time. For accuracy, both power and exponential fits represented the data almost equally well. The exponential fit seemed preferable since it described the tail of the

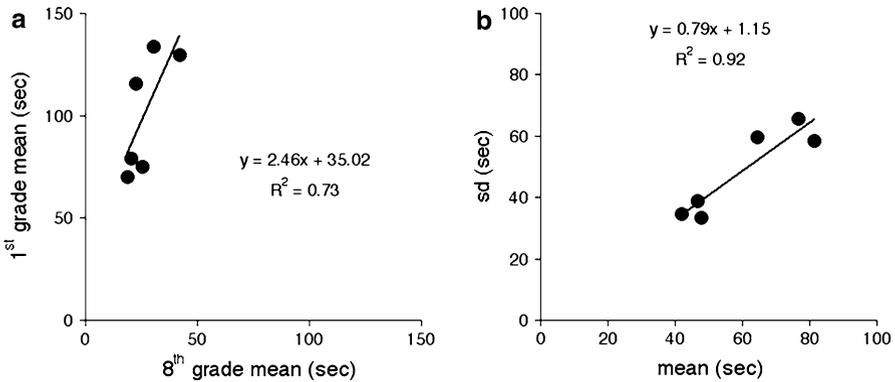


Fig. 3 Panel (a) reports a Brinley plot of the condition means for first and eighth graders. Panel (b) shows the relationship between standard deviations and mean group performances in all conditions

distribution better. It must be pointed out that studies on learning curves characteristically examine individual participants longitudinally (Heathcote, Brown, & Mewhort, 2000). A cross-sectional analysis, as in the present case, does not allow considering the mechanisms that may underlie learning, and the curves presented have a mainly descriptive aim.

Fitness of data to the RAM model

Figure 3a presents a plot of the condition means for the groups of the least (first grade) and most proficient (eighth grade) readers. The linear regression ($2.34x + 32.1$) accounts for 73% of variance (This estimate is the lowest one obtained when comparing children in first grade to those in second, third, fourth grade, etc.). Similar results were present when comparing groups of children enrolled in second through seventh grade.

Figure 3b plots the condition means for the general sample of children against the standard deviations on the same means. In this case, the linear regression ($0.47x + 6.0$) accounts for 92% of the variance.

Comments

With increasing experience in reading, reading times generally decrease. However, the relative difficulty of the various stimulus conditions remains similar. **According to the RAM (Faust et al., 1999) this indicates that a substantial part of the change in reading times (at least about 70% of the variance) is due to the influence of a global factor in information processing not to task-specific effects.** Similarly, the covariation between means and standard deviations is consistent with the view of a global factor underlying the data. These findings support the idea of partialling out the role of the global factor when evaluating the role of specific factors (such as length, frequency, and lexicality) on reading.

Reading time: effect of stimulus length

Raw data analyses

Reading time data are presented in Fig. 4 (top part). The figure contrasts reading time for long and short stimuli, separately for high frequency (a), low frequency (b), and non-words (c). An inspection of the figure indicates that:

- in all conditions, reading time decreases rapidly in early grades and then progressively less in higher grades. Note that, in all cases, the relationship with grade is well described by a power function. In all three ANOVAs the grade effect was significant (with $p < .001$);
- long words require longer reading times than short words in all cases, as expected (for all three ANOVAs the length effect was significant with $p < .001$);
- the difference between short and long words is smaller at higher grades for high frequency words, an effect that appears less pronounced for low frequency words and, even less, for non-words. In all three ANOVAs, the grade by length interaction was significant ($p < .001$): Tukey comparisons indicated that differences between short and long stimuli were always significant for all grades and types of stimuli.

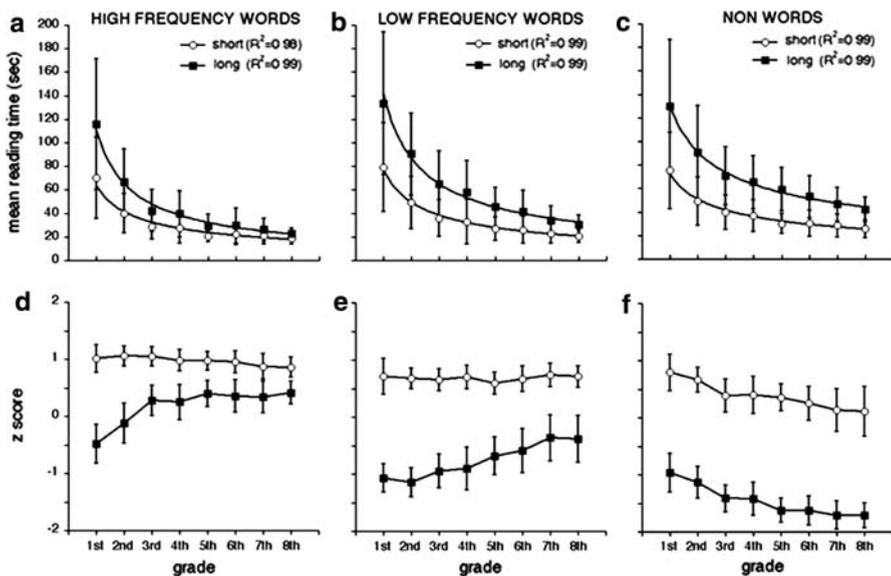


Fig. 4 Reading times (M and SD) for long and short stimuli as a function of grade. The top part of the figure reports raw data (interpolated by a power function); the bottom part presents the corresponding z scores. Data are reported for high frequency (a), low frequency (b) words and non-words (c)

Z score analyses

Figure 4 (bottom part) presents the same contrasts as the top part of the figure for reading time measures transformed in z values. An inspection of the figure indicates that:

- Short stimuli were “easy” in all grades and conditions, as indicated by the presence of positive z values. Long stimuli were comparatively “difficult”; except for high frequency stimuli above third grade, they were characterized by negative values. The length effect was significant ($p < .001$) in all three ANOVAs.
- The difference between short and long stimuli (length effect) was smaller for high frequency than for low frequency words and non-words. The length effect changed with grade more for high frequency than for low frequency words; no change in the length effect was apparent for non-words. The grade by length interaction was significant ($p < .001$) in the analyses of high frequency and low frequency words and non-significant in the analyses of non-words.

Z score differences

This length effect as a function of grade can be observed more directly in part a of Fig. 5. It shows the difference in z values between short and long words for the different stimuli used. The length effect was about 1.5 z value units for all stimuli in first grade; subsequently, it was smaller at higher grades for high frequency words and to a lesser extent for low frequency words. No change in the length effect was apparent with age for non-words.

Comments

The results confirmed the influence of *stimulus length* in modulating reading speed acquisition over and above the effect of the global changes in information processing ability.

Length markedly affected reading speed in early learning stages for all types of stimuli. These findings parallel those obtained with the vocal RT paradigm in naming (Burani et al., 2002; Zoccolotti et al., 2005). **At higher grades, the effect of length decreased considerably for high frequency words and less so for low frequency words.** This held for both standard and z -score analyses. In contrast, the change in the length effect with grade on non-words was present for raw data but disappeared when z scores were used. **Therefore, for non-words the stimulus length effect was constant at all grades,** and the decreasing effect observed in raw values should be interpreted as under-additivity, i.e., as due to a global improvement in information processing associated with age.

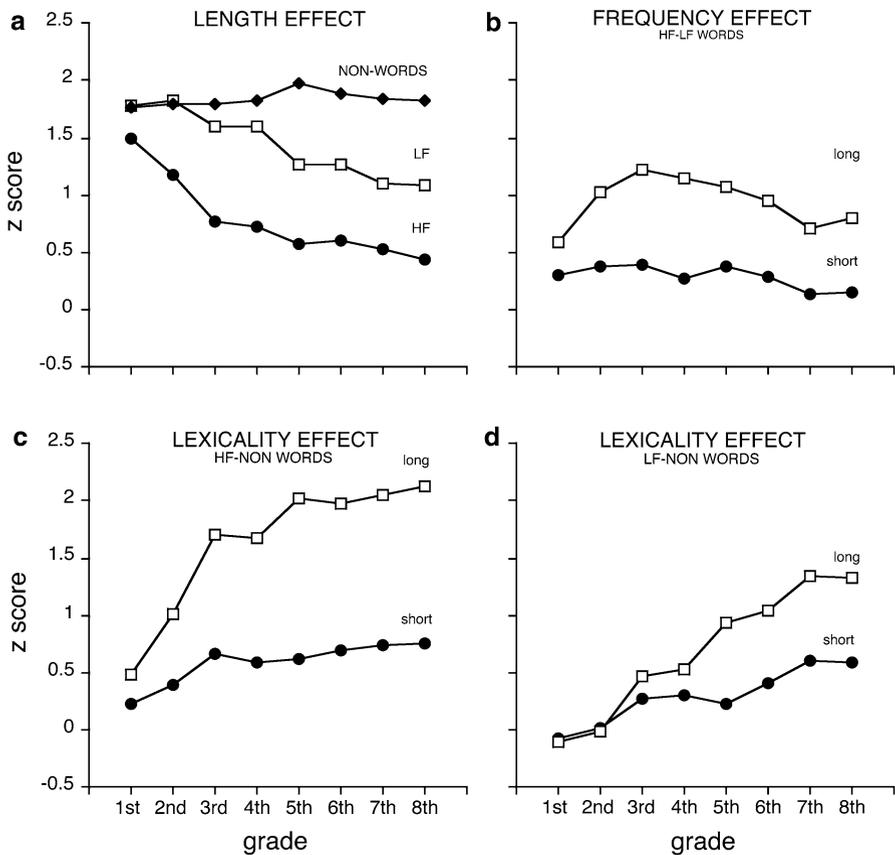


Fig. 5 The figure shows the three effects tested (length, lexicity and frequency) in terms of the difference in z value between the critical conditions: **(a)** difference between short and long stimuli separately for high frequency, low frequency, and for non-words; **(b)** difference between high frequency and low frequency words, separately for short and long stimuli; **(c)** difference between high frequency words and non-words, separately for short and long stimuli; **(d)** difference between low frequency words and non-words, separately for short and long stimuli

Reading time: effect of lexicity

Raw data analyses

Raw data highlighting the word frequency effect and the lexicity effects are presented in Fig. 6. Part *a* of Fig. 6 contrasts reading time for high frequency words, low frequency words and non-words in short stimuli as a function of grade; part *b* presents the same effects in long stimuli.

As to the contrast of raw data between words and non-words (lexicity effect), an inspection of Fig. 6 *a* and *b* indicates that:

- High frequency words were read faster than non-words at all grades; this lexicity effect tended to increase with grade level (for all main effects and

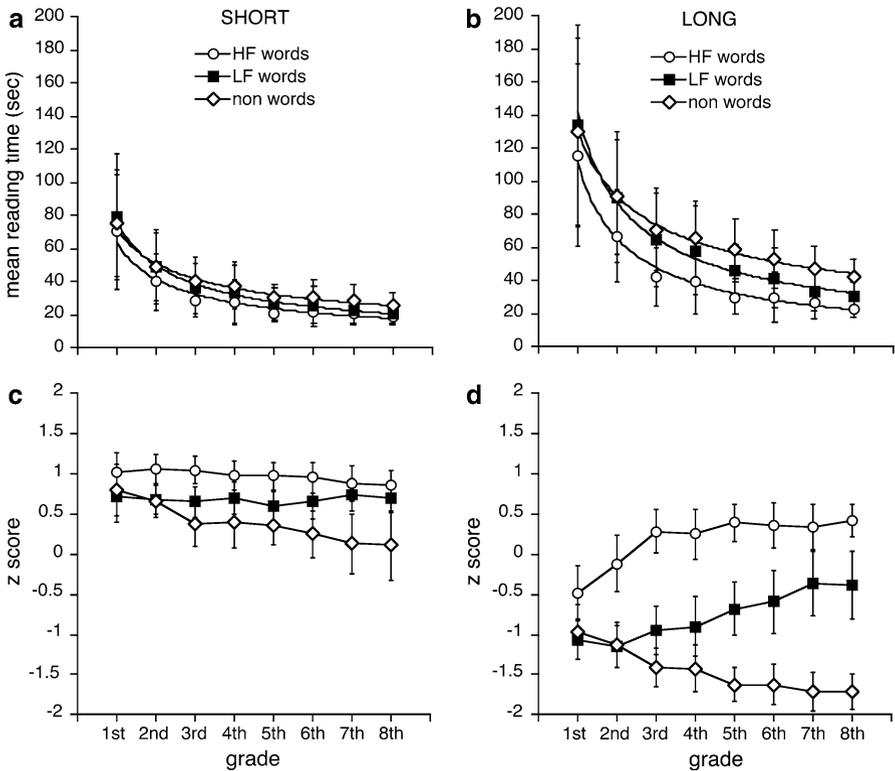


Fig. 6 Top part: mean (and SD) raw reading times for high frequency, low frequency and non-words as a function of grade (a power function is used to interpolate the data). Data are reported for short (a) and long (b) stimuli. Bottom part: z scores for high frequency, low frequency and non-words as a function of grade. Data are reported for short (c) and long (d) stimuli

interactions at least $p < .01$). This pattern held true for both short and long stimuli.

- In the case of long stimuli, low frequency words were not differentiated from non-words in first and second grade (Tukey test); in higher grades, a progressively larger lexicality effect was apparent (the main effects and interactions were significant with at least $p < .001$). In the case of short stimuli, the pattern was similar; however, there was no difference between low frequency words and non-words in first, second, third, and fifth grade.

Z score analyses

The lexicality effect as a function of grade can be appreciated in terms of z scores in Fig. 6c (short stimuli) and in Fig. 6d (long stimuli). An inspection of the two plots indicates that:

- High frequency words were read faster than non-words (Fig. 6c, d). The difference was present at all ages but was generally larger at higher grades. This pattern, present for both short and long words, was more evident for long stimuli (for main effects and interactions with lexicality, all $ps < .0001$).
- The pattern of the contrast for low frequency words and non-words was similar to that for high frequency words (this is more evident comparing the differential plots in Fig. 5c, d). However, a difference between low frequency words and non-words was found in third grade while no difference was detected in first or second grade (Tukey test).

Z score differences

The lexicality effect is expressed as the difference in z values in parts c (high frequency words versus non-words) and d (low frequency words versus non-words) of Fig. 5. The contrast between high frequency words and non-words increased with grade level; a similar effect, which started later in third grade, was present when low frequency words and non-words were contrasted. These two patterns were present for both short and long stimuli but they were more marked in the latter.

Comments

The *lexicality* effect was present at all ages when high frequency words were contrasted with non-words, confirming the influence of the lexicon on reading speed by the end of first grade. When low frequency words and non-words were contrasted, their differentiation was evident in third grade, indicating the enlargement of the lexicon by this age. At higher grades, a differentiation in the ability to read words and non-words was detected. This effect was clearer for z than for raw scores. Hence, in this case, under-additivity reduced the size of the effect. Although this pattern holds for all stimuli tested, the lexicality effect was generally greater for long than for short stimuli.

Reading time: effect of word frequency

Raw data analyses

An inspection of Fig. 6 (top) indicates that:

- High frequency words are read faster than low frequency words at all ages, including first grade. The main effect of frequency was significant for both short and long words ($p < .001$).
- The grade by frequency interaction was significant for both short and long stimuli ($p < .001$).
- This pattern was present for both short and long words, although differences were larger for the latter stimuli.

Z score analyses

The same contrasts as in Fig. 6a, b are presented at the bottom (Fig. 6c, d) in terms of z values. Concerning frequency contrast, an inspection of the figure indicates that:

- High frequency words yielded higher z scores than low frequency words. For short words, this difference was present at all ages but was smaller at higher grades. Also for long words the difference was present at all ages; however, it was smaller at both low and high grades (this effect is difficult to see in Fig. 6 but it is clear in the differential plot of Fig. 5b). The main frequency effects and the grade and frequency interaction were significant for both short and long stimuli ($p < .001$).

Z score differences

The frequency effect is expressed as the difference in z values in part *b* of Fig. 5. For short words, the effect is almost 0.5 z value units throughout elementary school; a small decrease with grade is detected in middle school. In long words, an inverted U slope emerged when high and low frequency words were contrasted; the frequency effect was maximal (more than a 1 z value unit) in third graders and smaller in lower and higher grades (about 0.5 z value unit).

Comments

Word frequency modulated reading speed at all ages, including first grade. However, as seen in the previous section, performance on low frequency words did not differ from that of non-words until third grade. Thus, it appears conservative to interpret the difference between high and low frequency words in the first two grades as a lexicality, not a genuine frequency, effect.

For long words, the effect of frequency was highest in third grade and smaller at later grades. The presence of this peak presumably indicates that at this age children are just starting to deal with low frequency words. Indeed, long high frequency words seem to be acquired first, as shown by the steep progressive reduction in the length effect obtained with these stimuli (Fig. 5a). As these words are consolidated in the lexicon the frequency effect increases, and reaches its full potency in third grade when children begin to acquire long low frequency words. Subsequently, mastering of low frequency items makes them more and more similar to high frequency words. Previous research showed that the word frequency effect was large and stable between third and fifth grade (Burani et al., 2002); this finding is generally consistent with the present pattern of results, which provides information on a larger range of grades.

The frequency effect was barely detectable for short words. There may be various reasons for this. The faster times for reading short words might have reduced the range of variation. Or, children may have used qualitatively different ways of decoding short and long words, and this may have affected reference to the orthographic lexicon. Proficient middle school readers are able to process words

with up to five letters in parallel but rely on sequential analysis for longer words (Spinelli et al., 2005). This limited number of characters processed in a glimpse may result in earlier consolidation in the lexicon for short low frequency words relative to long low frequency words.

Accuracy

Figure 7 presents the mode of the distributions of errors in the various conditions: short and long non-words and words of high and low frequency. Although the data present considerable variability, they show a characteristic pattern of improvement across the conditions. First, children learn to read without errors the high frequency, and then the low frequency words. The data also show the effect of word length: long words are read without errors later than short words. Results for long non-words are similar to the ones obtained with long low frequency word.

Comments

The analysis of accuracy in reading words showed a large improvement between first and second grade. Afterwards, only small changes were present indicating a ceiling effect for this parameter. Changes with grade were less marked for long low frequency words and non-words. The accuracy results generally confirmed the previously reported (Tressoldi, 1996) performance increase with reading experience.

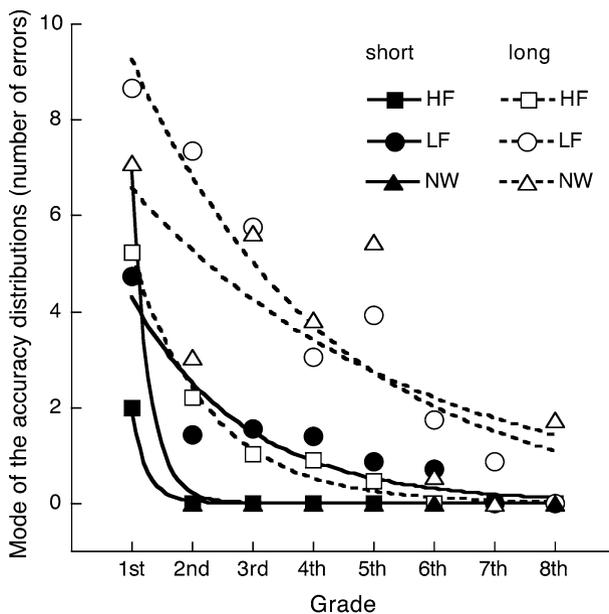


Fig. 7 Modes of the accuracy distributions as a function of grade. Data are fit by an exponential function. Results for short (filled symbols) and long (open symbols) stimuli are compared across conditions: high frequency words (squares), low frequency words (circles), and non-words (triangles)

Although we did not make a formal analysis of the length, frequency, and lexicality effects on accuracy, the change in the pattern of these effects with age generally matched the results obtained for speed. This indicates that there is no trade-off between time and errors as a function of grade.

Relationship between reading speed and accuracy

The correlations between reading time and accuracy are presented in Table 2 separately for each age group.

An inspection of the table indicates that:

- all but two correlation coefficients are positive, indicating more errors for slower children;
- correlations are generally higher at intermediate grades, the actual peak depending upon word length and frequency: speed and accuracy are significantly correlated at earlier grades for short than for long words and for high than for low frequency words. This pattern is less evident for non-words;
- reading time and accuracy are correlated at the beginning (first grade) only in the case of high frequency short words; they are not correlated at the highest grades tested in any of the conditions.

Comments

At early grades, accuracy and speed were only loosely related. Possibly, when reading performance is not fully established, different children may vary in their idiosyncratic tendency to be accurate or fluent. Note that the lack of relationship between accuracy and speed is not due to metric problems; in fact, the earliest grades present the largest spread in performance (approximating a normal distribution also in the case of accuracy). With greater reading experience, speed and accuracy were most highly correlated. This may indicate the emergence of a relatively homogenous

Table 2 Spearman correlation coefficients between speed and accuracy measures separately for condition and grade

Grade	HF short	HF long	LF short	LF long	NW short	NW long
1st	0.48	0.41	0.36	0.34	0.43	0.20
2nd	0.74	0.49	0.39	0.31	0.53	0.32
3rd	0.41	0.63	0.53	0.42	0.66	0.51
4th	0.35	0.66	0.66	0.56	0.59	0.39
5th	-0.19	0.39	0.30	0.60	0.38	0.30
6th	0.24	0.33	0.45	0.46	0.44	0.45
7th	0.13	0.23	0.09	0.21	0.16	0.25
8th	-0.01	0.21	0.07	0.26	0.05	0.16

Note: The numbers in bold are statistically significant considering the Bonferroni correction criterion for 48 comparisons; in bold $p < .001$

type of processing across individuals. It may be noted that, at least for stimuli with a lexical value, this occurred at earlier grades for generally easier conditions and at later grades for more difficult conditions; this pattern is in keeping with the idea that speed–accuracy correlation emerges with sufficient mastering of the stimulus materials. **Afterwards the size of the correlation reduced progressively, with no correlation being present for the two highest grades.** This latter finding may be due to a restriction of range for the accuracy measure; in fact, we have seen that errors are few at the highest grades even for the most difficult conditions.

Overall, correlational analysis confirmed that reading speed and accuracy should not be considered as equivalent measures of reading performance, although, at specific moments in reading acquisition, they show a relatively high degree of coherence.

Discussion

Throughout elementary and middle school, the Italian children in this study progressively optimized their reading behavior. Near-perfect accuracy was achieved relatively early, as one might expect for a language with quite regular orthography (Seymour et al., 2003). Changes in reading speed (as function of grade) were less marked and were detected at all ages tested. The present results are based on a cross-sectional study. Although the size of the sample is sufficiently large to accommodate for small sampling biases, a confirmation of these developmental trends based on a longitudinal study would be important.

Changes in reading ability as a function of grade seem to reflect both the influence of a global factor in information processing and the specific factors examined (length, frequency, and lexicality).

The idea that reading improves with years of schooling along with global ability in information processing may seem intuitively reasonable. However, with few exceptions (e.g., Kail & Hall, 1994), most authors focused on detecting the specific factors that contribute to reading acquisition (and disruption) and they neglected to study the role of a general factor affecting the acquisition of written language. Consequently, little is known about the characteristics of the global factor in information processing that contributes to reading. The Rate–Amount Model provides a useful framework for isolating the contribution of such a general factor (as well as the identification of specific factors), and this study represents a step in this direction.

A few preliminary remarks are in order. **First, the amount of variance in reading time explained by the global factor was relatively high even when the highest and the lowest grades were compared. This indicates that all reading conditions tested are loaded with the global factor in information processing.** This calls for caution in interpreting individual conditions at face value. For example, performance in reading a list of non-words expresses the child's general ability to deal with orthographic material *and* to use the non-lexical routine to convert graphemes to phonemes. These two components are inevitably interwoven in traditionally scored performances and their relative contributions cannot be discriminated. Second, for

all conditions tested, reading time was faster at higher grades, closely following a power relationship. Note that the power relation with grade held for all conditions tested, including non-words. Therefore, the increase in speed is evident also in the case of stimuli (non-words), for which no reading experience and no lexical entries were present. We have recently completed a study on the influence of global and specific components in dyslexia (Zoccolotti, De Luca, Judica, & Spinelli, 2008). Similar to the present results, the global factor accounted for naming of both words and non-words (but not pictures). It is intriguing that reading acquisition might have something in common with a large series of visual-motor and perceptual skills, such as mirror tracing, manufacturing cigars, reading inverted text, and scanning for visual targets (Newell & Rosenbloom, 1981). Most recent research on reading focussed on the cognitive underpinnings of this complex behaviour and neglected the contribution of perceptual and motor learning to reading acquisition. From a similar, but not identical, perspective to that of the present study, Nazir and colleagues (Nazir, Ben-Boutayab, Decoppet, Deutsch, & Frost, 2004) proposed that low-level perceptual factors significantly contribute to reading acquisition. Overall, the present results show that global changes in performance, which cannot be easily accounted for by specific factors (such as length, lexicality, and frequency), accompany reading. RAM appears as a useful reference to frame these relative influences. Further work aimed at characterizing the nature of the global factor in performance accompanying reading is currently ongoing in our laboratory.

Once the influence of the global factor in processing was taken into account, it was possible to evaluate the specific role of the factors tested (stimulus length, word frequency, and lexicality). The data showed that they played a role at different times in reading acquisition.

In transparent orthographies it is generally believed that children begin to read by emphasizing alphabetic analysis (Seymour et al., 2003), while no evidence of logographic analysis is detectable (Wimmer & Hummer, 1990). Alphabetic decoding is reflected in the influence of stimulus length over reading times. The present results showed that length was a powerful factor in modulating performance at early stages of learning and became progressively less critical at later stages in the case of words. In the case of non-words, the “pure” effect of length (once the global factor in processing was taken into account) was large but did not change as a function of grade. The developmental change in performance for non-words present in the raw data is not specific but is due to the changes in global factor processing. Overall, in the absence of lexical information alphabetic decoding is required independent of reading experience.

In languages with shallow orthographies, including Italian, the correct decoding of nearly all words can be achieved by using the grapheme-to-phoneme routine. Based on this, it was originally proposed that reading is actually achieved only by the non-lexical procedure (Frost, Katz, & Bentin, 1987). However, evidence indicates that proficient Italian readers use the lexical route, which allows more rapid and fluent word decoding (e.g., Bates et al., 2001).

The present results offer some information on the development of the lexical procedure. At the end of first grade children read words faster and more accurately than non-words indicating activation of the lexicon at early stages in the learning

process. Notably, the differences between words and non-words (lexicality effect) increased progressively with age, an effect attenuated by the presence of under-additivity. In other words, the difference in reading words and non-words evident in the raw data was larger if the global improvement in information processing common to all orthographic stimuli was taken into account.

In first grade the children's lexicon was limited to high frequency words, and low frequency words were read as fast as non-words. In third grade a frequency difference was detected indicating the progressive improvement of the lexical procedure. By this age the difference between words and non-words was present also for low-frequency words. In the subsequent grades, with increased exposure to print, the differences between high and low frequency words diminished, presumably indicating the progressive consolidation of the low frequency items within the lexicon. In the present study stimuli were matched based on adult frequency. In future research it would be important to extend this matching to children's written word frequency.

One of the aims of this study was to make a methodological contribution to the problem of measuring reading development in the Italian language. For a number of reasons it is difficult to evaluate reading deficits in Italian. A classical cognitive approach is to compare error patterns across different stimulus materials and tasks. Applying this approach to languages such as English or French produced crucial breakthroughs in differentiating reading disturbance patterns. In regular orthographies such as Italian or German reading error analysis may be difficult in developmental cases. First, by fourth or fifth grade proficient readers' performances are nearly flawless and the error distribution seriously deviates from normality. The results clearly showed that the mean does not represent well the central tendency of the score distributions. If means (and SDs) cannot be reliably estimated, it becomes impossible to evaluate the relative distance of an individual dyslexic's performance from that of the reference group. Characteristically, dyslexics make few errors in absolute terms but may be entirely out of range when compared to normal readers. In other terms, even a 5% error level may place a sixth grader approximately 10 standard deviations below the norm. This makes it difficult to compare "simple" and "complex" conditions, such as words versus non-words, or high frequency versus low frequency words, independently of the specific dimension tested. In these cases only the complex condition alone can contribute variance to the measurement (a problem referred to as the cow-canary paradox; Capitani, Laiacona, Barbarotto, & Cossa, 1999).

Reading times do not present the problem of skewed reading accuracy distributions. However, even for reading time, applying standard parametric analyses to the data may alter evaluation of the influence of a given parameter. Thus, one should expect the effect of any variable influencing performance to be larger in individuals with longer reading times than in individuals with generally shorter reading times. This criticism applies to developmental studies on reading irrespective of the orthographic characteristics of the language. The procedure adopted in the present study, based on RAM (Faust et al., 1999), was sensitive in detecting developmental effects that were artifactually produced (i.e., changes in the length effect for non-words) or attenuated (i.e., changes in the lexicality effect) by variations in global processing ability.

It is conceivable that the procedure proposed here might be useful for evaluating the performance profile of dyslexics compared to proficient readers. In the former, one should expect a large difference in general ability as well as specific effects in some variables affecting word recognition. Disentangling these two components may contribute to understanding reading deficits in languages with both opaque and transparent orthographies.

Appendix

Experimental reading task items

Practice trial for non-words. reme, soate, cata, borla, same, custa, rilo, fuodi, valo, reito, focca, cino, sagno, nore, madre, pada, sevio, irna, lasta, dibò.

Short non-words. dete, borma, voro, bitto, lasa, sargo, dase, panto, gese, binca, anci, vunto, bine, nutto, goce, tucca, arlo, borte, mata, donso, luso, molpa, zato, vondo, moso, vasso, lona, parne, ceva, mosmo.

Long non-words. primosta, fanziane, mirigante, sonviglio, podiglia, bileggio, tegresto, strebafè, grocelso, stilappo, panchefa, bismacca, risumapa, oblimione, soribanto, gracetto, pralotto, montrillo, canfrosto, prigiosa, flenesta, lonaggio, tresilio, presemma, raspetto, varniale, nisciope, banziona, trelidente, alcelisto.

Practice trial for words. rete, soave, data, borsa, fame, costa, riso, fuori, vago, reato, bocca, cono, sogno, noce, madre, paga, serio, urna, pasta, divo.

High frequency short words. lira, mondo, modo, forma, zero, luogo, voce, conto, uomo, volta, casa, spesa, tipo, linea, piano, donna, fase, acqua, mano, posto, vita, grado, sede, mezzo, capo, ruolo, caso, fondo, crisi, campo.

High frequency long words. capacità, politica, sindacato, operazione, rapporto, famiglia, occasione, movimento, articolo, istituto, settimana, università, ministro, industria, obiettivo, bilancio, elezione, compagnia, intervento, processo, rispetto, decisione, situazione, campagna, successo, posizione, lavoratore, immagine, confronto, controllo.

Low frequency short words. olmo, botto, lena, cargo, sito, dosso, cava, balia, anca, dorso, tana, baule, lima, abate, nova, oncia, saga, bivio, fuso, suola, noce, sodio, agro, fermo, covo, zinco, duna, turca, duce, sonda.

Low frequency long words. frenesia, bimestre, antipatia, cervellone, addebito, prodigio, riassunto, telescopio, tonalità, barricata, alcolismo, rendiconto, aliscafo, retaggio, femminismo, noleggio, regresso, attivista, oreficeria, rodaggio, annuario, eversione, fotogramma, accumulo, umorista, stupidità, trampolino, asparago, panchina, marmellata.

References

- Ackerman, P. T., & Dykman, R. A. (1993). Phonological processes, confrontation naming, and immediate memory in dyslexia. *Journal of Learning Disabilities*, 26, 597–609.
- Arduino, L. S., & Burani, C. (2004). Neighborhood effects on nonword visual processing in a language with a shallow orthography. *Journal of Psycholinguistic Research*, 33, 75–95.

- Bamber, D. (1979). State trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology*, *19*, 137–181.
- Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, and Computers*, *34*, 424–434.
- Bates, E., Burani, C., D'Amico, S., & Barca, L. (2001). Word reading and picture naming in Italian. *Memory and Cognition*, *29*, 986–999.
- Bonifacci, P., & Snowling, M. J. (2008). Speed of processing and reading disability: A cross-linguistic investigation of dyslexia and borderline intellectual functioning. *Cognition*, *107*, 999–1017.
- Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpholexical reading develop in readers of a shallow orthography? *Brain and Language*, *81*, 568–586.
- Capitani, E., Laiacona, M., Barbarotto, R., & Cossa, F. M. (1999). How can we evaluate interference in attentional tests? A study based on bi-variate non-parametric tolerance limits. *Journal of Clinical and Experimental Neuropsychology*, *21*, 216–228.
- Coltheart, M., & Leahy, J. (1996). Assessment of lexical and nonlexical reading abilities in children: Some normative data. *Australian Journal of Psychology*, *48*, 136–140.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual-route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256.
- Cossu, G. (1999). The acquisition of Italian orthography. In M. Harris & G. Hatano (Eds.), *Learning to read and write. A cross-linguistic perspective*. Cambridge, U.K: Cambridge University Press.
- Cossu, G., Gugliotta, M., & Marshall, J. C. (1995). Acquisition of reading and written spelling in a transparent orthography: Two non parallel processes?. *Reading and Writing*, *7*, 9–22.
- De Brauwier, J. D., Verguts, T., & Fias, W. (2006). The representation of multiplication facts: Developmental changes in the problem size, five and tie effects. *Journal of Experimental Child Psychology*, *94*, 43–56.
- De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudo-words: An eye movement study of developmental dyslexia. *Brain and Language*, *80*, 617–626.
- De Luca, M., Di Pace, E., Judica, A., Spinelli, D., & Zoccolotti, P. (1999). Eye movement patterns in linguistic and non-linguistic tasks in developmental surface dyslexia. *Neuropsychologia*, *37*, 1407–1420.
- Di Filippo, G., De Luca, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2006). Lexicality and stimulus length effects in Italian dyslexics: Role of overadditivity effect. *Child Neuropsychology*, *12*, 141–149.
- Ellis, A. W. (1985). The production of spoken words: A cognitive neuropsychological perspective. In A. W. Ellis (Ed.), *Progress in the psychology of language* (Vol. 2). Hillsdale, NJ: Erlbaum.
- Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implications for group differences in response latency. *Psychological Bulletin*, *125*, 777–799.
- Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographic depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, *13*, 104–115.
- Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin and Review*, *7*, 185–207.
- IBM Italia. (1989). *VELI, Vocabolario Elettronico della Lingua Italiana*. Milan: IBM.
- Job, R., Peressotti, F., & Cusinato, A. (1998). Lexical effects in naming pseudowords in shallow orthographies: Further empirical data. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 622–630.
- Judica, A., De Luca, M., Spinelli, D., & Zoccolotti, P. (2002). Training of developmental surface dyslexia improves reading performance and shortens eye fixation duration in reading. *Neuropsychological Rehabilitation*, *12*, 177–197.
- Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin*, *109*, 490–501.
- Kail, R., & Hall, L. K. (1994). Processing speed, naming speed, and reading. *Developmental Psychology*, *30*, 949–954.
- Kail, R., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*, 199–225.
- Landerl, K. (2001). Word recognition deficits in German: More evidence from a representative sample. *Dyslexia*, *7*, 183–196.
- Logan, G. D. (1992). Shapes of reaction times distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *18*, 883–914.

- Nazir, T. A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., & Frost, R. (2004). Reading habits, perceptual learning, and recognition of printed words. *Brain and Language*, *88*, 294–311.
- Newell, A., & Rosenbloom, P. S. (1981). Mechanisms of skill acquisition and the law of practice. In J. R. Anderson (Ed.), *Cognitive skills and their acquisition*. Hillsdale: Lawrence Erlbaum Associates.
- Orsolini, M., Fanari, R., Tosi, V., De Nigris, B., & Carrieri, R. (2006). From phonological recoding to lexical reading: A longitudinal study on reading development in Italian. *Language and Cognitive Processes*, *21*, 576–607.
- Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly*, *27*, 29–52.
- Salthouse, T. A., & Hedden, T. (2002). Interpreting reaction time measures in between-group comparisons. *Journal of Clinical and Experimental Neuropsychology*, *24*, 858–872.
- Sartori, G., Job, R., & Tressoldi, P. (1995). *Batteria per la valutazione della dislessia e della disortografia evolutiva*. Florence: Organizzazioni Speciali.
- Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.
- Spinelli, D., De Luca, M., Di Filippo, G., Mancini, M., Martelli, M., & Zoccolotti, P. (2005). Length effect in word naming latencies: Role of reading experience and reading deficit. *Developmental Neuropsychology*, *27*, 217–235.
- Spring, C., & Davis, J. M. (1988). Relations of digit naming speed with free components of reading. *Applied Psycholinguistics*, *9*, 315–334.
- Tressoldi, P. E. (1996). L'evoluzione della lettura e della scrittura dalla 2a elementare alla 3a media. Dati per un modello di sviluppo e per la diagnosi dei disturbi specifici. *Età evolutiva*, *56*, 43–55.
- Trueswell, J., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285–318.
- Van IJzendoorn, M. H., & Bus, A. G. (1994). Meta-analytic confirmation of the nonword reading deficit in developmental dyslexia. *Reading Research Quarterly*, *29*, 266–275.
- Weekes, B. S. (1997). Differential effects of number of letters on word and nonword naming latency. *The Quarterly Journal of Experimental Psychology*, *50A*, 439–456.
- Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics*, *14*, 1–33.
- Wimmer, H., & Hummer, P. (1990). How German speaking first graders read and spell: Doubt on the importance of the logographic stage. *Applied Psycholinguistics*, *11*, 349–368.
- Wimmer, H., & Mayringer, H. (2002). Dysfluent reading in the absence of spelling difficulties: A specific disability in regular orthographies. *Journal of Educational Psychology*, *94*, 272–277.
- Wimmer, H., Mayringer, H., & Landerl, K. (1998). Poor reading: A deficit in skill-automatization or a phonological deficit? *Scientific Studies of Reading*, *2*, 321–340.
- Zoccolotti, P., De Luca, M., Di Pace, E., Gasperini, F., Judica, A., & Spinelli, D. (2005). Word length effect in early reading and in developmental dyslexia. *Brain and Language*, *93*, 369–373.
- Zoccolotti, P., De Luca, M., Di Pace, E., Judica, A., Orlandi, M., & Spinelli, D. (1999). Markers of developmental surface dyslexia in a language (Italian) with high grapheme-phoneme correspondence. *Applied Psycholinguistics*, *20*, 191–216.
- Zoccolotti, P., De Luca, M., Judica, A., & Spinelli, D. (2008). Isolating global and specific factors in developmental dyslexia: A study based on the rate and amount model (RAM). *Experimental Brain Research*, *186*, 551–560.