Lexical and sub-lexical effects on accuracy, reaction time and response duration: impaired and typical word and pseudoword reading in a transparent orthography

Robert Davies · Javier Rodríguez-Ferreiro · Paz Suárez · Fernando Cuetos

Published online: 06 June 2012
© Springer Science+Business Media B.V. 2012

Abstract In an opaque orthography like English, phonological coding errors are a prominent feature of dyslexia. In a transparent orthography like Spanish, reading difficulties are characterized by slower reading speed rather than reduced accuracy. In previous research, the reading speed deficit was revealed by asking children to read lists of words. However, speed in list reading sums the time required to prepare an utterance, reaction time (RT), with the time required to say it, response duration (RD). Thus, the dyslexic speed deficit in transparent orthographies could be driven by slow RTs, by slow RDs, or both. The distinction is especially important if developmental readers rely on phonological coding to achieve lexical access because the whole word would have to be encoded before it could be identified. However, while the factors that affect reading RT have been extensively investigated, no attention has been paid to RD. We studied the performance of typically developing and dyslexic Spanish children in an oral reading task. We analysed the impact of word frequency and length on reading accuracy, RT, and RD. We found that accuracy, RT, and RD were affected by word frequency and length for both control and dyslexic readers. We also observed interactions between effects of
reader group—dyslexic, typically developing (TD) younger or TD older readers—and effects of lexicality, frequency, and word length. Our results show that children are capable of reading aloud using lexical and sub-lexical coding processes in a transparent orthography.

**Keywords**  Reading · Development · Dyslexia · Transparent · Orthography · Speed

**Introduction**

Dyslexic reading in transparent orthographies is characterized by speed not accuracy problems, in contrast to dyslexic reading in opaque orthographies where accuracy problems are more prominent (Wimmer, 1993). The observation that dyslexic reading in transparent orthographies is not characterized by accuracy problems is explained by assuming the same cause used to explain, also, the observation that reading acquisition happens more quickly in transparent than in opaque orthographies: spelling-sound consistency. Numerous studies have shown that where spelling-sound mappings are highly consistent, as in transparent orthographies like Spanish or Italian, the typical course of reading development ends in accurate foundation level reading within the first school year (Seymour, Aro, & Erskine, 2003): a much faster rate than that seen in English (see also consistent results obtained in cross-linguistic comparisons reported by Bruck, Genesee, & Caravolas 1997; Ellis & Hooper, 2001; Frith, Wimmer, & Landerl, 1998; Goswami, Gombert, & de Barrera, 1998; Wimmer & Goswami, 1994). The explanation for the difference in acquisition rate, as for the apparent limitation of reading difficulties to speed problems, is that transparent orthographies permit accurate phonological coding just so long as grapheme-phoneme correspondences are mastered. In comparison, in an opaque orthography like English, mastering grapheme-phoneme correspondences does not afford accurate phonological coding, due to the variation in, especially, the pronunciation of vowel letters.

Wimmer’s (1993) observation that dyslexics reading in German were distinguished by slower reading speed not by lesser reading accuracy, in comparison to controls, has been repeated in a number of studies (e.g., Wimmer & Mayringer, 2002), and in other languages (Dutch; Yap & Van der Leij, 1993; Greek; Porpodas, 1999; Italian; Zoccolotti, De Luca, Di Pace, Judica, Orlandi, & Spinelli, 1999; Spanish; Davies, Cuetos, & Glez-Seijas, 2007; Gonzalez & Valle, 2000). Wimmer and colleagues (Bergman & Wimmer, 2008; Wimmer et al., 2010) have suggested that slowed dyslexic reading in a transparent orthography like German results from the slow activation of phonology.

In the Dual-Route Cascaded (DRC) model of reading aloud (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001), following presentation of a letter string, phonological coding takes place through two routes, a lexical route involving the activation of lexical orthographic and lexical phonological units, and a sub-lexical route involving the activation of grapheme-phoneme correspondences (GPCs). The phoneme units representing the sound of a word are activated jointly following the transmission of sufficient activation from activated lexical phonological or GPC
representations. Phonemes are activated in parallel across all phoneme positions by
the lexical route but are activated serially, left-to-right, phoneme-by-phoneme, by
the sub-lexical route. There are alternative conceptions of phonological coding, in
particular, accounts that assume reading processes involve connectionist systems of
distributed representations without lexical units as such (Harm & Seidenberg, 2004;
Seidenberg & McClelland, 1989). However, we will frame our discussion of reading
in transparent orthographies with respect to the DRC because current debates
usually continue to distinguish lexical and sub-lexical reading, and because our data
do not speak to the hypothesized differences between connectionist and DRC
accounts.

In terms of the DRC (Coltheart et al., 2001), Bergman and Wimmer blamed
slowed reading on both (1) slower lexical phonological coding, that is, slower
activation of phonological word representations given orthographic word activation,
and (2) slower grapheme to phoneme conversion in the sub-lexical route. The speed
deficit account of dyslexia in transparent orthographies ties slower reading to slower
phonological coding but there is an omission from general theoretical accounts of
reading which may pose a critical limitation for our understanding of reading in
transparent orthographies. Current computational models of reading aloud (Colt-
heart et al., 2001; also Harm & Seidenberg, 2004; Seidenberg & McClelland, 1989)
are focused on simulating the effects of factors like lexical frequency on reaction
times (RTs) recorded in studies of reading aloud. For example, in the DRC, naming
latencies are equated with the moment when all a word’s phonemes are activated
above threshold, when phonological coding has been completed. The simulations
that are reported concern how factors like frequency affect phoneme activation as an
effect on response latency. The model is silent on the factors that affect response
duration. However, as we will argue, it is likely that we should also consider
response duration (RD) and the factors that influence how long it takes to say a
word.

Consideration of the factors that affect RD may be especially important to
understand reading in a transparent orthography. This is because, given the rule
governed nature of grapheme-phoneme correspondences in a transparent orthog-
raphy like Spanish, a reader could start to say a word even before they have
completely prepared the phonological code for the word. The phonological code for
the remainder of the word could be left to be accurately computed online because
pronunciation in a transparent orthography is not subject to the irregularities that
obtain in English. This line of thinking opens the possibility that one should be able
to detect effects of reader skill (comparing dyslexic and typically developing
children) and effects of the lexical or sub-lexical properties of words on RD as well
as RT in reading aloud in Spanish.

The original observations showing the speed deficit in various languages relied
on the list reading task, in which children are asked to read words printed as a list on
the page (e.g., in Spanish, Davies et al., 2007; in German, Wimmer, 1993). This task
has the advantage that it mimics natural reading. However, list reading times sum
both the time required for response preparation (RT) and the time required to
produce the response (RD). While there is substantial evidence concerning the
factors that affect reading RTs, including those recorded for children, there is no

\[ \text{Springer} \]
evidence on the factors that influence RDs in developmental reading in any language, let alone in Spanish. The first aim of our investigation, therefore, was to examine the effects of item attributes on reading RTs and RDs.

What should our expectations be? There is ample evidence to suggest that RT is influenced by lexical and sub-lexical factors in children’s reading in transparent orthographies. Though most of the relevant data come from work in Italian, the results should adequately delimit our expectations concerning Spanish, given the similarity of the Italian and Spanish orthographies, as well as the congruence of indications from other measures of reading performance in Spanish. Burani, Marcolini, and Stella (2002) found that word frequency and length affected the reading of 8- and 10-year old Italian children in lexical decision and word naming tasks (see, for consistent results on frequency in an older developmental sample, Barca, Burani, Di Filippo, & Zoccolotti, 2006). For word naming, Burani et al. (2002) observed that while the frequency effect did not vary by age, the length effect was found to be smaller for older children. The decrease in length effect with increasing age was also observed by Zoccolotti et al. (2005); see, also, De Luca, Barca, Burani, & Zoccolotti, 2008; Zoccolotti et al., 1999), who tested dyslexic (mean age = 8.4 years) and typically developing readers (aged 6.8–8.7 years).

The frequency effect is argued to reflect the functioning of a lexical route, with frequency scaling the activation function for lexical units in the dual-route model of reading (Coltheart et al., 2001; see Seidenberg & McClelland, 1989, and Adelman, Brown, & Quesada, 2006, for alternative accounts of the frequency effect). Zoccolotti et al. (2005) proposed that the Burani et al. (2002) findings on frequency suggest that a lexical reading route is already functioning by the third grade. The frequency effect appears to be smaller in more skilled readers, when comparing dyslexic with typically developing children (Barca et al., 2006). However, frequency effects continue to be observed in the reading of Italian-speaking adults (Barca, Burani, & Arduino, 2002; Burani, Arduino, & Barca, 2007), suggesting that a lexical reading route is established fairly early and then remains in operation through the lifetime of a reader in a transparent orthography.

The length effect has been argued to reflect the serial phonological coding of graphemes by the sub-lexical reading route (Coltheart et al., 2001; Weekes, 1997). Zoccolotti et al. (2005) suggested that because the length effect diminished substantially between first and second grade readers, lexical coding largely displaces sub-lexical coding in determining RTs in word naming. In addition, it is suggested that because the length effect in dyslexic readers resembled that seen in younger controls, sub-lexical coding may remain a dominant strategy for children with reading difficulties.

Findings in Spanish indicate similar effects on other performance measures compared to those reported for Italian reading RT. Defior, Justicia, and Martos (1996) reported effects of both word frequency and length on reading accuracy in typically developing and dyslexic Spanish-speaking children aged 6–12 years. An effect of frequency on accuracy was also noted for the reading of Spanish dyslexic children (aged 9–10 years) by Rodrigo López and Jiménez González (1999). In addition, Davies et al. (2007) reported effects of word frequency, length, and orthographic neighbourhood size (N-size) on word list reading times. These effects
are largely replicated in analyses of Spanish-speaking adults’ reading RTs (Cuetos & Barbón, 2006), indicating the development and continuing use of lexical and sub-lexical reading routes in Spanish.

We are not aware of any study of the factors that influence response duration in developmental reading in any language. However, some previous research does offer clues as to what effects can be expected. Firstly, it appears that the duration of whole word spoken responses are shortened if target words are preceded by semantically related primes in reading (Balota, Boland, & Shields, 1989; see also Kawamoto, Goeltz, Agbayani, & Groel, 1998). Secondly, it has been reported that the duration of a spoken word response in a lexical decision task is shorter for high compared to low frequency words (Balota & Abrams, 1995). To be clear, in Balota and Abrams’s (1995) study, stimulus words varied in frequency but participants were required to say a single word (e.g., “normal” in their third experiment) in response to a stimulus, and it is the duration of that response that varied with stimulus frequency. Nevertheless, taken together, these findings contradict the separation between response preparation and response execution that is usual in reading models.

The observation of a frequency effect on response duration is inconsistent with the assumption that the frequency effect is confined to a locus in lexical activation prior to the events surrounding response execution (Balota & Abrams, 1995). This is because the observation suggests that lexical frequency continues to influence response execution even after response onset. Similarly, the priming effect on response duration reported by Kawamoto et al. (1998) is inconsistent with the assumption that a response is initiated (as in the DRC, Coltheart et al., 2001) when phonological coding has been completed. This is because priming should not influence pronunciation if pronunciation is simply the articulation of a phonological code, a set of phonemes simultaneously and entirely activated at response onset. Results like those we have reviewed imply a cascaded flow of activation from stages that are dominated by access to and processing of lexical representations to later stages that are focused on response production (Balota & Yap, 2006). For our study, previous research suggested we should find word frequency and length effects in word naming in RDs as well as in RTs.

The present study

In the present study we examined the factors that influence oral reading performance in Spanish-speaking children. We presented words or pseudowords in discrete trials rather than word lists. For the first time, to the best of our knowledge, both reaction time and response duration measures of children’s reading were subject to analysis. Our analyses were designed to estimate the effect of item attributes on RT, RD, and accuracy, including the effects of lexicality, length and lexical frequency. We examined reading in Spanish dyslexic children and in two groups of typically developing readers, one matched on chronological age to the dyslexic readers, and the other composed of younger readers (in the first grade of primary school).
Method

Subjects

We recruited nine dyslexic children (7 males, 2 females) and 20 children of typically developing reading ability (12 females, 8 male) from schools in the city of Oviedo and surrounding areas. Participant characteristics are summarized in Table 1. The nine dyslexic children attended the speech and language therapy clinic managed by one of the authors (PS), where they were receiving remedial instruction in reading or spelling averaging 1 h per week. The 20 controls were recruited from public schools in Oviedo. For these children, selection to the study was based on teacher ratings of typical literacy development. No child was included if she was known to have received a diagnosis of neurological abnormality.

We collected information on children’s intelligence using the Spanish version of the Wechsler Intelligence Scale for Children-R (WISC-R; Wechsler, 1974), the Escala de Inteligencia para Niños-Revision (EIWN-R; Wechsler, 1982). No child scored <80 on the performance scale of the EIWN-R.

We collected information on children’s reading level using the word and pseudoword sub-tests of the PROLEC-R standardized literacy test (Cuetos, Rodríguez, Ruano, & Arribas, 2007). In the PROLEC-R reading sub-test, a list of

| Test                                | Participant group |  |  |  |  |
|-------------------------------------|-------------------|---|---|---|---|
|                                     | Dyslexic n = 9   | Older TD n = 10 | Younger TD n = 10 |
|                                     | Mean (SD)        | Mean (SD)       | Mean (SD)       |
| Age (months, rounded to nearest month) | 132 (12)         | 124 (11)        | 77 (6)          |
| EIWN-R General intelligence        | 107.1 (16.3)     | 112.0 (11.3)    | 100.9 (14.8)    |
| EIWN-R Verbal intelligence         | 108.7 (24.6)     | 111.8 (13.0)    | 102.0 (12.0)    |
| EIWN-R Performance intelligence    | 101.3 (9.4)      | 107.9 (10.6)    | 99.4 (15.7)     |
| PROLEC-R Word reading accuracy (total correct/40) | 29.7 (5.0) | 39.6 (0.5) | 38.9 (0.9) |
| PROLEC-R Word reading speed (s)     | 81.6 (27.6)      | 25.4 (2.9)      | 51.0 (12.3)     |
| PROLEC-R Non-word reading accuracy (total correct/40) | 25.6 (6.2) | 37.5 (0.5) | 37.0 (0.9) |
| PROLEC-R Non-word reading speed (s) | 105.8 (42.9)     | 43.6 (7.1)      | 70.7 (9.0)      |
40 printed words, varying in length, frequency, and syllable structure, and a list of 40 pseudowords, varying in length and syllable structure, is presented for reading aloud. Reading accuracy and speed are measured for each list. The standardization sample for the PROLEC-R consisted of about 150 pupils (about 50% of each gender) for each year of primary school, with samples drawn from schools across Spain.

The criteria for inclusion in the participant sample were as follows. Children were included in the control group of Typically Developing (TD) readers if their PROLEC-R word or pseudoword reading accuracy scores fell within a range equal to the mean plus or minus 2 SD for the standardization sample for their school year. Dyslexic children were included if their accuracy scores were less than or equal to the mean minus 2 SD. The dyslexic children in our study were found to read words aloud at a level of accuracy 2 SD below the mean for the standardization sample for their school year. All but one of the dyslexic children read pseudowords at a level of accuracy below the mean minus 2 SD for their age; the exception child’s pseudoword accuracy was 1.5 SD below the mean for his age. We note, also, that all but one of the dyslexic children took longer to read the word list than the mean time for the standardization sample plus 2 SD; the exception child’s word reading speed was 1 SD greater than the mean time for his age. All but three of the dyslexic children took longer to read the pseudoword list than the mean time for the standardization sample plus 2 SD; the exception children’s pseudoword reading times were all at least 1 SD greater than the mean time for his age. In sum, the dyslexic children read less accurately and more slowly than typically developing readers of the same age.

The TD children were split into two groups of 10 children each: a chronological age match (older TD) group and a younger group (younger TD). The older TD children were matched to the dyslexic children on age and intelligence (EIWN-R, general, verbal, and performance indices), \( p > 0.10 \) for all comparisons. However, the older TD children were significantly more accurate and faster in PROLEC-R word and pseudoword reading \( (p < 0.001 \) for each comparison) than the dyslexic children. The younger TD controls were in the 1st year of primary school, compared to the dyslexic and older TD children, who were mostly in the fifth or 6th years. The younger TD children were significantly younger than the dyslexic or the older TD children \( (p < 0.001) \) but did not differ from either of the other groups on intelligence \( (all \ ps > 0.05) \). The younger TD children were significantly faster and more accurate in word and pseudoword reading than the dyslexic children \( (p < 0.001 \) for each comparison).

The TD younger cannot be said to present a match on reading ability to the dyslexic children. However, we could not have found ability matches to our sample of dyslexic children without testing pre-school readers, entailing an unacceptable confound between comparison groups and differences in experience of instruction. It was better, in our view, to bring younger TD readers into the sample in order to afford a comparison between older dyslexic and younger TD readers that would align reading performance in the dyslexic group with the performance of typical readers in the 1st year of primary school.
Materials

We selected 80 words and constructed 80 pseudowords for use in our study. The stimulus words were short words (10 four-letter and 10 five-letter words) or long words (10 six-letter and 10 seven-letter words) and were of high or low frequency, with 20 words selected for each cell in a fully factorial, length by frequency, design. Four or five letter words and pseudowords were two syllables in length, six and seven letter items were three syllables in length. The stimulus pseudowords were constructed from a set of base-words matched to the stimulus words on frequency and length. Words and base-words consisted of monomorphemic nouns, verbs and adjectives; most items were nouns. Controlling the selection of base-words was done to ensure that the pseudowords would be as similar to the words as possible and thus that responses would require lexical access for words. Matching words and base-words on frequency and word class was expected to ensure high similarity because pseudoword construction was by single letter substitution.

The estimates of lexical frequency values for words and for the base-words corresponding to pseudoword stimuli were taken from an analysis of texts given to children for reading in school (Martínez & García, 2004). We used as our index of lexical frequency the estimates of the accumulated frequency of occurrence of words experienced in print up to the sixth grade, when children are 11–12 years old in Spain. Several other item attributes were controlled across the four manipulated cells, including the imageability and Age-of-Acquisition (AoA) of words, as well as the orthographic neighbourhood size of both words and pseudowords. Imageability data were gathered from the Valle-Arroyo (2001) norms which include average values obtained from the ratings of 135 adults on a 7-point scale. We collected AoA ratings for words (on a 7-point scale) from 20 adult volunteers, undergraduates attending the University of Oviedo. Estimates of the orthographic neighbourhood size (N-size) of words and pseudowords computed using the BuscaPalabras database (Davis & Perea, 2005). In BuscaPalabras, neighbourhood size is determined by counting the number of words that can be formed by substituting a single letter at any of the letter positions within a word or pseudoword string, counting as neighbours only entries in the LEXESP adult frequency corpus (Sebastián, Martí, Carreiras, & Cuetos, 2000). Characteristics of the words presented are summarized in Table 2.

Words and pseudowords did not differ on length nor on N-size ($F(1,158) = 0.89, p = 0.35$). Words of different length conditions (two or three syllables) did not differ on frequency ($F(1,78) = 0.68, p = 0.411$), AoA ($F(1,78) = 1.3, p = 0.256$), imageability ($F(1,78) = 0.857, p = 0.357$) or N-size ($F(1,78) = 0.51, p = 0.822$).

To improve the sensitivity of our analyses and militate against the effect of skew in variable distributions, we transformed word frequency measures to log 10(frequency +1).

Apparatus and procedure

The children were tested in two sessions, in which standardized reading or intelligence tasks were also administered. Children were instructed to read items as quickly and as accurately as possible.
Stimuli were presented and responses recorded using DMDX (Forster & Forster, 2003) on a Windows XP laptop. Stimuli were presented in Arial 10-point type. Children were seated at 30 cm from the display screen. The stimuli subtended an average of 4.29° of visual angle at that distance.

Each item was presented twice to each child, once in each of the two test sessions. This was done to allow for a check that the effects of word attributes like item length or lexical frequency might be modulated by the presence or absence of pseudowords in the stimulus list. We manipulated list composition by presenting pure or mixed lists of words and pseudowords. Children therefore saw each critical word or pseudoword stimulus once in pure lists of word or pseudoword trials only and once in mixed lists of word and pseudoword trials. As we did not find interactions between the potential effects of word attributes and of list composition, in analyses that are not reported, we do not discuss the manipulation of list composition or block mix further.

Four experimental programs were created. Within each experimental program, 80 items were presented in two blocks of 40 trials each. Items were pseudorandomly assigned to the two blocks such that stimuli high or low in (word or base-word) frequency and length were equally likely to be represented in each block, for each program. The order of administration of the four experimental programs was counterbalanced across participants using a Latin Square design. Then, for each participant, in each test session, the order of presentation of trials within each block was randomized, as was the order of the blocks themselves.

Two programs were administered in each test session. A short break was permitted between each block. Each experimental program began with six practice trials. Practice items were selected to have similar characteristics to the critical items. The experimenter answered any questions the participants had following the practice but no further feedback was given thereafter.

An experimental trial had the following sequence of events (timed with respect to screen refreshes): firstly, there appeared a blank screen for 512 ms (event timing is coordinated by DMDX with respect to screen refresh rates or ticks). Then, a black
asterisk was presented in the centre of a grey field screen for 512 ms. The stimulus replaced the asterisk and was presented for 3,072 ms. Responses made during the three second response intervals were recorded digitally to hard disk. Each test session lasted 30 min in total.

Results

Data extraction

We recorded a total of 9,280 responses. We analysed sound spectrograms of the recorded responses, using the CheckVocal application (Protopapas, 2007), to extract accuracy, RT and RD. For our analysis of accuracy, we analysed the accuracy of all responses. For our analysis of RT and RD, we analysed just correct responses, excluding all observations corresponding to incorrect responses. Of the 728 errors (8% of the total): 187 were null responses (participants took longer than 3,000 ms to respond); 158 were word substitution errors; and 383 were pseudoword errors. We analysed the RTs and RDs of only the 8,522 correct responses remaining after these exclusions.

Analysis strategy

Our investigation had a repeated measures design, with the same items presented to different participants, requiring the use of mixed-effects modelling to accurately estimate effects of theoretical interest while properly accounting for error variance (Baayen, 2008; Baayen, Davidson, & Bates, 2008). Mixed-effects models estimate both fixed effects, that is, replicable effects of theoretical interest, for example, of word frequency or reader ability, and random effects, for example, unexplained effects due to random variation between items or between participants. Mixed-effects modelling supports accurate estimation of interactions between reader type and item attribute effects without running the risk of spurious over-additivity due to differences between readers or between groups of readers in average RT or RD (see Faust, Balota, Spieler, & Ferraro, 1999).

Analysis of accuracy

A summary is presented in Table 3 of the percentage accuracy of responses made by participants in each group in each condition. It can be seen that the level of accuracy in word naming is very high (comparable to that seen e.g., in Seymour et al., 2003), above 80% in all conditions for all children, largely above 90% for the control group TD children.

We analysed the accuracy of responses to words or pseudowords using Generalized-Mixed effects Modelling (GLMM), that is, the mixed-effects extension of multiple logistic regression (Baayen, 2008) in which the aim is to estimate the log odds that a response would be accurate given a set of predictors. Random (intercepts) effects of both subjects and items were required for all reported models.
We analysed, firstly, the accuracy of responses to both words and pseudowords. We examined the effects on accuracy of: reader group (comparing dyslexic, younger TD, and older TD children); word and pseudoword length in letters (4–7 letters); and lexicality (comparing responses to words or to pseudowords); as well as the effects of interactions between the effect of reader group and the effects of length or lexicality. A summary of the results is presented in Table 4. Note that in the tabled model summary, z values indicate that a response is more likely to be correct for unit change in the corresponding factor whereas negative z values indicate that a response is less likely to be correct for unit change.

We found that older but not younger TD children were significantly more likely to produce correct responses to words or to pseudowords than dyslexic children. Shorter items (words and pseudowords) were more likely to be read correctly than longer items. Correct responses were more likely to be elicited by words than by pseudowords. We also found an interaction such that the lexicality effect on accuracy was greater for older TD children compared to dyslexic readers.

We conducted a separate analysis of children’s responses to words alone (Table 4). Lexical frequency was now included as a predictor in this analysis, and lexicality was dropped. The accuracy of responses to words was affected by frequency alone: more frequent words were more likely to elicit correct responses.

### Analysis of reaction times and response durations

We report, firstly, our analysis of RTs then our analysis of RDs. We transformed the raw RT and RD scores to log (base 10) values, before analysis, to militate against the influence of the marked skew associated with chronometric data (following Baayen, Feldman, & Schreuder, 2006). We analysed the RT and RD data using linear mixed-effects modelling. Random (intercepts) effects of both subjects and
Table 4  Generalized linear mixed effects models of log odds accuracy as well as linear mixed effects models of logRTs and logRDs of correct responses

|                          | Accuracy            | Words and non-words | Words only |
|--------------------------|---------------------|---------------------|------------|
|                          | z value             | z value             |            |
| (Intercept)              | 6.532***            | 3.559***            |            |
| Dyslexia versus Old TD   | 2.565*              | 0.719               |            |
| Dyslexia versus Young TD | −0.537              | 0.225               |            |
| Letters                  | −3.378***           | −1.857              |            |
| Lexicality               | 4.484***            |                     |            |
| Frequency                |                     | 3.279**             |            |
| Group (Dys. vs. Old)*letters | −0.834         | 0.229               |            |
| Group (Dys. vs. Young)*letters | 1.103          | 1.032               |            |
| Group (Dys. vs. Old)*lexicality | 3.38***         |                     |            |
| Group (Dys. vs. Young)*lexicality | 0.783          |                     |            |
| Group (Dys. vs. Old)*frequency |                  | 0.623               |            |
| Group (Dys. vs. Young)*frequency | −1.172         |                     |            |

|                          | Reaction times      | t value             | t value     |
|--------------------------|---------------------|---------------------|-------------|
| (Intercept)              | 104.83***           | 92.82***            |            |
| Dyslexia versus Old TD   | −4.27***            | −3.83***            |            |
| Dyslexia versus Young TD | −0.94               | −1.18               |            |
| Letters                  | 6.26***             | 4.87***             |            |
| Lexicality               |                     | −11.48***           |            |
| Frequency                |                     | −4.97***            |            |
| Group (Dys. vs. Old)*letters | −3.58***          | −3.01**             |            |
| Group (Dys. vs. Young)*letters | −0.47           | 1.19                |            |
| Group (Dys. vs. Old)*lexicality | 1.83            |                     |            |
| Group (Dys. vs. Young)*lexicality | 6.29***         |                     |            |
| Group (Dys. vs. Old)*frequency |                 | 1.67                |            |
| Group (Dys. vs. Young)*frequency | 1.76            |                     |            |

|                          | Response durations  | t value             | t value     |
|--------------------------|---------------------|---------------------|-------------|
| (Intercept)              | 93.47***            | 72.76***            |            |
| Dyslexia versus Old TD   | 1.07                | 0.43                |            |
| Dyslexia versus Young TD | 1                   | 1.41                |            |
| Letters                  | 21.62***            | 14***               |            |
| Lexicality               | −7.5***             |                     |            |
| Frequency                |                     | −3.13**             |            |
| Group (Dys. vs. Old)*letters | −7.66***          | −4.35***            |            |
| Group (Dys. vs. Young)*letters | 1.9             | 1.24                |            |
| Group (Dys. vs. Old)*lexicality | 5.75***         |                     |            |
| Group (Dys. vs. Young)*lexicality | 3.62***        |                     |            |
| Group (Dys. vs. Old)*frequency |                  | 2.59**              |            |
| Group (Dys. vs. Young)*frequency | −0.08          |                     |            |

*** if $p < 0.001$; ** if $p < 0.01$
items were required for all reported models. A summary of average RTs and RDs is reported in Table 3.

**Analyses of RTs**

Following the same procedure used in the analysis of accuracy data, we first analysed the latency of all 8,522 correct responses to words or pseudowords considered together. For models of chronometric data, where $t$ statistics associated with effects are positive this indicates that, in comparison to the overall average, the logRT or logRD increased for unit increase in the predictor variable. Where $t$ statistics are negative this indicates that the logRT or logRD decreased for unit change in the predictor. Following Baayen (2008), we report Markov chain Monte Carlo (MCMC)-derived $p$ values for effects. We note that MCMC- and $t$-derived $p$ values for our analyses largely coincide but the former are slightly more conservative than the latter (Baayen et al., 2008).

A summary of the models of RT is presented in Table 4. We first analysed the effects on the RTs of responses to words and pseudowords of: reader group (comparing dyslexic, younger TD, and older TD children); word and pseudoword length in letters (4–7 letters); and lexicality (comparing responses to words or to pseudowords); as well as the effects of interactions between the effect of reader group and the effects of length or lexicality. We found that, compared to dyslexic readers, response latencies tended to be shorter if readers were older TD but not if they were younger TD children. A length effect was obtained, with short items eliciting shorter RTs in all the groups. Words elicited shorter latencies than pseudowords. In addition, we found interaction effects such that, compared to dyslexic readers, the effect of item length was less for older TD readers, while the effect of lexicality was less for younger TD readers.

We then analysed the latency of 4,384 correct responses to just words (Table 4). We found that, compared to dyslexic readers, response latencies to words tended to be shorter if readers were older TD but not if they were younger TD children. We found both word length and frequency effects: shorter and more frequent words tended to elicit shorter RTs. We again found an interaction between the effect of item (word) length and reader group such that the length effect was smaller for older TD readers compared to dyslexic readers.

**Analyses of RDs**

We first analysed the effects on the RDs of responses to words and pseudowords of: reader group (comparing dyslexic, younger TD, and older TD children); word and pseudoword length in letters (4–7 letters); and lexicality (comparing responses to words or to pseudowords); as well as the effects of interactions between the effect of reader group and the effects of length or lexicality (Table 4). We found no effect of reader group but we found significant effects of length and lexicality. Shorter items tended to elicit shorter durations. Words elicited shorter RDs than pseudowords. These effects were modulated by interactions involving reader group. The effect of item length on RDs was significantly smaller for older TD readers compared to
dyslexic readers. The effect of lexicality was significantly smaller for both control groups compared to dyslexic readers.

We then analysed the duration of correct responses to just words. We found no effect of reader group on response duration. We found both word length and frequency effects: shorter and more frequent words tended to elicit shorter RDs. We again found an interaction between the effect of word length and reader group such that the length effect was smaller for older TD readers compared to dyslexic readers. We also found an interaction between the effect of word frequency and reader group such that the frequency effect was smaller for older TD readers compared to dyslexic readers.

Discussion

Our study explored the factors that influence reading aloud in typically developing and dyslexic readers in a transparent orthography, Spanish. The presentation of words and pseudowords as well as the manipulation of length and frequency allowed us to evaluate the effects of lexicality and item attributes, addressing important questions about the processes involved in children's reading in transparent orthographies. We found effects of lexicality, length, and word frequency in all analyses, demonstrating the use of both a sub-lexical reading route whose function is reflected in the length effect and a lexical route reflected in the frequency effect. These results are consistent with previous observations on Italian children (e.g., Burani et al., 2002) and adults (e.g., Burani et al., 2007). The combination of evidence for a lexicality effect with evidence for a frequency effect conclusively demonstrates the impact of lexical knowledge on reading in a transparent orthography. The observation of these effects in RDs as well as RTs indicates that that lexical influence extends beyond the initiation of responses, consistent with the view that activation cascades through the phonological process in reading (Balota & Yap, 2006).

The length of words and pseudowords affected the accuracy, latency and duration of the responses of the three groups of participants, reflecting the use of serial non-phonological coding processes (Coltheart et al., 2001; Weekes, 1997) by readers of a transparent orthography. This finding is consistent with a great deal of evidence previously reported for Italian by Zoccolotti and colleagues (e.g., Zoccolotti et al., 2005), among others. However, the analysis of interactions between effects of length and reader group revealed that the pattern of effects was a bit more complicated. Word reading RTs and RDs were affected by letter length in all groups of participants, implying that not even the more experienced children could avoid the effect of the number of letters on the time needed to prepare or to pronounce words. (Note that item length measured in letters very precisely corresponds to item length in phonemes in Spanish). Older TD readers, however, showed a significantly smaller effect of word length on their reading RTs and RDs. The effect on word reading latencies of the interaction between reader group and length suggests that the use of sub-lexical information and sequential coding processes, when confronted
with words, is more prominent in these younger and less skilled readers than in the older and more skilled TD group.

The analyses of the influence of lexical frequency on reading performance also revealed interesting results. Word reading accuracy and RTs for all three groups were affected by lexical frequency. A frequency effect was also observed to influence the duration of the responses of the younger TD and dyslexia groups but did not affect the older TD group as much. Thus, frequency seems to be a strong predictor of reading speed in younger typical or older dyslexic developing readers of Spanish. These results demonstrate the use of lexical information in the preparation and execution of phonological coding in a transparent orthography. However, the frequency effect on RDs evidently decreased for older TD readers. We think that this indicates that the reading expertise of the older TD participants tended to be associated with a greater coding efficiency consequent on learning from accumulated reading experience. A diminishing frequency effect with increasing experience is reported by Zevin and Seidenberg (2002) in a connectionist simulation of reading development, and is attributed to the gradual optimization of connection weights, a process of adaptation to experience that tends to narrow the space in which the frequency effect can appear.

When considering the influence of lexical frequency on RTs in developmental reading in Italian, in comparisons of younger and older TD children, the evidence is inconsistent. Whereas Barca et al. (2007) reported a larger frequency effect for younger compared to older children, Burani et al. (2002) observed no significant interaction between the effects of frequency and age group or relative reading skills. Those results were obtained with respect to RT. Our results suggest that other dependent measures, like RD, might also be sensitive to the impact of reader age or ability on the frequency effect in reading.

In sum, just as has been found for healthy adults reading in transparent orthographies, for children reading in Spanish, performance depends on a mixture of lexical and sub-lexical knowledge. This is not altogether surprising if one turns from the emphasis on reading accuracy that has perhaps stemmed from the influence due to the concerns of studying reading in English, an inconsistent orthography (Share, 2008). It makes perfect sense if one focuses on reading fluency which, here, is measured with respect both to the time it takes to prepare an oral response (RT) and the time it takes to say it (RD). Our work extends the evidence base by showing the subtlety of lexical and sub-lexical effects, in particular, indicating how they are modulated by differences between readers.

The RD observations have a wider, quite practical, implication if lexical access to meaning proceeds, at least, initially in development, through phonological coding of printed words (see Harm & Seidenberg, 2004, for a review of relevant evidence). This will apply especially to dyslexic children reading in transparent orthographies. Our results show that dyslexic children take longer to prepare and to say spoken responses to longer words while age-matched peers are less affected by word length. Dyslexic children also take longer to prepare and to say spoken responses to new words (pseudowords compared to words) while age-matched peers, and also younger TD readers, are less affected by lexicality. We have shown that RDs as well as RTs are affected by reader skill and by word attributes like length or frequency.
This suggests that phonological coding processes may continue after response onset, that is, that the phonological specification for a word’s pronunciation may not be fully prepared at response onset. If the rate of access to meaning is limited by the rate at which readers can complete phonological coding, if also the rate of phonological coding must count both response preparation and response duration, then the implication of our findings is that dyslexic children will experience slower access to meaning. This slowness will be accentuated for longer words and for newly encountered words. It will be less characteristic of the performance of typically developing readers. Typically developing children will, as they grow older, become increasingly less vulnerable, by comparison with their dyslexic peers, to word length or to word novelty.

Acknowledgments This study was supported by grant MCI-PSI2009-09299 from the Spanish Government.

References

Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity not frequency determines word naming and lexical decision times. *Psychological Science, 17*, 814–823.

Baayen, R. H. (2008). *Analyzing linguistic data*. Cambridge, U.K.: Cambridge University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55*, 290–313.

Balota, D. A., & Abrams, R. A. (1995). Mental chronometry: Beyond onset latencies in the lexical decision task. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21*, 1289–1302.

Balota, D. A., Boland, J. E., & Shields, L. W. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory and Language, 28*, 14–36.

Balota, D. A., & Yap, M. J. (2006). Attentional control and flexible lexical processing: Explorations of the magic moment of word recognition. In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 229–258). Hove, U.K.: Psychology Press.

Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behaviour research methods, instruments and computers, 34*, 424–434.

Barca, L., Burani, C., di Filippo, G., & Zoccolotti, P. (2006). Italian developmental dyslexic and proficient readers: Where are the differences? *Brain and Language, 98*, 347–351.

Bergmann, J., & Wimmer, H. (2008). A dual-route perspective on poor reading in a regular orthography: Evidence from phonological and orthographic lexical decisions. *Cognitive Neuropsychology, 25*, 653–676.

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

Burani, C., Arduino, L. S., & Barca, L. (2007). Frequency, not age of acquisition, affects Italian word naming. *European Journal of Cognitive Psychology, 19*, 828–866.

Burani, C., Marcolini, S., & Stella, G. (2002). How early does morpho-lexical reading develop in readers of a shallow orthography? *Brain and Language, 81*, 568–586.

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.

Cuetos, F., & Barbón, A. (2006). Word naming in Spanish. *European Journal of Cognitive Psychology, 18*, 415–436.

Cuetos, F., Rodriguez, B., Ruano, E., & Arribas, D. (2007). PROLEC-R. Madrid: TEA Ediciones, S.A. Davies, R., Cuetos, F., & Glez-Seijas, R.-M. (2007). Reading development and dyslexia in a transparent orthography: A survey of Spanish children. *Annals of Dyslexia, 57*, 179–198.
Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighbourhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods, 37*, 665–671.

De Luca, M., Barca, L., Burani, C., & Zoccolotti, P. (2008). The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Cognitive and Behavioral Neurology, 21*, 227–235.

Defior, S., Justicia, F., & Martos, F. J. (1996). The influence of lexical and sublexical variables in normal and poor Spanish readers. *Reading and Writing: An Interdisciplinary Journal, 8*, 497.

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

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.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124.

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

Gonzalez, J. E. J., & Valle, I. H. (2000). Word identification and reading disorders in the Spanish language. *Journal of Learning Disabilities, 33*, 44–60.

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

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review, 111*, 662–720.

Kawamoto, A. H., Goeltz, K., Agbayani, J. T., & Groel, K. (1998). Locus of semantic priming effects in speeded naming. *Psychonomic Bulletin & Review, 5*, 676–682.

Martínez, J. A., & García, M. E. (2004). *Diccionario de frecuencias del castellano escrito en niños de 6 a 12 años* (Frequency dictionary of written Spanish read by children of 6–12 years of age). Salamanca, Spain: Servicio de Publicaciones de la Universidad Pontificia de Salamanca.

Porpodas, C. D. (1999). Patterns of phonological and memory processing in beginning readers and spellers of Greek. *Journal of Learning Disabilities, 32*, 406–416.

Protopapas, A. (2007). CheckVocal: A program to facilitate checking accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*, 859–862.

Rodrigo López, M., & Jiménez González, J. E. (1999). An analysis of the word naming errors of normal readers and reading disabled children in Spanish. *Journal of Research in Reading, 22*, 180–197.

Sebastián, N., Martí, M. A., Carreiras, M. F., & Cuetos, F. (2000). *LEXESP, Léxico informatizado del español* (Computerized lexicon of Spanish). Barcelona, Spain: Ediciones de la Universitat de Barcelona.

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

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

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an “Outlier” orthography. *Psychological Bulletin, 134*, 584–615.

Valle-Arroyo, F. (2001). *Imaginabilidad en Español* (Imageability in Spanish). Oviedo: Publicaciones de la Universidad de Oviedo.

Wechsler, D. (1974). *Wechsler Intelligence Scale for Children – Revised WISC-R, Manual*. New York, NY: Psychological Corporation.

Wechsler, D. (1982). *Escala de Inteligencia para Niños-Revision (EIWN-R)* (Intelligence Scale for Children—Revision). New York, NY: Psychological Corporation.

Weekes, B. S. (1997). Differential effects of number of letters on word and pseudoword naming latency. *The Quarterly Journal of Experimental Psychology, 510A*, 439–456.

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular reading system. *Applied Psycholinguistics, 14*, 1–33.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition, 51*, 91–103.

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., Schurz, M., Sturm, D., Richlan, F., Klackl, J., Kronbichler, M., et al. (2010). A dual-route perspective on poor reading in a regular orthography: An fMRI study. *Cortex, 46*, 1284–1298.

Yap, R., & van der Leij, A. (1993). Word-processing in dyslexics—an automatic decoding deficit. *Reading and Writing: An Interdisciplinary Journal, 5*, 261–279.

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and Language, 47*, 1–29.

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.