

**A cross-linguistic, longitudinal study of the foundations of decoding and reading comprehension ability**

Caravolas, Marketa; Lervåg, Arne; Mikulajová, Marína; Defior, Sylvia; Málková, Gabriela Seidlová; Hulme, Charles

Scientific Studies of Reading

DOI:

[10.1080/10888438.2019.1580284](https://doi.org/10.1080/10888438.2019.1580284)

Published: 31/10/2019

Peer reviewed version

[Cyswllt i'r cyhoeddiad / Link to publication](#)

Dyfyniad o'r fersiwn a gyhoeddwyd / Citation for published version (APA):

Caravolas, M., Lervåg, A., Mikulajová, M., Defior, S., Málková, G. S., & Hulme, C. (2019). A cross-linguistic, longitudinal study of the foundations of decoding and reading comprehension ability. *Scientific Studies of Reading*, 23(5), 386-402.
<https://doi.org/10.1080/10888438.2019.1580284>

Hawliau Cyffredinol / General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

A cross-linguistic, longitudinal study of the foundations of decoding and reading
comprehension ability

Markéta Caravolas¹, Arne Lervåg², Marína Mikulajová³, Sylvia Defior⁴, Gabriela Seidlová-
Málková⁵, Charles Hulme⁶

¹ Bangor University, ² University of Oslo, ³ Comenius University, ⁴ University of Granada, ⁵
Charles University, and ⁶ University of Oxford

Accepted for publication in *Scientific Studies of Reading*

Corresponding Author:

Markéta Caravolas

Adeilad Brigantia

Bangor University

Bangor, LL57 2AS

United Kingdom

Email: m.caravolas@bangor.ac.uk

A cross-linguistic, longitudinal study of the foundations of decoding and reading
comprehension ability

Abstract

The present study investigated the moderating role of orthographic consistency on the development of reading comprehension in four language groups (English, $n=179$; Spanish, $n=188$; Czech, $n=135$; Slovak, $n=194$) from kindergarten to grade 2. In all languages, early variations in phoneme awareness/letter knowledge, RAN, and emerging decoding skills, but not oral language, predicted variations in decoding skills at the end of grade 1; these in turn predicted reading comprehension in grade 2. For the three consistent orthographies (Spanish, Slovak and Czech), kindergarten language skills were another significant predictor of grade 2 reading comprehension. This effect was absent in the English sample where variations in decoding skills were a more powerful predictor. These results provide the first longitudinal evidence for effects of orthographic consistency on the development of reading comprehension and provide support for the simple view of reading.

Many studies of learning to read in English support the simple view of reading according to which, decoding skills and oral language comprehension are strong predictors of reading comprehension (Gough & Tunmer, 1986; Hoover & Gough, 1990; Keenan, Betjeman & Olson, 2008; Salceda, Alonso, & Castilla-Earls, 2014). This model has also been validated in a variety of other alphabetic orthographies (e.g., de Jong & van der Leij, 2002; Florit & Cain, 2011; Kendeu, Papadopoulos, & Kotzapolou, 2013; Muller & Brady, 2001). However, it has been suggested that, in orthographies with relatively consistent letter-sound mappings, decoding is a less powerful predictor of reading comprehension than in English (e.g., see reviews in Florit & Cain, 2011; García & Cain, 2014; Salceda et al, 2014) because decoding is acquired more quickly in these orthographies (e.g., Caravolas, 2018; Caravolas, Lervåg, Defior, Seidlová Málková, & Hulme, 2013; Leppänen, Niemi, Aunola, & Nurmi, 2004; Parrila, Aunola, Leskinen, Nurmi, & Kirby, 2005; Verhoeven & van Leeuwe, 2011; Wimmer & Goswami, 1994). As a consequence, oral language skills may be a stronger predictor of reading comprehension in consistent orthographies than in English.

This claim is supported by a meta-analysis (Florit & Cain, 2011) contrasting studies of the simple view of reading in English and in relatively consistent alphabetic orthographies, where the association between decoding and reading comprehension in relatively consistent orthographies was weaker than in English. Conversely, the associations between oral language comprehension and reading comprehension skills were stronger in the relatively consistent orthographies than in English. Within the consistent orthographies, the relative weighting of the two predictors was additionally influenced by the nature of the decoding measure. When decoding was assessed with efficiency measures, the oral language contribution did not, on average, exceed that of decoding. These general patterns fit with studies of English populations which show that decoding ability becomes a weaker predictor, while oral language skills become a stronger predictor of reading comprehension, as children

gain in reading experience and skill (e.g., Adlof et al., 2010; Fuchs et al., 2012; García & Cain, 2014; Gough, Hoover, & Peterson, 1996; Quinn, 2016; Kim, Wagner, & Foster, 2011; Salceda et al., 2014; Tilstra, et al., 2009).

Studies of the development of reading comprehension in orthographies other than English are still relatively rare and include a wide variety of languages. For example, Florit and Cain's (2011) meta-analysis included 13 non-English studies, representing only eight European languages. What is more, the comparisons with English were necessarily indirect, and used different measures of each construct across studies. Thus, the basis for cross-linguistic generalisations is still tenuous.

Nevertheless, several recent single-language studies with alphabetic orthographies that are more consistent than English, suggest that decoding plays a somewhat weaker role by 2nd or 3rd grade, relative to oral language skills, in predicting reading comprehension (Portuguese: Cadime et al., 2016; Finnish: Lepola, Lynch, Kiuru, Laakkonen, & Niemi, 2016; Norwegian: Lervåg, Hulme, & Melby-Lervåg, 2017; Italian: Tobia & Bonifacci, 2015). Differences among these studies in terms of their designs and measures, however, preclude strong generalizations about the relative weightings of the constructs. To date no study has assessed decoding and oral language skills as predictors of reading comprehension longitudinally in children learning to read in different languages using directly comparable measures.

Issues in the Measurement of Constructs in the Simple View of Reading

Studies with English readers have shown that when reading comprehension is tested with cloze or multiple choice tests, decoding tends to be the stronger predictor, whereas tests with open-ended questions and retelling formats tend to elicit stronger contributions from listening comprehension (e.g., Keenan et al., 2008). However, in their systematic review of studies with English-speaking children, Salceda et al. (2014) concluded that, in the aggregate, decoding emerged as a strong predictor of both types of reading comprehension measures.

How decoding is assessed (using untimed or speeded measures, or with word versus nonword lists) may also affect its strength as a predictor of reading comprehension. This issue has particular relevance in cross-linguistic comparisons. García and Cain (2014) and Salceda et al. (2014) concur that, while tests of untimed word reading accuracy yield the most robust association with reading comprehension in English (such measures being less sensitive in languages with consistent orthographies), *all* types of decoding (i.e., word-level reading) measures are nevertheless strongly predictive for this language group. Indeed, some studies have demonstrated untimed and timed word reading tests to be similarly predictive of reading comprehension in English (e.g., Fuchs et al., 2012; Kim, Wagner, & Lopez, 2012).

Finally, measures of oral language comprehension must be considered (Lervåg et al., 2017; Kim, 2015; Savage, 2006; Silva & Cain, 2015). In their original model, Gough and Tunmer (1986) used *listening comprehension* as a global indicator of oral language comprehension skill. However, numerous studies have since explored a broader range of receptive and expressive oral language abilities, including vocabulary knowledge and morphological and syntactic skills, which have emerged as independent predictors of reading comprehension in some studies (e.g., Cain, Oakhill, & Bryant, 2004; Clarke, Snowling, Truelove, & Hulme, 2010; Muter, Hulme, Snowling, and Stevenson, 2004; Cutting & Scarborough, 2006; Savage, Ouellette, 2006; Tilstra et al., 2009). Salceda et al. (2014) reported measures of vocabulary, morphosyntax and listening comprehension to be equally strong predictors ($\rho = .49-.50$) of reading comprehension. Moreover, recent studies in English (e.g., Hulme et al., 2015) and Norwegian (Lervåg et al., 2017) confirm that a latent oral language comprehension construct can be reliably estimated from a combination of indicators including vocabulary, morphological and syntactic awareness (as well as verbal memory and inferencing skills).

Longitudinal studies of the development of reading comprehension.

Seeking a fuller developmental understanding, two longitudinal studies have mapped the development of reading comprehension from its cognitive and linguistic foundations (Hulme et al., 2015; Torppa et al., 2016). Using latent variables, both studies tracked children from before or during the earliest stages of learning to read to examine how the core foundational skills underpinning word-level reading ability (i.e., phoneme awareness and letter knowledge skills that underlie letter-sound encoding in alphabetic orthographies, and, rapid naming – a proxy for a mechanism enabling fluent mappings between visual objects and their pronunciations) relate to later measures of reading. The study by Hulme et al., (2015) tracked English children with and without family risk of dyslexia, from nursery school (age 3-4 years), until Year 3 (age 7-9 years) when reading comprehension could be reliably assessed. In line with a large body of evidence (Caravolas & Samara, 2015), structural equation models revealed a direct influence of the core foundational skills on later decoding skills. Also, oral language skills (vocabulary, syntactic knowledge, morphological knowledge, speech and articulation skills) measured at 3 to 4 years of age, and decoding skills measured in the second year of formal schooling (5-6 years), predicted children's reading comprehension in the third year of schooling. Moreover, the model confirmed that oral language skills directly predicted children's attainments on measures of the core foundational skills at the start of schooling, thus also contributing to reading comprehension indirectly via the components of word-level skills.

Torppa et al. (2016) tracked Finnish participants from kindergarten to grade 3. Like Hulme et al., (2015), they found that the three foundational code-related skills measured in kindergarten predicted reading fluency, but not listening comprehension, in first grade. In turn, grade 1 word reading predicted second grade reading comprehension. Kindergarten measures of vocabulary knowledge predicted grade 1 listening comprehension, reading fluency, and reading comprehension. In line with the simple view, grade 1 reading fluency

and listening comprehension both predicted grade 2 reading comprehension, with a relatively stronger contribution from listening comprehension. By grade 3, however, and in contrast with Hulme et al. (2015), only listening comprehension, and not word reading, continued to account for significant variance in reading comprehension. The authors attributed the latter pattern to limited variations in decoding in the highly consistent Finnish orthography (see Lervåg et al. (2017) for similar findings). These longitudinal studies suggest that foundational, code-related skills exert their effects on reading comprehension indirectly via decoding ability, while early oral language skills directly influence *both* emergent decoding skills and reading comprehension.

The present study

Here we examine the longitudinal influences of the precursors of word reading ability (phoneme awareness, letter knowledge, rapid naming) and of oral language skills on individual differences in decoding and reading comprehension in children learning to read in English, Spanish, Czech and Slovak. In line with the studies of Hulme et al. (2015) and Torppa et al. (2016) and our own earlier findings (Caravolas et al., 2012, 2013) we hypothesized that individual differences in decoding at the end of grade 1 would be strongly predicted in all languages by kindergarten measures of phoneme awareness, letter knowledge, and rapid naming. We also anticipated that decoding in grade 1 would be predicted by oral language skills (Duff & Hulme, 2012; Hulme et al., 2015; Torppa et al., 2016). Furthermore, in light of the relative difficulty of learning to decode in English, we predicted a greater contribution of oral language skills to decoding in English, where children may make greater use of lexico-semantic knowledge in recognizing written words, than might be the case in the more consistent orthographies. For reading comprehension, we expected that earlier decoding *and* oral language skills would predict reading comprehension in grade 2 in all languages. To investigate whether these two constructs were differentially predictive of

reading comprehension for children at different levels of ability, we also tested for curvilinear effects of decoding and oral language skills on reading comprehension. On the assumption that orthographic consistency moderates the rate of development of decoding skills, we expected grade 1 decoding to play a less powerful role as a predictor of reading comprehension in the consistent orthographies (Spanish, Czech and Slovak) than in English, and oral language skills to play a correspondingly bigger role. Finally, we tested whether the kindergarten measures had indirect effects on grade 2 reading comprehension via grade 1 decoding.

Method

Orthographic consistency of the languages under study

To investigate the moderating role of orthographic consistency on the early development of reading comprehension, we compared four groups of children: those learning the inconsistent orthography of English, and those learning the relatively consistent orthographies of Spanish, Czech and Slovak. (see Caravolas et al., 2012; for estimates of the consistency of these different orthographies).

Participants

A total of 696 children (179 English; 188 Spanish; 135 Czech; 194 Slovak) participated in all tests at Time 1, 645 participated at Time 2 (end-grade 1), and 625 at Time 3 (mid-grade 2). For some measures there were small amounts of missing data and thus the Ns for individual measures varied slightly. Details of participant sampling and other demographic and schooling information are provided in Caravolas et al., (2012). Table 1 reports the summary statistics for sample sizes, age, gender, and nonverbal reasoning ability of the groups

participating in the present study. Participants were monolingual speakers of the language of their country. At initial testing (mid-school year), all children were attending kindergarten (Reception Year in England). The English children, although younger, had started to receive formal literacy instruction some 5-6 months prior to the start of the study. For the other groups, formal literacy instruction had not started. The group discrepancies in age reflect differences in the age of school entry between countries. From first grade onward, all groups received phonically-based instruction. Data were collected concurrently in each country at each data collection point. Here we report results from the initial time point (Time 1) in February/March of their reception/kindergarten year, approximately 16 months later at end-grade 1 (Time 2), and approximately 6 months later in mid-grade 2 (Time 3). The study was carried out in compliance with the British Psychological Society's Guidelines for the Minimum Standards for Ethical Approval for Psychological Research (2004) and with the Convention on Human Rights in Biomedicine or Biomedical Research.

Time 1 Measures

Letter Knowledge. In four separate trials, children were asked to say the names and sounds for each letter of their alphabet (upper and lower case). The number of letters correct was summed to give measures of letter-name and letter-sound knowledge. These were then combined into a composite measure.

Phoneme Awareness. *Phoneme Isolation* and *Phoneme Blending* tests were used, as described in detail in Caravolas et al. (2012). *Phoneme Isolation* comprised two blocks of 16 nonword items in which the child was asked to isolate and pronounce the initial (Block 1) or the final (Block 2) consonant phoneme of spoken nonword monosyllables; in each Block, syllable complexity was manipulated such that the consonant to be isolated was a singleton (CVC or CVC) or part of a consonant cluster (CCVC or CVCC). Testing was discontinued after four consecutive errors in a block. *Phoneme Blending* required children to blend spoken

phonemic segments into words. Ten (11 in Czech and Slovak) mono- and bisyllabic items with increasingly complex syllable structures were administered in a fixed order without corrective feedback. The test was discontinued after six consecutive responses that showed no overlap with any of the sounds in the target word. The scores from both tasks were combined into a phoneme awareness composite measure.

Rapid Automatised Naming (RAN) of Objects and Colours. Parallel versions of these tasks were created, as described in Caravolas et al. (2012); across languages the stimuli were identical and corresponded to names of comparable length, phonological complexity, familiarity and frequency. Children named five items at speed that were repeated eight times over five lines of an A4 card. Two trials of each task were administered, with items arranged in a different quasi-random order. Children first named each of the stimuli on a separate card prior to the start of each task. The average naming time for the two trials of each task was calculated. Error rates were very low across languages (0.39 to 1.98 for objects; 0.52 to 2.14 for colours).

1 Minute Word Reading. This test has been described in detail in Caravolas (2018; Caravolas et al., 2012). Children read aloud as quickly as possible a list of 140 (144 in English) high-frequency words. The lists contained progressively longer and more complex words, starting with single letter words (e.g., ‘a’ in English) and up to two- and three-syllable words (up to five-syllables in Spanish). The distribution of words by syllable number varied across languages to reflect their distribution in each language. The number of words read correctly in 60 seconds was recorded as the reading efficiency score.

Vocabulary. The *Vocabulary* subtest of the Wechsler Preschool and Primary Scale of Intelligence for Children, WPPSI-III in English (Wechsler, 2003) and WPPSI in Spanish (Wechsler, 2001) was administered. For Czech and Slovak, we adapted the English versions and created standardized scores based on extended kindergarten-aged samples.

Morphological Knowledge. This test consisted of blocks of items assessing knowledge of noun phrase, verb phrase, and derivational morphology (see Appendix A for details). The morphological manipulations were selected to be age-appropriate and relevant in each language. Item-level matching across languages was not appropriate in this task but the structure and task demands were the same. The test consisted of 24 (Czech, Slovak) to 28 (English, Spanish) items. For each item, the administrator uttered a stem sentence followed by a second sentence in which the final word was omitted. The child was required to supply the missing word, which demanded a variation in morphological ending from the one used in the stem sentence. The noun phrase morphology items were accompanied by pictures, whereas the remaining test items were delivered aurally, without picture support. Each correct answer was awarded one point. All correct responses were summed across all blocks and were converted to percentages.

Syntactic Knowledge. Part of the Test for Reception of Grammar (TROG-2; Bishop, 2003) was used to measure children's receptive knowledge of syntax. In English and Spanish (Bishop, 2003; Mendoza, Carballo, Muñoz, & Fresneda, 2005), the published versions were used, while adaptations were created for Czech and Slovak. In particular, blocks G, K, S and T of the English TROG-2 were selected to assess children's understanding of subject relatives, passive voice, object relatives, and embedded clauses. Each block consisted of four items (16 items in total) in which children heard a sentence and were required to select one of four pictures representing the meaning of the sentence. Each correct answer was awarded one point and the total number of items correct was calculated.

Time 2 Measures

1 Minute Word Reading. As at Time 1.

1 Minute Pseudoword Reading. This test has been described in detail in Caravolas (2018). Children read lists of pseudowords, which were derived from the real words in the 1

Minute Word Reading measure and thus maintained the same numbers of graphemes and syllable structures. Again, the number of possible items was 140 (144 in English), and all were phonologically legal, pronounceable pseudowords. The number of pseudowords read correctly in 60 seconds was recorded as the pseudoword reading efficiency score.

Time 3 Measure

Reading Comprehension. The items for this task were selected from the Gates-MacGinitie Reading tests – Levels 1 and 2 (fourth edition) (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) and were adapted in each language. Children silently read five short texts, each consisting of four short passages. Each passage was associated with three pictures printed on a sheet and children were instructed to select the one that agreed with what they had read in the passage.

The cross-language adaptation was carried out in line with key Guidelines for Translating and Adapting Tests (2018) and Hambleton’s (2002) recommendations for successful test adaptation. Specifically, the English source material was examined from a cross-linguistic and cross-cultural perspective by reading experts fluent in English and in the relevant native languages (Spanish, Czech, Slovak) at a joint meeting. A list of plausible items was selected on the basis of: their level of difficulty (i.e., age- and grade-appropriateness considered against published reading curriculum targets in each country); their linguistic ‘transferability’ (i.e., the possibility to convey the same concepts in comparable syntactic constructions-- e.g., active vs. passive voice--, and semantic depth— e.g., concretely expressed vs. inferred information); and their cultural relevance (e.g., exclusion of an item about baseball, a sport not well known in the participating countries, vs. inclusion of items about animals deemed generally appropriate). A short list of items was next considered for potential linguistic translation issues (e.g., the level of inferencing, the syntactic difficulty and sentence length). Each group then delegated a translator to carry out

the translation. The resulting texts were scrutinised by members of each language group and the PI, and any differences in opinion about the optimal wording, structure etc., were resolved within and across language groups.

Details of the test characteristics in each language are provided in Table S1 in the online Supplemental materials. One story consisting of three practice items was used to familiarise children with the task. Order of presentation was graded in difficulty and fixed in all languages. The test was administered to small groups of 4-8 pupils, and although not administered under time pressure, the maximum allotted time was 15 minutes. The majority of children completed the test within this time (two Czech, two Slovak, and seven English children failed to complete in 15 minutes). Each correct answer was awarded one point, to a possible maximum score of 20.

Results

Means, standard deviations, ranges and reliabilities for all measures in the English, Spanish, Czech and Slovak groups are shown in Table 2. Table S2 (available in Supplemental Material online) shows the estimated correlations between all variables for each group. We have described the relative performances by group on the measures of letter knowledge, phoneme awareness and RAN, as well as for word reading efficiency at Time 1 (mid-kindergarten) in a previous paper (Caravolas et al., 2012). Briefly, Table 2 shows similar levels and patterns of performance across languages for the core foundational measures with the exception of letter knowledge, where children's scores reflected national differences in alphabet instruction (or lack thereof) in the preschool years. Word reading at Time 1 was extremely low in all groups, with fewer than four words (of which three single-letter words) read per minute for Czech and Slovak, and fewer than nine words (all comprising one or two letter words) per minute for English and Spanish. On the measures of oral language, all

groups showed similar scaled vocabulary scores that approximated the population average. In morphological knowledge and syntactic knowledge, some between-language fluctuations in mean attainments emerged, and these were consistent with system-wide differences in morphological and syntactic complexity among the languages (children in the more consistent languages, which are also morphologically more complex than English, showed higher attainments than the English group). However, direct comparisons between groups on the latter two measures warrants caution as it was not possible to create identical tests across languages. Despite these differences, in all languages the oral language measures showed similar patterns of correlations with the other measures in the battery. The Time 2 measures of reading efficiency (one minute word and pseudoword reading) had high test-retest stability and were similar across languages in terms of performance levels (with the exception of the relatively lower English pseudoword reading efficiency score). The reading comprehension measure yielded similar levels of accuracy across the high consistency groups, and a lower mean attainment in the English group; nevertheless, all language versions had reasonable to good internal consistency.

Predicting reading comprehension

We used structural equation models (SEM) to examine the longitudinal predictors of decoding efficiency at the end of grade 1, and reading comprehension in mid-grade 2. A multiple-group factor analysis (CFA) of all latent variables with two or more indicators (Phoneme/Letter Knowledge, RAN, Language and Decoding Time 2) with all four samples included revealed a lack of metric ($\chi^2 (18) = 73.323, p < .001$) and scalar ($\chi^2 (36) = 901.168, p < .001$) invariance across the groups. However, pairwise comparisons between orthographies showed that there was metric ($\chi^2 (6) = 3.659, p < .723$) but not scalar ($\chi^2 (12) = 72.111, p < .001$) invariance between the English and Spanish groups. No metric or scalar invariance was found between any of the other pair of groups. Therefore, we estimated one

two-group model for the English and Spanish samples where the factor loadings but not the intercepts were fixed to be equal across groups, and, equivalent single-group models for the Czech and Slovak samples. In these models, we present 95% confidence intervals around the path coefficients, but comparing these path weights across the groups (other than the English-Spanish comparison) warrants caution since the weightings of the indicators of the latent variables vary across groups. As word reading at Time 1 and reading comprehension at Time 3 were measured with only one observed variable, we fixed their residuals to reflect the reliability (alpha) of the variable within each group. Also, as word reading at Time 1 was heavily skewed, we log transformed it within all the four groups. All analyses were carried out using full information maximum likelihood (FIML) with Mplus version 8 (Muthèn & Muthèn, 1998-2017).

For each language group, we regressed reading comprehension at Time 3 on language at Time 1 and decoding efficiency at Time 2. Decoding efficiency at Time 2 was regressed on Time 1 measures of word reading, a latent phoneme awareness/letter knowledge construct (phoneme blending, phoneme identification and letter knowledge), a latent RAN construct (RAN colours and RAN objects) and a latent language construct (vocabulary, morphological knowledge and syntactic knowledge). In these models, we observed strong collinearity between the latent phoneme awareness/letter knowledge construct and word reading in all four groups. The correlation between these two constructs ranged from .974 in Spanish to .793 in Slovak leading to large standard errors for the path coefficients. As a result, neither the latent phoneme awareness/letter knowledge construct nor word reading at Time 1 were significant predictors of decoding efficiency at Time 2 in the Spanish, Czech and Slovak samples although deleting the two paths gave a significant loss in model fit. In the English sample, the latent phoneme awareness/letter knowledge construct did predict decoding efficiency at T2 ($\beta = .776, p = .008$) but word reading at T1 did not ($\beta = -.057, p = .812$).

In order to deal with the above collinearity, we standardized the constructs and constrained these two path coefficients to be equal in each of the four groups. Wald tests confirmed that these constrained models did not differ significantly from the original models where the two paths were freely estimated (English: $\chi^2(1) = 1.743$, $p = .187$, Spanish: $\chi^2(1) = .001$, $p = .981$, Czech: $\chi^2(1) = .004$, $p = .948$ and Slovak: $\chi^2(1) = .592$, $p = .442$). In addition, we estimated curvilinear effects of decoding efficiency and oral language in order to account for possible nonlinear relationships between these constructs and reading comprehension. However, as these curvilinear effects models require numerical integration in Mplus, which cannot be used to estimate multi-group models, we estimated four equivalent single-group models. Further, in order to compare the confidence intervals for the comparable regression paths across the English and Spanish samples, we fixed the unstandardized factor loadings to be equal to the factor loadings obtained from a two-group model with only linear paths. This two-group model fitted the data well, $\chi^2(102) = 134.719$, $p = .017$, RMSEA = .042 (90% CI = .019-.060), TLI = .981, SRMR = .085. Further, as no curvilinear effects of oral language were found in any of the language groups, this curvilinear parameter was deleted from the final models that are shown in in Figures 1a-d.

Decoding at Time 2 was a unique predictor of reading comprehension in all groups but oral language at Time 1 was only a significant predictor of reading comprehension in the three consistent orthographies (Spanish, Czech and Slovak). These two predictors accounted for 40.7% (Czech) to 68.7% (Spanish) of the variance in reading comprehension. Decoding was a strong predictor in the English group and a moderate predictor in the three consistent orthographies. In the Spanish and Slovak groups, there was a significant negative curvilinear relationship between decoding efficiency and reading comprehension indicating that variations in word-level reading efficiency are less strongly related to reading comprehension for children with good than poor decoding skills.

Phoneme awareness, letter knowledge and Time 1 word reading combined together predicted decoding efficiency at Time 2 after accounting for the effects of RAN and oral language at Time 1. RAN also uniquely predicted decoding at Time 2, but this was not the case for the language construct. The variance explained in decoding efficiency seemed higher in English ($R^2 = .743$) than in the three consistent orthographies (R^2 ranging from .270 in Spanish to .433 in Czech).

Comparing the 95% confidence intervals shown in Figure 1a and 1b for the English and the Spanish sample respectively, we can see that language in Grade 1 is a stronger predictor of reading comprehension in the Spanish than in the English sample. There is however, an overlap in the confidence intervals in both the linear and curvilinear parameter that constitutes the path from decoding to reading comprehension. However, this parameter is difficult to compare as it is significant in the Spanish but not in the English sample. Concerning the prediction of decoding skills, early decoding and phoneme awareness/letter knowledge are stronger as a combined predictor in the English than in the Spanish sample. No differences were found between these two samples concerning RAN and language as predictors of later decoding skills.

The fit of these four models without the curvilinear effects (full model fit for models with the curvilinear relationships between language cannot be estimated as they require numerical integration in Mplus) was good to moderate for all four orthographies, English, $\chi^2(55) = 70.638, p = .076, RMSEA = .040$ (90% CI = .000-.065), TLI = .987, SRMR = .083; Spanish, $\chi^2(56) = 62.952, p = .244, RMSEA = .026$ (90% CI = .000-.054), TLI = .994, SRMR = .086; Czech, $\chi^2(45) = 74.530, p = .004, RMSEA = .070$ (90% CI = .040-.097), TLI = .954, SRMR = .043; and Slovak, $\chi^2(45) = 101.115, p < .001, RMSEA = .082$ (90% CI = .060-.103), TLI = .927, SRMR = .051. The moderate fit of the Czech and the Slovak models seems to reflect the fact that word reading at Time 1 was more related to word reading

efficiency at Time 2 than it was with pseudoword reading efficiency. As this uneven association between phoneme awareness/letter knowledge and the two word-level reading indicators did not affect the research questions, we chose to keep the model as it was.

Indirect effects on reading comprehension

To test for indirect effects of the predictors at Time 1 on reading comprehension at Time 3 through decoding efficiency at Time 2, we bootstrapped confidence intervals with 1000 bootstraps in the linear versions of the single-group models. The results of these estimations are shown in Table 3. There were significant indirect effects of both RAN and word reading and phoneme awareness/letter knowledge combined in the English, Spanish and Czech samples. The indirect effect of language was not significant in any of the groups. The indirect effects in the Slovak sample did not reach significance, although its pattern of results was similar to that of the other consistent orthography groups.

Discussion

This is the first longitudinal, cross-linguistic study using directly comparable measures to assess the foundations of decoding skills, and the relative importance of decoding and language comprehension skills as predictors of later reading comprehension. The basic pattern of findings is very clear. In all four languages (English, Spanish, Slovak and Czech) early (reception/kindergarten class) variations in word reading, phoneme awareness/letter knowledge, and RAN predicted variations in decoding skills some 16 months later at the end of grade 1. Reading comprehension in mid-grade 2 was predicted in all cases by variations in decoding skill in grade 1. For the three consistent orthographies, kindergarten language skills were also a significant predictor of reading comprehension in grade 2. However, language skills were not a predictor of reading comprehension in the English sample where variations in decoding skills were the only unique predictor. While a

strong effect of decoding was anticipated, we had expected language to play some role as a predictor of reading comprehension in English. The absence of a unique effect of language skills here probably reflects the fact that we studied children in the very early stages of reading development and used a multiple choice reading comprehension measure, which may have elicited a stronger reliance on decoding (Keenan et al., 2008).

Our finding that decoding was a stronger predictor of reading comprehension in English than in more consistent orthographies aligns with much earlier research (e.g., Cadime et al., 2016; Florit & Cain, 2011; Lepola et al., 2016; Lervåg et al., 2017; Salceda et al., 2014; Florit & Cain, 2011). In the consistent orthographies, we found the direct contributions of decoding *and* oral language skills to reading comprehension to be moderately predictive. Contrary to some earlier non-English findings (e.g., Cadime et al. 2016; Lepola et al., 2016), in Czech decoding was a strong predictor of reading comprehension and was only a marginally weaker predictor than oral language in Spanish and Slovak. A significant negative quadratic component in Spanish and Slovak furthermore indicated that decoding skill was more predictive of reading comprehension for poor than for good decoders, while good readers' skills were more homogeneous already by the end of first grade. The somewhat different pattern of relationships in the Czech group is difficult to explain, and may simply reflect a random sampling difference. Alternatively, perhaps the somewhat lower Time 2 decoding scores of the Czech group, like those of the English group, precluded a differentiation between the top decoders in their waning dependence on word reading skills for reading comprehension relative to their less skilled peers. Such an interpretation is in keeping with the 'simple view' that the effect of decoding on reading comprehension decreases with increasing decoding skill and experience (e.g., Fuchs et al., 2012; García & Cain, 2014; Kim, Wagner, & Foster, 2011), and it suggests that eventually (given a sufficient level of decoding efficiency) children learning any alphabetic orthography will show a

curvilinear effect like the one observed here for Spanish and Slovak. This interpretation awaits future replications, and is speculative given that direct statistical comparisons across groups were only partially possible in this study.

Our examination of the early precursors of decoding skills showed that phoneme awareness/letter knowledge, RAN and early word reading were predictors of decoding skill 16 month later in all four languages. This pattern is broadly consistent with recent longitudinal, cross-linguistic studies, which however reported interactive relationships between the above foundational skills and reading over time (e.g., Landerl et al., 2018; Peterson et al., 2017). Our pattern of results also replicates the longitudinal studies of Hulme et al. (2015) and of Torppa et al. (2016), which tracked children through similar phases of development. Our cross-linguistic comparison did show, however, that the proportion of variance explained in decoding by the three foundational skills and reading itself was considerably higher in English ($R^2 = .709$) than in the other languages. This underscores once more the critical role played by phoneme awareness, letter knowledge and RAN for the mastery of decoding skills in English at this early stage of reading development (Hulme & Snowling, 2013).

Regarding effects of grapheme-phoneme consistency, we found that children learning relatively consistent orthographies tended to be more advanced word readers by the end of first grade, and comprehended better what they read in second grade than children learning English. While factors other than orthographic consistency may have contributed to these differences, we attempted to minimize the most likely extraneous influences. For example, all participants were being taught to read by phonics methods, all were attending kindergarten (Reception Year) at the start of the study and had comparable starting levels of reading and related skills (see details in Caravolas et al., 2012, 2013). Despite these similarities, the

groups differed in age by up to 12 months which may have influenced the relative weightings of language and decoding skills as predictors of reading comprehension.

In earlier analyses, based on the same longitudinal study, examining the predictors of initial (mid-grade 1) and later word reading skills (Caravolas et al., 2012), we found that none of the kindergarten/reception year oral language measures emerged as a unique predictor beyond the core, code-related skills of phoneme awareness, letter knowledge and RAN. In the present study, we again found no direct effects of oral language on decoding at a somewhat later (end-grade 1) phase of reading development; moreover, the core predictors of word-level reading (but not language skills) exerted an indirect effect, via grade 1 decoding, on reading comprehension outcomes in grade 2. This influence was observed in all four languages, although it was not significant in the Slovak sample. In other earlier studies with these groups, we confirmed that, despite similar foundations, children learning consistent orthographies learned to read words and pseudowords at a faster rate than their English peers (Caravolas, 2018; Caravolas et al., 2013). In the present study, we demonstrated that the faster rate of reading development extends also to reading comprehension, our English cohort attaining lower accuracy (60%) than the other three groups (75%) on a reading comprehension test.

The patterns reported here are broadly consistent with theoretical expectations about the moderating influences of orthographic consistency on reading development. Nevertheless, the modelling of a longer-term and more complex account of reading development flagged up areas for consideration in future studies. First, we observed some between-language differences in the proportion of variance explained by the models. These differences contrast with some of our earlier studies demonstrating invariant results across languages (Caravolas et al., 2012, 2013). Importantly, those studies focused on earlier and shorter phases of literacy development, and on simpler models in terms of the number of constructs to be modelled.

Second, the inclusion of Time 1 reading as an autoregressor in the present models introduced statistical artefacts of multicollinearity with phoneme awareness and letter knowledge. The groups' reading levels at Time 1 were extremely low. Thus the reading measure probably estimated early cipher knowledge (e.g., Ehri, 1992) rather than reading per se, and obscured the true roles of phoneme awareness and letter knowledge in promoting later reading skill.

Furthermore, while the inclusion of several latent variables (PA/LK, RAN, language and T2 Decoding) in the present study allowed us to test for measurement invariance across languages, this process revealed some of the challenges in conducting complex cross-linguistic comparisons. We took great care to create comparable measures across languages at each time point. The factor structure was comparable across all languages, although tests of factorial invariance showed that we achieved metric invariance for the latent constructs (PA/LK, RAN, language and T2 Decoding) only for English and Spanish (two languages differing significantly in terms of orthographic consistency). We can say with confidence that for English and Spanish there were differences in the pattern of predictors of reading comprehension, while comparisons between the pattern of predictive relationships between the other languages need to be interpreted with more caution. Despite the challenges presented by investigations of increased complexity in the types and number of constructs measured and with longer developmental windows, it will be interesting to see in future studies whether the observed cross-linguistic differences in the patterns of relationships that we ascribe here to the effects of orthographic consistency will generalize to a broader range of languages, including those with orthographies with relatively low and/or intermediate levels of orthographic inconsistency.

With respect to the simple view of reading, the present results confirmed some but not all predicted between-orthography differences. While oral language did exert a stronger direct influence on reading comprehension in consistent orthographies than in English (while in

English decoding had an unequivocally stronger direct and indirect effect), it was not a clearly stronger predictor than decoding within these languages (cf. Florit & Cain, 2011). This relatively balanced pattern is also anticipated among older English readers (e.g., Quinn, 2016). Importantly, however, the longitudinal design of the present study sheds light on the way in which orthographic consistency may moderate the distribution of labour between code-related skills (the associations between orthography and phonology) and semantic (and broader oral language) knowledge during reading development.

Conflict of Interest

The authors declared no conflicts of interest with respect to the authorship or publication of this article.

Acknowledgements

Acknowledgments

This study was supported by Grant PITN-215961 – ELDEL from the Marie Curie, Seventh Framework Programme. We are grateful to research fellows of the Enhancing Literary Development in European Languages network, Petroula Mousikou, Naymé Salas, Miroslava Schöffelová, Eduardo Onochie-Quintanilla, and Miroslav Litavský for their contributions to this study.

References

- Adlof, S., Catts, H., & Lee, J. (2010). Kindergarten predictors of second vs. eighth grade reading comprehension impairments. *Journal of Learning Disabilities, 43*, 332—345. doi:10.1177/0022219410369067.
- Bishop, D. V. M. (2003). *Test for Reception of Grammar (TROG-2)*. Pearson Assessment.
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977–990.
- Cadime, I., Rodrigues, B., Santos, S., Viana, F., Chaves-Sousa, S., Cosme, M., & Ribiero, I. (2016). The role of word recognition, oral reading fluency and listening comprehension in the simple view of reading : a study in an intermediate depth orthography. *Reading and Writing, Online First*, DOI 10.1007/s11145-016-9691-3
- Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31–42. doi:10.1037/0022-0663.96.1.31.
- Caravolas, M. (2018). The Growth of Word and Pseudoword Reading Efficiency in Alphabetic Orthographies: Impact of Consistency. *Journal of Learning Disabilities, 51*, 422-433.
- Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*, 1398-1407. doi: 10.1177/0956797612473122.
- Caravolas, M., Lervåg, A., Mousikou, P., Effrim, C., Litavský, M., Onochie-Quintanilla, E., Salas, N., Schöffelová, M., Defior, S., Mikulajová, M., Seidlová-Málková, G., & Hulme,

- C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, 23(6), 678–686. doi:10.1177/0956797611434536
- Caravolas, M., & Samara, A. (2015). Learning to read and spell words in different writing systems. In A. Pollatsek and R. Treiman (Eds.). *Oxford Handbook of Reading*. OUP.
- Clarke, P. J, Snowling, M. J., Trulove, E., & Hulme, C. (2010). Ameliorating children’s reading comprehension difficulties: A randomized controlled trial. *Psychological Science*, 21, 1106-1116. DOI: 10.1177/0956797610375449
- Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, 10, 277–299.
- de Jong, P. F., & van der Leij, A. (2002). Effects of phonological abilities and linguistic comprehension on the development of reading. *Scientific Studies of Reading*, 6, 51–77.
- Duff, F. & Hulme, C. (2012). The role of children’s phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading*, 16, 504-525, DOI: 10.1080/10888438.2011.598199
- Ehri, L. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107-143). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, 23, 553–576. DOI: 10.1007/s10648-011-9175-6
- Fuchs, D., Compton, D., Fuchs, L., Bryant, J., Hamlett, C., & Lambert, W. (2012). First-grade cognitive abilities as long-term predictors of reading comprehension and

disability status. *Journal of Learning Disabilities*, 45, 217-231.

doi:10.1177/0022219412442154

García, J. R. & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84, 74-111. DOI: 10.3102/0034654313499616

Gough, P., Hoover, W. & Peterson, C. (1996). Some observations on a simple view of reading. In Cornoldi, C.; Oakhill, J., eds. *Reading comprehension difficulties: Processes and interventions*. Mahwah, NJ: Lawrence Erlbaum; p. 1-13.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7, 6–10. DOI: 10.1177/074193258600700104

Hambleton, R. (2002). Adapting achievement tests into multiple languages for international assessments. In A. C. Porter and A. Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement*. National Research Council.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing*, 2, 127–160. DOI: 10.1007/BF00401799

Hulme, C., Nash, H., Gooch, D., Lervag, A., & Snowling, M. (2015). The foundations of literacy development in children at familial risk of dyslexia, *Psychological Science*, 26, 1877–1886. doi: 10.1177/0956797615603702

Hulme, C., & Snowling, M.J., (2013) The interface between spoken and written language: Developmental disorders. *Proceedings of the Royal Society of London B*, 369, 20120395. <http://dx.doi.org/10.1098/rstb.2012.0395>.

International Test Commission (2018). Guidelines for Translating and Adapting Tests (Second Edition), *International Journal of Testing*, 18: 101–134, DOI: [10.1080/15305058.2017.1398166](https://doi.org/10.1080/15305058.2017.1398166).

- Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*, 281–300. DOI: 10.1080/10888430802132279
- Kendeu, P., Papadopoulou, T., & Kotzapolou, M. (2013). Evidence for the early emergence of the simple view of reading in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal, 26*, 189-204.
- Kessler B. (2009). *Weslalex User's Guide: West Slavic lexicon of child-directed printed words*. Retrieved from <http://spell.psychology.wustl.edu/weslalex>
- Kim, Y. S. (2015). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. *Child Development, 86*(1), 128-144. doi:10.1111/cdev.12293.
- Kim, Y., Wagner, R., & Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. *Scientific Studies of Reading, 15*, 338-362. doi:10.1080/10888438.2010.493964.
- Kim, Y., Wagner, R., & Lopez, D. (2012). Developmental relations between reading fluency and reading comprehension: A longitudinal study from grade 1 to grade 2. *Journal of Experimental Child Psychology, 113*(1), 93–111. doi:10.1016/j.jecp.2012.03.002.
- Landerl, K., Freudenthaler, H., Heene, M., de Jong, P., Desrochers, A., Manolitsis, G., Parilla, R., & Georgiou, G. (2018). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Scientific Studies of Reading*, <https://doi.org/10.1080/10888438.2018.1510936>
- Lepola, J., Lynch, J., Kiuru, N., Laakkonen, E., Niemi, P. (2016). Early oral language comprehension, task orientation, and foundational reading skills as predictors of grade

- 3 reading comprehension. *Reading Research Quarterly*, 51, 373–390,
doi:10.1002/rrq.145
- Leppänen, U., Niemi, P., Aunola, K., & Nurmi, J. E. (2004). Development of reading skills among preschool and primary school pupils. *Reading Research Quarterly*, 39, 72-93.
DOI: 10.1598/RRQ.39.1.5
- Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2017). Unpicking the developmental relationships between oral language skills and reading comprehension: It's simple, but complex. *Child Development*, doi:10.1111/cdev.12861
- MacGinitie, W., MacGinitie, R., Maria, K., & Dreyer, L., & Hughes, K. (2000). *Gates-MacGinitie Reading Tests*® Fourth Edition. Houghton Mifflin Harcourt.
- Martínez, J.A. & García, M.E. (2004). *Diccionario de frecuencias del castellano escrito en niños de 6 a 12 años*. Salamanca: Servicio de Publicaciones Universidad Pontificia de Salamanca.
- Mendoza, E., Carballo, G., Muñoz, J., & Fresneda, M. D. (2005). Test de comprensión de estructuras gramaticales (CEG). *Lexicografía y enseñanza de la lengua española*, 151.
- Müller, K., & Brady, S. (2001). Correlates of early reading performance in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal*, 14, 757–799.
- Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, 40, 665-681. DOI: 10.1037/0012-1649.40.5.665
- Muthén, L. K., & Muthén, B. O. (1998-2015). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

- Parrila, R., Aunola, K., Leskinen, E., Nurmi, J. E., & Kirby, J. R. (2005). Development of individual differences in reading: Results from longitudinal studies in English and Finnish. *Journal of Educational Psychology, 97*, 299-319. DOI: 10.1037/0022-0663.97.3.299
- Peterson, R., Arnett, A., Pennington, B., Byrne, B., Samuelsson, S., & Olson, R. (2017). Literacy acquisition influences children's rapid automatized naming. *Developmental Science*, DOI: 10.1111/desc.12589
- Quinn, J. (2016). *Predictors of Reading Comprehension: A model-based meta-analytic review*. Retrieved from ProQuest Dissertations Publishing, 2016. 10161140.
- Salceda, J., Alonso, G., & Castilla-Earls, A. (2014). The simple view of reading in elementary school: A systematic review. *Revista de Logopedia, Foniatría y Audiología, 43*, 17-31.
- Savage, R. S. (2006). Reading comprehension is not always the product of nonsense word decoding and linguistic comprehension: Evidence from teenagers who are extremely poor readers. *Scientific Studies of Reading, 10*, 143–164.
- Silva, M., & Cain, K. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology, 107*(2), 321–331. doi:10.1037/a0037769.
- Verhoeven, L., & van Leeuwe, J. (2011). Role of gender and linguistic diversity in word decoding development. *Learning and Individual Differences, 21*, 359-367. DOI: 10.1016/j.lindif.2011.02.004
- Tilstra, J., McMaster, K., van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading, 32*, 383–401.

- Tobia, V., & Bonifacci, P. (2015). The simple view of reading in a transparent orthography: The stronger role of oral comprehension. *Reading and Writing*, 28, 939–957.
doi:[10.1007/s11145-015-9556-1](https://doi.org/10.1007/s11145-015-9556-1).
- Torppa, M., Georgiou, G. K., Lerkkanen, M.-K., Niemi, P., Poikkeus, A.-M., & Nurmi, J.-E. (2016). Examining the Simple View of Reading in a Transparent Orthography: A Longitudinal Study From Kindergarten to Grade 3. *Merrill-Palmer Quarterly*, 62, 179-206. Retrieved from <http://digitalcommons.wayne.edu/mpq/vol62/iss2/4/>,
- Wechsler, D. (2001). *WPPSI-III: Escala de Inteligencia de Wechsler para Preescolar y Primaria* [WPPSI-III: Wechsler Preschool and Primary Scale of Intelligence]. Madrid, Spain: Tea.
- Wechsler, D. (2003). *Wechsler Preschool and Primary Scale of Intelligence—Third UK Edition (WPPSI-IIIUK)*. Oxford, England: Psychological Corp.
- Wimmer, H. & Goswami, U. (1994). The influence of orthographic consistency on reading development – word recognition in English and German children. *Cognition*, 51 (1), 91-103.

Appendix 1

Morphological Knowledge Procedure. The *first block of items* assessed knowledge of noun phrase morphology. In the English test, the first eight items assessed knowledge of plural formation (regular and irregular nouns) and the remaining four assessed knowledge of genitive-case marking. The Spanish test consisted of four items that required knowledge of plural endings and another four that required knowledge of the feminine form of adjectives. The Slovak test consisted of 12 items overall: four requiring knowledge of plural endings, four assessing knowledge of production of genitive and another four assessing children's knowledge of grammatical agreement. In the Czech test, four items required knowledge of plural endings, four assessed knowledge of production of adjectives from nouns or verbs and another four assessed knowledge of grammatical agreement. The *second block of items* assessed knowledge of verb phrase morphology. In English, the first four items tested knowledge of third person singular formation in the present tense and the remaining eight items assessed knowledge of simple past formation (regular and irregular verbs). The Spanish test consisted of six items that assessed knowledge of present tense formation in all persons and eight items that assessed specific knowledge of plural, present and past tense formation. In the Slovak and Czech tests, the first four items assessed knowledge of third person singular formation in the present tense and another four assessed knowledge of simple past formation. The *third block* of items assessed knowledge of derivational morphology. The English test comprised eight items, the Spanish test six items, the Czech and Slovak tests four items, respectively. For each morphological aspect, two practice items were used. The aim of the practice items was to familiarise children with the task and to model, tacitly, the knowledge they needed to use to complete the sentences.

Figure Captions

Figure 1a. Structural equation model for English children predicting the (a) early decoding efficiency skills at the end of grade 1 (Time2) from word reading, phoneme awareness/letter knowledge, rapid-automatized-naming (RAN) speed, and general language skills at the onset of literacy instruction (Time 1) and (b) early reading comprehension skills at the middle of grade 2 (Time 3) from general language skills Time 1 at the onset of literacy instruction (Time 1) and early decoding efficiency skills at the end of grade 1 (Time2). The path weights [and 95% CIs] are reported for all paths; non-significant paths are designated by dotted lines.

Figure 1b. The same path model as Figure 1a above fitted to the data of the Spanish children. This model, in addition, includes a non-linear path between Time 2 decoding efficiency and Time 3 reading comprehension.

Figure 1c. The same path model as Figure 1a above fitted to the data of the Czech children.

Figure 1d. The same path model as Figure 1a above fitted to the data of the Slovak children. This model, in addition, includes a non-linear path between Time 2 decoding efficiency and Time 3 reading comprehension

Table 1

Participant details and mean general ability scores for each group at Time 1.

Language	N (of which boys)	Mean Age ^a (SD)	Range	WPPSI Block Design ^d (SD)
English	179 (94)	60.20 (3.67)	53-66	10.27 (3.12)
Spanish	188 (103)	66.76 (3.65)	61-73	11.68 (3.41)
Czech	135 (64)	72.00 (3.73)	64-80	10.35 (2.86)
Slovak	194 (101)	71.85 (3.78)	62-81	10.22 (2.96)

Note: ^a – age in months at Time 1; ^b – mean scaled scores.

Table 2.

Descriptive statistics and reliabilities of Time 1 and Time 2 predictor and outcome measures for each language group.

	Mean	S.D.	Range	$\alpha/r^{a,c}$
English				
Letter Names (T1)	16.26	15.24	0 – 52	.97
Letter Sounds (T1)	38.70	11.99	2 – 52	.96
Phoneme Isolation (T1)	35.10	17.03	0 – 64	.94
Phoneme Blending (T1)	5.60	2.37	0 – 10	.76
RAN Colours Speed (s) (T1)	66.34	21.51	32.00 – 138.18	.84 ^a
RAN Objects Speed (s) (T1)	57.69	16.47	31 – 104.78	.84 ^a
Vocabulary (T1) (scaled score)	9.63	2.61	3-15	--- ^b
Morphological Knowledge (T1) (percent)	47.05	16.10	0 – 82	.80
Syntactic Knowledge (T1) (max=16)	7.71	2.67	2 – 15	.58
One Minute Reading (T1)	8.75	13.27	0 - 92	.98 ^c
One Minute Reading (T2)	50.94	26.97	0 – 113	.91 ^c
One Minute Pseudoword Reading (T2)	22.45	15.04	0-69	.85 ^c
Reading Comprehension (T3)	11.67	4.83	1 – 20	.89
Spanish				
Letter Names	24.92	17.54	1 -58	.98
Letter Sounds	18.57	16.59	0 – 56	.98
Phoneme Isolation	22.10	23.49	0 – 64	.97
Phoneme Blending	9.0	2.30	0 - 9	.87

RAN Colours (s)	64.19	20.72	33.5 - 128	.79 ^a
RAN Objects (s)	54.40	12.14	33.5 - 89	.78 ^a
Vocabulary (T1)	9.13	3.43	1-19	-- ^b
Morphological Knowledge (T1)	65.40	13.98	10 - 93	.87
Syntactic Knowledge (T1)	7.45	2.47	2 - 15	.44
One Minute Reading (T1)	7.53	12.51	0 - 64	.87 ^c
One Minute Reading (T2)	53.71	15.49	7 - 94	.88 ^c
One Minute Pseudoword Reading (T2)	37.16	9.14	5 - 69	.76 ^c
Reading Comprehension (T3)	14.01	3.46	6-20	.74

Czech

Letter Names	8.28	10.51	0 – 45	.95
Letter Sounds	22.90	19.06	0 – 63	.98
Phoneme Isolation	34.15	20.34	0 – 64	.96
Phoneme Blending	2.36	3.71	0 – 11	.95
RAN Colours (s)	57.40	17.23	26.5 – 135.5	.65 ^a
RAN Objects (s)	51.65	11.41	25.5 – 86	.80 ^a
Vocabulary (T1)	10.32	3.06	1-17	-- ^b
Morphological Knowledge (T1)	64.03	16.71	21 – 96	.79
Syntactic Knowledge(T1)	11.70	2.45	5 – 16	.64
One Minute Reading (T1)	3.76	10.41	0 - 87	.91 ^c
One Minute Reading (T2)	50.13	18.38	17 – 120	.87 ^c
One Minute Pseudoword Reading (T2)	37.05	12.31	13 - 75	.87 ^c
Reading Comprehension (T3)	15.40	3.13	7 - 20	.72

Slovak

Letter Names	3.11	5.97	0 - 37	.94
--------------	------	------	--------	-----

Letter Sounds	25.35	19.99	0 -64	.98
Phoneme Isolation	34.95	20.33	0 - 64	.97
Phoneme Blending	3.44	4.05	0 - 11	.95
RAN Colours (s)	57.25	20.14	25 –122.5	.89 ^a
RAN Objects (s)	53.24	12.61	24-87.5	.79 ^a
Vocabulary (T1)	9.97	2.91	1-16	-- ^b
Morphological Knowledge (T1)	65.90	16.08	0 - 96	.78
Syntactic Knowledge(T1)	10.71	2.34	4 - 16	.56
One Minute Reading (T1)	3.18	8.44	0 - 72	.97 ^c
One Minute Reading (T2)	46.63	15.58	18 - 103	.90 ^c
One Minute Pseudoword Reading (T2)	34.06	9.87	2 - 64	.79 ^c
Reading Comprehension (T3)	15.06	2.98	2 - 20	.68

Note: a = correlation between Trial 1 and Trial 2 ; b = estimate not available; c = test-retest correlation.

Table 3

Indirect effect of Word Reading, Phoneme Awareness/Letter Knowledge, RAN and Language at Time 1 on Reading Comprehension at Time 3 through Decoding Efficiency at Time 2

Predictor	English	Spanish	Czech	Slovak
	β (95% CI)	β (95% CI)	β (95% CI)	β (95% CI)
Word reading	.262 (.163, .335)	.054 (.020, .098)	.089 (.021, .148)	.051 (-.009, .105)
Phoneme awareness/letter knowledge	.262 (.163, .335)	.054 (.020, .098)	.089 (.021, .148)	.051 (-.009, .105)
RAN	-.216 (-.307, -.127)	-.125 (-.207, -.054)	-.139 (-.306, -.026)	-.120 (-.284, .004)
Language	-.077 (-.220, .078)	.044 (-.028, .116)	-.053 (-.157, .029)	-.027 (-.118, .050)

Note. Significant effect sizes in **bold**.

Table S1

Features of the reading comprehension measure in English, Spanish, Czech, and Slovak.

	Language			
	English	Spanish	Czech	Slovak
Number of words	345	376	290	284
Mean number of syllables/word (range)	1.24 (1 – 3)	1.84 (1 – 6)	1.97 (1 – 5)	1.93 (1 – 5)
Mean number of letters/word (range)	3.81 (1 - 9)	4.22 (1 – 12)	4.74 (1 – 14)	4.84 (1 – 15)
Average word frequency/million	7817.77 (0.02 –	9098.24 (1 – 45,113)	5963.15 ^a (1 – 41,639)	2815.87 ^a (1 – 30,270)
Primary grades (range)	68,006)			
Average word frequency/million	6826.94 (1 – 48,534)	8665.71 (1 – 38,533)	6371.97 ^a (9 – 45,446)	2843.36 ^a (10 – 29,319)
Grade 2 texts (range)				

Notes: ^a The word frequency counts were obtained from corpora of child-directed texts; in English (Zeno et al., 1999), Spanish (Martínez & García, 2004), Czech and Slovak (Kessler & Caravolas, 2011). The differences in mean word frequencies, especially between Czech and Slovak tests, which are in fact very similar, reflect the distortions caused by small corpus sizes (e.g., Brysbaert & New, 2009). The Slovak corpus in particular is quite small (180,577 tokens; 35,105 distinct word forms, 14,746 lemmas) relative to its Czech counterpart (387,702 tokens; 63,939 distinct word forms, 23,990 lemmas) (Kessler, 2009).

Table S2a

Estimated correlations between all variables. English in the upper diagonal and Spanish in the lower diagonal.

	1	2	3	4	5	6	7	8	9	10	11	12
1 Letter Knowledge Time 1	1	.580**	.632**	-.391**	-.363**	.298**	.440**	.376**	.624**	.697**	.568**	.800**
2 Phoneme Blending Time 1	.583**	1	.575**	-.350**	-.293**	.329**	.422**	.391**	.466**	.492**	.483**	.640**
3 Phoneme Isolation Time 1	.686**	.623**	1	-.422**	-.386**	.298**	.450**	.326**	.476**	.606**	.570**	.675**
4 RAN Objects Time 1	-.371**	-.240**	-.280**	1	.777**	-.302**	-.317**	-.240**	-.383**	-.593**	-.550**	-.456**
5 RAN Colour Time 1	-.188**	-.106	-.121	.639**	1	-.269**	-.313**	-.155*	-.346**	-.509**	-.476**	-.413**
6 Vocabulary Time 1	.361**	.259**	.355**	-.217**	-.078	1	.439**	.420**	.243**	.249**	.266**	.312**
7 Morphol. Knowledge Time 1	.404**	.321**	.432**	-.357**	-.220**	.495**	1	.351**	.312**	.410**	.316**	.474**
8 Syntactic Knowledge Time 1	.379**	.269**	.390**	-.296**	-.134	.486**	.446**	1	.353**	.303**	.275**	.393**
9 Reading comprehension Time 3	.324**	.200**	.326**	-.326**	-.211**	.451**	.401**	.348**	1	.719**	.607**	.618**
10 Word Reading Time 2	.421**	.253**	.322**	-.425**	-.359**	.287**	.253**	.251**	.532**	1	.833**	.739**
11 Nonword Reading Time 2	.290**	.202**	.238**	-.316**	-.342**	.230**	.179*	.182*	.447**	.819**	1	.657**
12 Word Reading Time 1	.797**	.662**	.702**	-.417**	-.202**	.305**	.378**	.402**	.268**	.413**	.296**	1

*Note: * = $p < .05$; ** = $p < .01$*

Table S2b

Estimated correlations between all variables. Czech in the upper diagonal and Slovak in the lower diagonal.

	1	2	3	4	5	6	7	8	9	10	11	12
1 Letter Knowledge Time 1	1	.704**	.751**	-.450**	-.475**	.329**	.352**	.379**	.215*	.557**	.490**	.831**
2 Phoneme Blending Time 1	.708**	1	.659**	-.341**	-.308**	.348**	.371**	.378**	.265**	.418**	.367**	.727**
3 Phoneme Isolation Time 1	.712**	.644**	1	-.464**	-.472**	.344**	.458**	.488**	.286**	.463**	.426**	.657**
4 RAN Objects Time 1	-.307**	-.334**	-.300**	1	.697**	-.205*	-.178*	-.404**	-.239**	-.392**	-.421**	-.361**
5 RAN Colour Time 1	-.355**	-.270**	-.397**	.599**	1	-.142	-.209*	-.340**	-.335**	-.469**	-.483**	-.410**
6 Vocabulary Time 1	.279**	.251**	.243**	-.260**	-.176**	1	.321**	.359**	.200*	.179*	.139	.288**
7 Morphol. Knowledge Time 1	.406**	.417**	.383**	-.226**	-.137	.237**	1	.525**	.218*	.197*	.171*	.383**
8 Syntactic Knowledge Time 1	.236**	.303**	.266**	-.242**	-.154*	.238**	.439**	1	.279**	.234**	.184*	.314**
9 Reading comprehension Time 3	.328**	.388**	.415**	-.221**	-.340**	.353**	.187*	.321**	1	.356**	.417**	.203*
10 Word Reading Time 2	.446**	.358**	.376**	-.249**	-.459**	.143*	.157*	.091	.307**	1	.912**	.548**
11 Nonword Reading Time 2	.321**	.260**	.308**	-.253**	-.484**	.160*	.152*	.078	.373**	.881**		.444**
12 Word Reading Time 1	.732**	.634**	.533**	-.308**	-.307**	.216**	.327**	.202**	.258**	.424**	.311**	1

*Note: * = $p < .05$; ** = $p < .01$.*

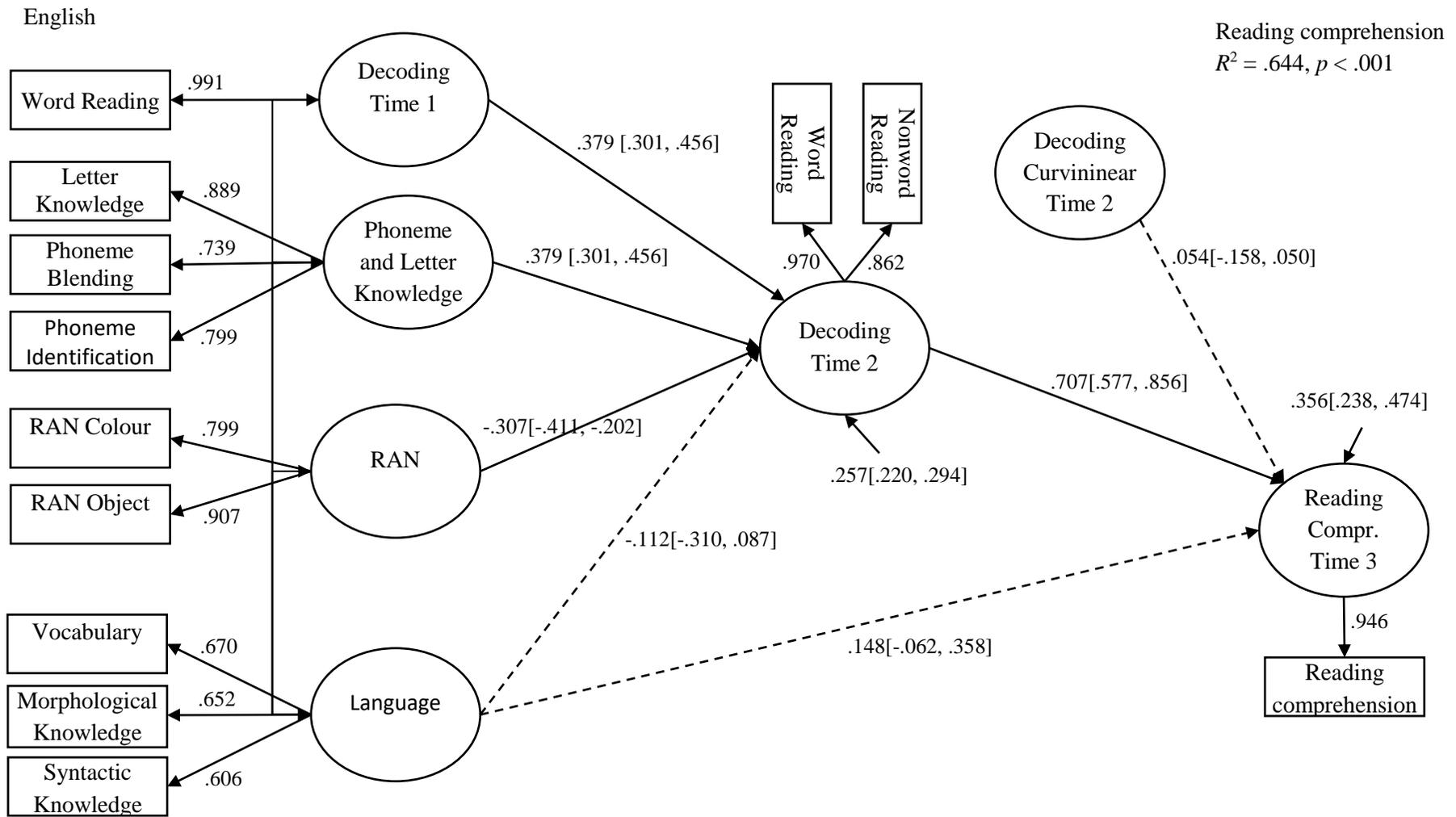


Figure 1a

Czech

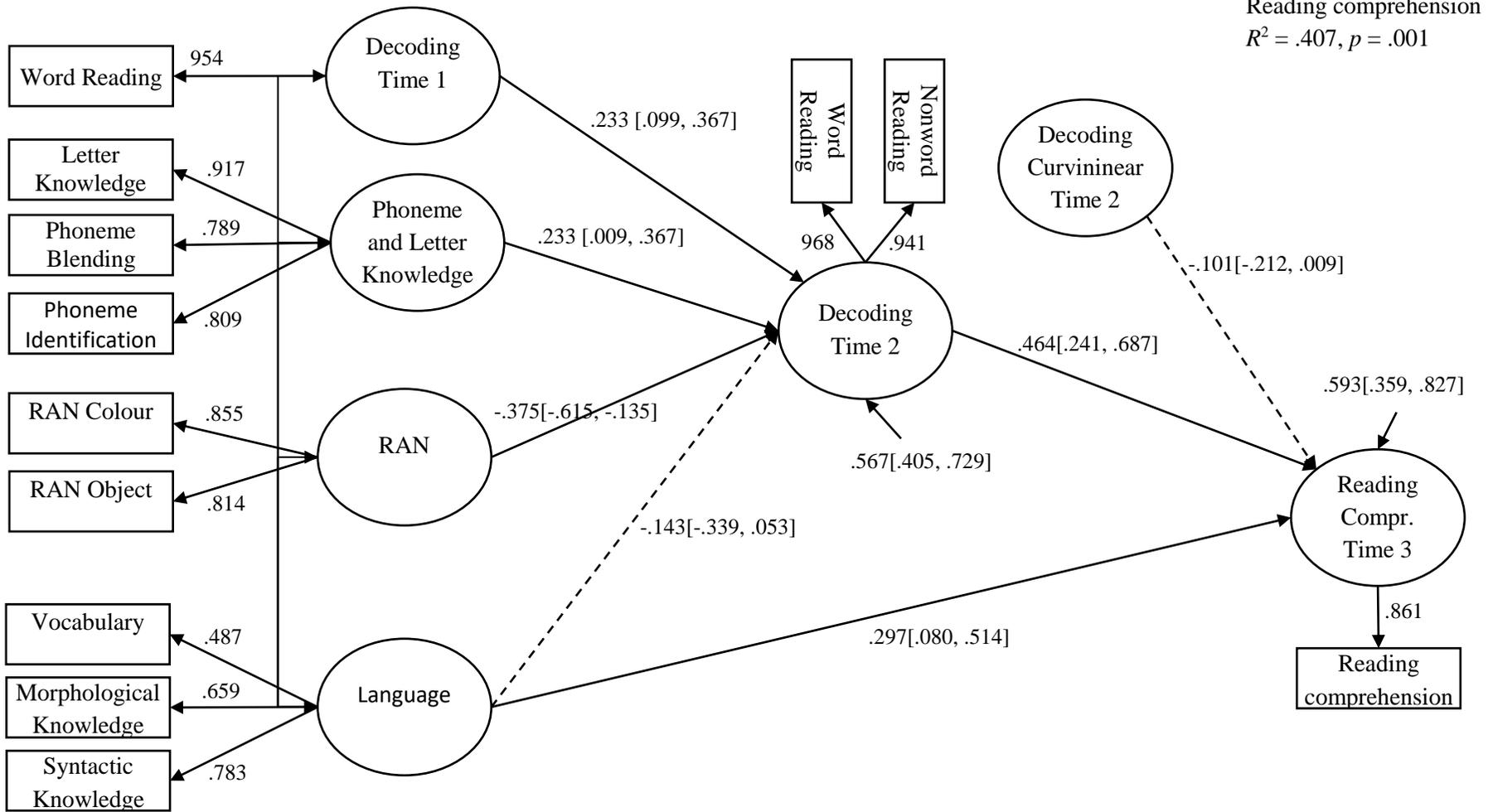


Figure 1c

