Routledge
Taylor & Francis Group

Scientific Studies of Reading

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hssr20

A Meta-analysis of Morphological Awareness Deficits in Developmental Dyslexia

George K. Georgiou, Ana Paula Alves Vieira, Kyriakoula M. Rothou, John R. Kirby, Andrea Antoniuk, Dalia Martinez & Kan Guo

To cite this article: George K. Georgiou, Ana Paula Alves Vieira, Kyriakoula M. Rothou, John R. Kirby, Andrea Antoniuk, Dalia Martinez & Kan Guo (2023) A Meta-analysis of Morphological Awareness Deficits in Developmental Dyslexia, Scientific Studies of Reading, 27:3, 253-271, DOI: 10.1080/10888438.2022.2155524

To link to this article: https://doi.org/10.1080/10888438.2022.2155524

SCIENTIFIC STUDIES OF READING
2023, VOL. 27, NO. 3, 253–271
https://doi.org/10.1080/10888438.2022.2155524

Routledge
Taylor & Francis Group

A Meta-analysis of Morphological Awareness Deficits in Developmental Dyslexia

George K. Georgiou[a], Ana Paula Alves Vieira[a], Kyriakoula M. Rothou[b], John R. Kirby[c], Andrea Antoniuk[a], Dalia Martinez[a], and Kan Guo[d]

[a]Faculty of Education, University of Alberta, Edmonton, Canada; [b]School of Early Childhood Education, Aristotle University, Thessaloniki, Greece; [c]Faculty of Education, Queen's University, Kingston, Canada; [d]School of Mathematical Sciences, Beijing Normal University, Beijing, P. R. China

ABSTRACT

Purpose: We performed a meta-analysis to examine if children with dyslexia experience deficits in morphological awareness (MA) and if the effect sizes are influenced by different moderators (age, aspect of MA measured, type of MA task, language, modality input, semantic knowledge, and selection criteria).

Method: We reviewed 40 studies published in English between January 1990 and August 2021, representing a total of 5,018 participants (age range = 5.1–13.9 years). Studies with adults or English language learners were excluded. There were 49 independent samples in the chronological-age (CA) – dyslexia (DYS) comparison and 18 independent samples in the comparison between DYS and reading–level (RL) controls.

Results: A random-effects model analysis revealed a large effect size for the CA–DYS comparison (g = 1.11) and a non-significant effect size for the RL–DYS comparison (g = −0.08). Age was the only significant moderator of the effect sizes.

Conclusion: These findings suggest, first, that individuals with dyslexia experience significant difficulties in MA and second, that the effect sizes are as large as those reported for phonological awareness, rapid automatized naming, and orthographic knowledge. The lack of a significant RL-DYS difference indicates that MA is not a core causal feature of dyslexia.

Dyslexia, defined as a persistent and unexpected difficulty in developing age- and experience-appropriate word reading skills, affects about 5%-10% of school–age children (Snowling et al., 2020). Several studies have shown that most individuals with dyslexia experience deficits in phonological processing skills such as phonological awareness and rapid automatized naming (RAN; Elliott & Grigorenko, 2014). In their meta-analysis, Melby-Lervåg et al. (2012) found that the average effect size in phonological awareness was $d = -1.37$ (dyslexics performing worse than their chronological-age controls). Likewise, Araújo and Faísca (2019) estimated the average effect size in RAN to be $d = -1.19$. Even though the deficits in phonological awareness and RAN are well documented, much less is known about the role of morphological awareness in dyslexia. Morphological awareness (MA) is defined as "the conscious awareness of the morphemic structure of words and the ability to reflect on and manipulate that structure" (Carlisle, 1995, p. 194). Examining whether individuals with dyslexia experience deficits in MA is interesting in view of conflicting evidence in the literature, some studies reporting a deficit and others reporting no deficit (see Deacon et al., 2008; Deacon et al., 2016, for

CONTACT George K. Georgiou georgiou@ualberta.ca Faculty of Education, University of Alberta, Edmonton-AB, 6-102 Education North, T6G 2G5; Kan Guo guokan@bnu.edu.cn School of Mathematics Sciences, Beijing Normal University, Beijing 100875, P. R. China

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10888438.2022.2155524

reviews). Thus, the main goal of this meta-analysis was to determine whether there is a deficit in MA and secondly, the size of that deficit, should there be one.

Theoretical connections between morphological awareness and reading

According to the Triangle Model of Reading (e.g., Plaut et al., 1996; Seidenberg & McClelland, 1989), phonology, orthography and semantics are three fundamental constituents of word reading and they interact with each other to facilitate word recognition. Kirby and Bowers (2017, 2018) added morphology to the Triangle Model of Reading, as a *binding agent* that connects phonology, orthography, and semantics. The binding agent idea was adopted from Perfetti's lexical quality hypothesis (see e.g., Perfetti, 2007; Stafura & Perfetti, 2017), as something that connects the three components of the Triangle Model and enhances representational quality. The link between morphology and semantics is rather obvious because by definition morphemes are units of meaning, and the link to orthography comes from the highly consistent spelling of morphemes. Morphology is also thought to be linked to phonology because it helps specify the pronunciation of certain graphemes (e.g., the <ea> in *read* and *react*). More recently, Levesque et al. (2021) also proposed the Morphological Pathways Framework according to which morphological processing contributes to the initial, implicit, visual decomposition of morphologically complex words into morphemic sub-units which, in turn, facilitates lexical access.

Morphological awareness deficits in dyslexia

Although there are good theoretical reasons linking morphology to reading, the literature around whether individuals with dyslexia experience MA deficits is mixed. On the one hand, some researchers have shown that individuals with dyslexia experience significant difficulties in MA (e.g., Casalis et al., 2004; Chung et al., 2011; Duranovic et al., 2014; Giazitzidou & Padeliadu, 2022;Joanisse et al., 2000; Chung et al., 2014; Rothou & Padeliadu, 2019). For example, Joanisse et al. (2000) compared English-speaking dyslexic children with chronological-age matched controls. Children were given two orally-presented inflectional MA tasks, namely past tense verb marking (i.e., children were asked to complete a sentence using the past tense of the verb as was done in the example: *Sally bakes a cake: Sally baked a cake; Sally drives a car: Sally . . . . a car*) and generation of the plural form of nouns (i.e., children were asked to complete a sentence: *Here is a fish. Now there are two of them. There are two . . .* [pause for response]). Individuals with dyslexia performed significantly below their peers in both tasks.

Three hypotheses have been proposed to explain deficits in MA. First, children with dyslexia may experience difficulties in MA because of underlying phonological processing deficits, namely limited ability to perceive or manipulate phonological information of morphemes (Deacon et al., 2008). Some studies have shown that individuals with dyslexia perform lower than chronological-age controls in MA tasks that also require phonological processing (i.e., phonological change in the derived forms or past-tense marking) (see e.g., Casalis et al., 2004; Robertson et al., 2013). Second, MA deficits may be related to deficits in broader oral language skills. Joanisse et al. (2000), for example, compared two subgroups of children with dyslexia with their reading-level matched controls. For the children identified as having phonological dyslexia, their deficits were isolated to their phonological awareness skills whereas the children with more global spoken language impairments had difficulties in vocabulary and word structure (morphology). Group comparisons further showed that although both subgroups of dyslexic children had difficulties in production of inflected nouns and verbs, only the language impaired dyslexic children underperformed the reading-level matched controls. Finally, individuals with dyslexia may have MA deficits because of underlying deficits in word reading. This could either be because the stimuli are written or because the word reading deficit limits their exposure to new words because they do less reading. This third possibility is based on the idea that after about Grade 4 new words are more likely to be encountered in print. To build up good representations of morphemes, either orally or in print, frequent exposure to these units in different words is required.

Furthermore, some studies have reported a bidirectional relationship between children's MA skills and word reading (e.g., Deacon et al., 2013; Kruk & Bergman, 2013). Assuming this is true, then deficits in MA might be a product of deficits in word reading.

On the other hand, some researchers have argued that morphological awareness may not be deficient in individuals with dyslexia (e.g., Cavalli et al., 2017; Elbro & Arnbak, 1996; Law et al., 2015; Quémart & Casalis, 2015; Suárez-Coalla et al., 2017; see alsoDeacon et al., 2008, for a review). Elbro and Arnbak (1996), for example, suggested that adolescents with dyslexia use recognition of root morphemes as a compensatory strategy in reading words. In their study, they compared adolescents with and without dyslexia in reading text in different conditions (e.g., one syllable at a time, one morpheme at a time). Adolescents with dyslexia performed significantly better in the morpheme condition than in the syllable condition, with the adolescents without dyslexia presenting the opposite pattern. Suárez-Coalla et al. (2017) also showed that Spanish-speaking children with dyslexia benefited as much as their chronological-age controls from the presence of high frequency base morphemes in reading and spelling of unfamiliar items. The idea that MA skills may compensate for the weak phonological skills of children with dyslexia is also supported by the psycholinguistic grain-size theory (Ziegler & Goswami, 2005). According to Ziegler and Goswami (2005), the availability of a specific processing unit during word reading depends on the orthographic consistency and the availability of the spoken unit in oral language. Arguably, to bypass their underlying phonological difficulties during word reading, individuals with dyslexia may resort to larger and more consistent grain-size units such as morphemes (see Kotzer et al., 2021, for a similar argument).

Moderators

When research findings regarding the presence or not of a deficit are as diverse as the ones for MA, it is reasonable to expect significant variability in the produced effect sizes. This would require further investigation of the role of possible moderators. For the purpose of this meta-analysis, we examined the role of seven moderators that are described in more detail below.

Age

As children get older, they encounter more unfamiliar, morphologically-complex words. Assuming earlier reading ability predicts future MA skills (Deacon et al., 2013), the limited reading skills of individuals with dyslexia should negatively impact the development of their MA skills. Thus, as children grow older the difference between children with dyslexia and chronological-age controls should widen.

MA task content

There are three types of morphological combinations that appear in MA tasks: inflections, derivations, and compounds (Kuo & Anderson, 2006). Derivational morphology appears to be more challenging for children than inflectional morphology because it is often accompanied by phonological (e.g., *heal–health*), orthographic (e.g., *try–tried*), semantic (e.g., *safe–unsafe*) or phonological plus orthographic (e.g., *five–fifth*) changes (Deacon et al., 2008). Given that inflectional morphology is more consistent, it might be easier for children including those with dyslexia, and, as a result, group differences in inflectional morphology might be smaller than in derivational morphology. Finally, there are three categories of MA task content used in Chinese: (1) compounding, (2) homophone, and (3) homograph (Liu & McBride-Chang, 2010). It remains unclear if Chinese children with dyslexia experience more severe deficits than controls in one or more of these categories.

Type of MA task

The MA tasks can be classified into the following categories: production tasks, judgment tasks, blending or segmenting tasks, and word/sentence analogy tasks. Production MA tasks require children to generate a new word (i.e., derivational MA or compounding MA) or a different form of the same word (i.e., inflectional MA) by applying specific morphological rules. In turn, judgment tasks require children to choose which of two or three options best completes a sentence (e.g., *Farm. My uncle is a* ____ (farmer/farming), Singson et al., 2000). In the blending or segmenting tasks, children are asked to combine morphemic units (i.e., base, suffix, or prefix) or to identify the correct morphemic unit given a derived word or an inflected word (Apel, 2014). Finally, the word/sentence analogy tasks are also a form of production tasks, because participants must recognize the relationship between words or sentences and then produce a new word. Because production tasks are more difficult in general, they may be particularly difficult for children with dyslexia, whose limited skills may allow them to recognize a correct answer or eliminate some incorrect options, but not allow them to generate a novel response.

Language

The ways in which morphology is used in different languages is different and this may account (at least partly) for the mixed results reported in the literature. For example, English is morphophonemic and prioritizes communicating morphology through the orthography. In contrast, Finnish is a transparent orthography that uses lots of compounding and Hebrew is root-derived. Finally, in Chinese, most words are compound words with two or more morphemes. Given that there is no agreed-upon dimension of morphological complexity (seeBorleffs et al., 2017, for a discussion), we decided to perform pairwise comparisons with English being the point of reference in all comparisons (i.e., English vs. Chinese; English vs. Finnish/Greek; English vs. French; and English vs. Arabic/Hebrew) because it can potentially reveal meaningful differences between languages.

Modality of input/output

MA tasks can be presented either orally or in writing (i.e., visually). Likewise, children's responses might be either oral or written (Apel, 2014). Oral MA tasks might be easier for children with dyslexia than written tasks because in oral tasks dyslexic children do not need to read words. Thus, we would expect larger differences between groups in MA tasks that are presented in writing or require a written response.

Semantic control

Because morphological awareness is intricately related to vocabulary knowledge (e.g., Spencer et al., 2015), whether or not vocabulary (the most common indicator of semantic knowledge) is assessed and matched across groups might be an important moderator. In general, we would expect larger MA deficits in studies in which vocabulary was not matched between children with dyslexia and their chronological-age controls.

Selection criteria

Participants with dyslexia are usually selected using two criteria. One criterion is the presence of an official diagnosis of dyslexia by clinicians (e.g., Casalis et al., 2004) and the other is through screening with standardized reading and/or spelling tasks (e.g., Vender et al., 2017). The criterion used to classify dyslexics might have an effect on the nature of MA deficits because children with an official diagnosis of dyslexia usually receive reading interventions in schools (some of which may include morphological awareness activities; see e.g., Bowers et al., 2010). Assuming the intervention is effective (e.g., Bar-Kochva et al., 2020), these children may have less pronounced difficulties than their peers who do not have such diagnosis. However, it is also possible that the formally diagnosed children have more profound difficulties.

The present study

In view of conflicting evidence regarding the presence of MA deficits in dyslexia (e.g., Casalis et al., 2004; Chung et al., 2010; Elbro & Arnbak, 1996; Law et al., 2015; Rothou & Padeliadu, 2019), we performed a meta-analysis. We aimed to answer the following two questions:

(1) Do children with dyslexia experience a deficit in MA compared to their chronological-age and reading-level matched controls, and, if yes, what is the size of the deficit?
(2) Do age, aspect of MA measured, type of MA task, language, modality input/output, semantic control, and selection criteria influence the size of the effects?

Method

Data collection

To find the studies that compared the MA (defined as the explicit awareness of morphological information and the ability to reflect on and manipulate this information, Carlisle, 1995) of children with and without dyslexia, the second and third authors first conducted a computerized database search in PsycINFO, ProQuest Educational, PubMed, Medline, ERIC, and Scopus. Publications, including master's theses and doctoral dissertations, between the years 1990 and August 2021 with terms related to dyslexia (*dyslexia* OR *dyslexic(s)* OR *reading difficulty(ies)* OR *learning disability(ies)* OR *reading disorder* OR *poor readers* OR *at risk readers*) crossed with terms related to morphological awareness (*morphology* OR *morphological knowledge* OR *morphological awareness* OR *morphological sensitivity* OR *morphological analysis* OR *morphological processing* OR *morphological derivation* OR *morpheme* OR *morpheme production* OR *morpheme judgment* OR *inflectional morphology* OR *past-tense morphology* OR *derivational morphology* OR *compound morphology* OR *compound* OR *compound awareness* OR *compounding*) were used as part of our initial search. We started our search from 1990 because the first studies on dyslexia and morphological awareness were published in the early 90′s.

Moreover, we reviewed the reference lists of previous meta-analyses on dyslexia (e.g., Araújo & Faísca, 2019; Kudo et al., 2015; Melby-Lervåg et al., 2012; Parrila et al., 2020; Peng et al., 2017) and dissertations on dyslexia and MA (e.g., Caglar-Ryeng, 2010). In addition, journals publishing studies on dyslexia (e.g., *Annals of Dyslexia, Dyslexia, Journal of Learning Disabilities, Learning Disabilities Quarterly, Reading and Writing, Scientific Studies of Reading, Journal of Experimental Child Psychology*, and *Journal of Research in Reading*) were hand searched. This first round of search helped us identify 640 studies. Two hundred and eighty-nine studies were excluded after reviewing their abstracts either because the study was a duplication of another study or because the study was described as qualitative, a case study, or a review. The interrater reliability for this part of the screening process was .99. The remaining 351 studies were subsequently perused considering the following inclusion criteria:

(1) The study reported empirical data on morphological awareness.
(2) The study included results that reflected the performance of individuals with dyslexia compared to those of typical readers matched either on chronological age and/or reading level.
(3) The study included participants who either had a formal diagnosis of dyslexia (i.e., a diagnosis made by a professional) or were identified based on standardized reading assessments.
(4) The study reported either the effect sizes or contained enough information to compute them. Note here that we contacted four authors whose work could have been considered in our meta-analysis had we had more information, but none of them replied to our e-mail.

Based on these criteria, 234 studies were further removed. Before finalizing the list of studies to be included in the analyses, we further applied the following exclusion criteria:

(1) Participants had a familial risk of dyslexia (e.g., Torppa et al., 2010), but no formal diagnosis ($n = 7$).
(2) Participants were described as learning disabled without any additional information if their learning disability was specific to reading (e.g., Sümer Dodur, 2021; $n = 9$). We contacted the authors of these studies but we did not receive any response.

(3) Participants were identified on the basis of low performance in tasks other than word reading. For example, Abu-Rabia et al. (2003) screened children for reading disabilities using a reading comprehension task (n = 6).
(4) Participants were described as reading disabled or poor readers, but a rather lenient cutoff score (e.g., a cutoff above the 16th percentile) was used to select them (e.g., Carroll & Breadmore, 2018; n = 20).
(5) Participants were adults with dyslexia or English Language Learners with dyslexia (e.g., Caglar-Ryeng, 2010; n = 11).
(6) Participants were individuals with dyslexia and another disability, but data were not disaggregated by groups (e.g., Wong et al., 2010, had a group with a specific language impairment and dyslexia) (n = 3).
(7) The study used only spelling to dictation or reading aloud tasks to measure morphological awareness (e.g., Diamanti et al., 2014; n = 21). Assuming children with dyslexia are selected based on low performance in word reading and spelling, then including these measures would artificially inflate the differences with the control groups.

Finally, publications from the same author(s) were checked to ensure no duplicated datasets were included in the meta-analysis. Our final sample consisted of 40 studies (total of 5,018 participants). Totals of 49 and 18 unique samples were identified for the chronological-age (CA) and reading-level (RL) controls, respectively (see Figure 1 for the PRISMA flow chart).

Recorded variables and coding procedure

All studies from our final sample were coded independently by the third author, who has experience conducting research in morphological awareness, and the fifth author, who is a doctoral student with experience in coding studies for meta-analyses. Each coder entered the data in two separate spreadsheets, one for the DYS-CA groups and one for the DYS-RL groups. The consensus rate between the coders was 98.02% for the DYS-CA coding and 98.13% for the DYS-RL coding. The few discrepancies were resolved after revisiting the original studies and after discussing the recorded data with the first author.

For each study in the final sample, we recorded information on the following moderators: (a) age, (b) MA task content, (c) type of MA task, (d) language, (e) modality of input/output, (f) semantic knowledge, and (g) selection criteria. With the exception of age which was a continuous moderator, all other moderators were categorical (see Appendices A and B in Online Supplemental Materials, for detailed information on the moderators).

Age

We coded the mean age of the samples in years and months. Age was coded only if the study provided it. Six studies did not provide information on the mean age of their participants and so were not included in that moderator analysis. The age ranged from 5 years and 1 month to 13 years and 9 months.

MA task content

This moderator included three categories: (1) derivation, (2) inflection, and (3) compounding (Deacon et al., 2016). Inflection refers to a change in grammatical features (e.g., adding the suffix *-s* changes the number in car-car<u>s</u>; adding the suffix *-ed* changes the tense in watch-watch<u>ed</u>). Inflectional morphology has most often been operationalized with sentence completion tasks that provide great context (e.g., *This is a car. Now there is another one. There are two [cars]*, see, Maynard et al., 2018, for a review). In turn, derivational morphology often comprises changes in grammatical category and/or meaning (e.g., adding the suffix *-ness* changes the adjective happy to the noun happi<u>ness</u>; adding the prefix *un-* changes the meaning to its opposite, <u>un</u>happy). Derivational morphology tasks include

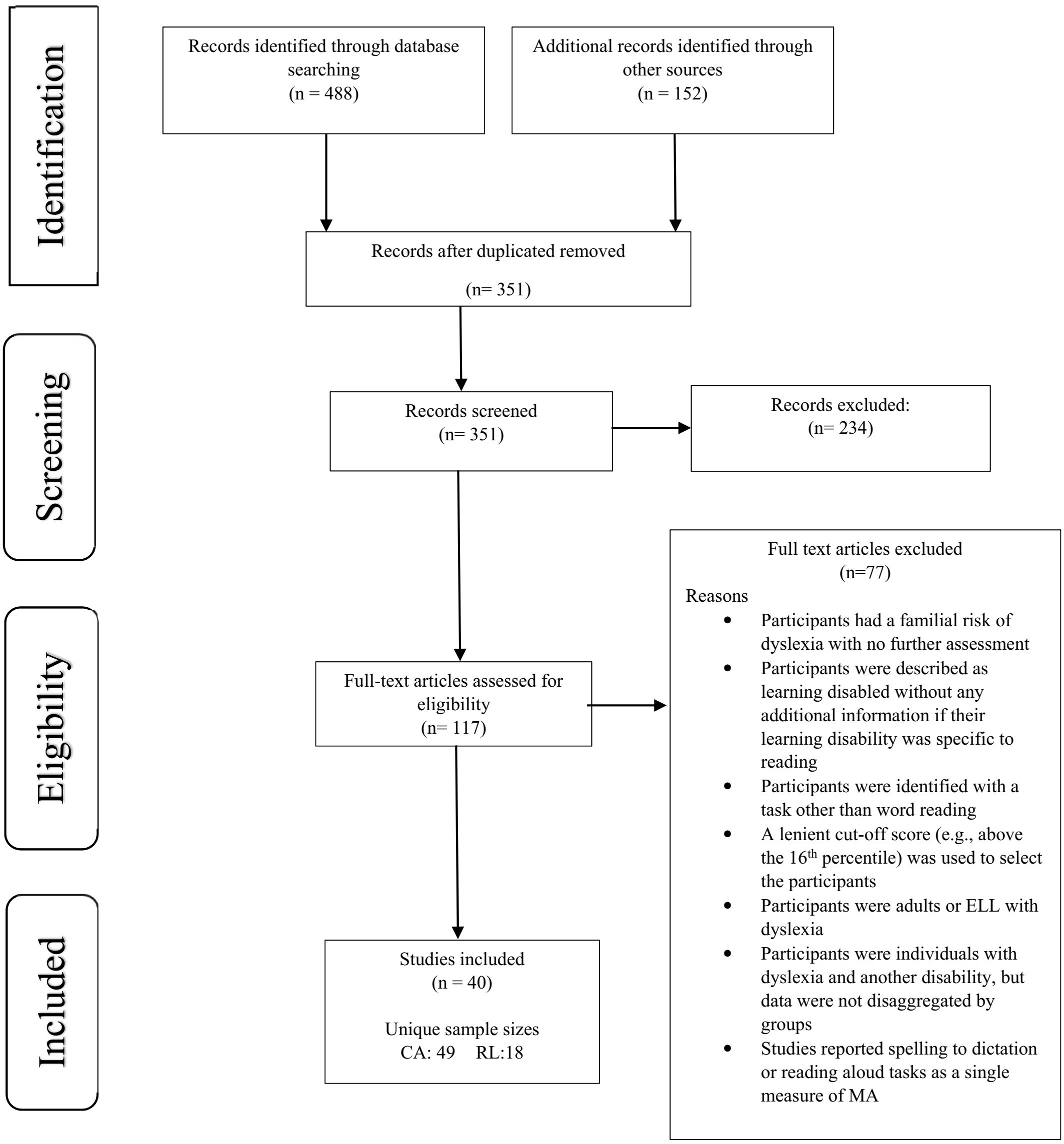

Figure 1. Flow diagram for the search and inclusion of studies. Note: CA = Chronological-age controls; RL = Reading-level controls; MA = Morphological awareness.

sentence completion (e.g., given the cue word *farm*, complete the sentence: *My uncle is a [farmer]*) or word analogy (e.g., long: length:: strong: strength, see Duncan et al., 2009, for further examples). Finally, compounding is the process of joining two roots to create a word both orthographically and semantically accepted (examples of compound words are: basketball, airplane, birthday).

For studies conducted in Chinese, we further coded the MA tasks in three categories: (1) compounding (e.g., *when we see the sun rising in the morning, we call it "sunrise," what would we then call the moon rising in the evening?[moonrise]*), (2) homophone (e.g., the experimenter would orally present the participants with the words 首先 "[shou xian]," meaning "at first," and 守护 "[shou hu]," meaning "protect," and then ask them to judge whether the words had similar meanings), and (3) homograph (e.g., the experimenter would show a word 草地 (cao di, lawn) and then ask children to produce two more words using the target morpheme 草 (cao), one sharing the same meaning with the target word and one having a different meaning from the target word). In contrast to the first two categories that involved oral material, homographs could be presented in an oral or written form.

Type of MA task

This moderator had four categories: (1) production (e.g., children were asked to complete a sentence with a base word given a derived word, as in *politeness*/ this boy is . . . [*polite*], Casalis et al., 2004; see also Duranovic et al., 2014, for an example of production on inflectional morphology),[1] (2) judgment (e.g., children were asked to select the correct word missing in the sentence, "*She hoped to make a good* ______ *[A. impressive/B. impressionable/C. impression/D. impressively]*, Siegel, 2008), (3) blending (e.g., children were provided with a base word and an affix and were instructed to pronounce the resulting word after joining the two parts) or segmenting (e.g., children were asked to segment words into as many meaningful parts as possible, see Casalis et al., 2004; Layes et al., 2017; or to identify the base of a suffixed word, see, Berthiaume & Daigle, 2014), and (4) word analogy (e.g., *teach: teacher, drive*: ___ [*driver*], see Apel, 2014; Kirby et al., 2012, for further examples).

Language

We created four contrasts with English being our point of reference in each pairwise comparison. We had four comparisons: English vs. Chinese, English vs. Finnish/Greek, English vs. French, and English vs. Arabic/Hebrew. One study in Bosnian and one in Italian were excluded from this analysis.

Modality of input/output

Presentation format included three categories: (1) oral (tasks were presented orally), (2) visual (tasks were presented visually, e.g., children were asked to circle or underline roots from a written list), and (3) both (e.g., participants were visually presented with a sentence that was also read to them). Because in most studies the presentation of the task and the children's response were in the same modality, we decided to calculate the effect sizes only for modality of input. Only in seven studies was the presentation visual but the response oral (i.e., Chung et al., 2010, 2011, 2013; Chung et al., 2014; Kalindi & Chung, 2018; Meng et al., 2019; Vender et al., 2017). We did not find any studies in which the instructions were given orally, but the response was not oral (e.g., students were not required to circle or underline a word, for example).

Semantic control

This moderator had three categories: (1) vocabulary was assessed and the groups were selected to have the same vocabulary scores (6 studies), (2) vocabulary was assessed but the groups differed on vocabulary scores (14 studies), and (3) vocabulary was not assessed (23 studies). Of the studies that included a measure of vocabulary, 11 used a measure of expressive vocabulary, six a measure of receptive vocabulary, one used both, and one did not provide a description of the vocabulary task.

Selection criteria

The way individuals with dyslexia were selected was coded into two categories: (1) individuals were selected because they had a formal diagnosis of dyslexia from an appropriate clinician, or (2) individuals with dyslexia were identified following a screening process.

Statistical analysis

The *metafor* package for the *R* statistical program (Viechtbauer, 2010) was used to perform the analyses. Effect sizes for studies involving group comparisons were computed with Hedge's *g*. All effect sizes used in analyses were from independent samples. When multiple measures were reported for one sample, an averaged effect size was used.[2] Overall effect size and 95% CI were estimated by calculating a weighted average of individual effect sizes in a random-effects model, which assumes that variation between studies can be systematic and not simply due to random error. Whether or not the overall effect size differed from zero was tested with a z test. For studies including both CA and RL control groups, a separate effect size was calculated for each of the two comparisons. Forest plots were used to present the distributions of the effect sizes visually.

To examine the variation in effect sizes between studies, the Q test of homogeneity was used (Hedges & Olkin, 2014). A significant value on this test indicates a reliable variability between the effect sizes in the sample of studies. I^2 is the proportion of total variation between effect sizes that is caused by real heterogeneity rather than by chance, which was used to determine the magnitude of the heterogeneity. Moderator variables were also explored as potential sources of variance in the effect sizes. When multiple measures were reported for one sample or one coded test type, an average effect size was used. Linear models were used to predict the study's outcome from the moderator variables, both for the continuous and categorical moderators. The degree of difference between the subsets of studies was tested with a Q test and by comparing the correlation magnitude with CIs between the study subsets. When categorical variables were used as moderators, we did not do the dummy coding manually. Instead, we used a formula together with a *factor* function to let the *metafor* package handle the dummy coding automatically. If the moderator was found to be statistically significant, we further used an ANOVA approach to test for all pairwise differences.

Publication bias

To test for publication bias, we first computed Rosenthal's Fail-Safe N and we also conducted the Rank Correlation and Egger's Regression tests. These tests examine the relationship between the size of the effects from each study and the associated standard error. If the results of the Rank Correlation and Egger's Regressions tests are significant, this suggests possible publication bias. Furthermore, funnel plots were created to assess whether the studies were distributed asymmetrically around the mean effect size, which may also indicate the presence of publication bias (Borenstein et al., 2009). In the funnel plot, the standard error is plotted on the y axis and the effect size on the x axis. In the absence of publication bias, this plot should be expected to form an inverted funnel. In the presence of bias, the funnel will be asymmetric. Finally, to examine the impact of studies that might be missing from the analysis, the "trim and fill" method for random-effects models was used (Duval & Tweedie, 2000). The "trim and fill" method imputes values in the funnel plot to make it symmetrical and calculate an estimated overall effect size on this basis. Considering the relatively low agreement between various publication bias approaches, we followed Lin et al.'s (2018) recommendation and we performed multiple tests as mentioned above to examine for possible publication bias.

Results

Meta-analytic results

The random-effects model demonstrated that the overall mean effect size of differences between the CA and DYS groups was significant (see Table 1 and Figure 2 for the forest plot). The overall mean effect across the 49 effect sizes was large, $g = 1.11$ ($z = 13.921$, $p < .001$, 95% CI = [0.957, 1.271]), favoring the CA group. For the RL-DYS comparison, the overall mean effect size ($g = -0.08$) was not significant (see Table 1 and Figure 3 for the forest plot). The heterogeneity analysis further showed that the variation between studies was significant for both the CA-DYS ($Q = 223.228$, $I^2 = 78.58\%$, $p < .001$) and the RL-DYS ($Q = 46.404$, $I^2 = 66.75\%$, $p < .001$) comparisons.

Similar results were obtained when using robust variance estimation (RVE) meta-regression including all possible effect sizes instead of only averaged effect sizes for separate samples. More specifically, when we reran the analyses using robumeta, the overall mean effect for the CA-DYS and RL-DYS comparisons was $g = 1.12$ ($p < .001$, 95% CI = [0.965, 1.300]) and $g = -0.075$ ($p = .826$, 95% CI = [−0.265, 0.214]), respectively.

Table 1. Meta-analytic results: overall standardized mean differences for the control and dyslexic groups.

								Heterogeneity		
Comparison	*k*	*n*	g^a	S.E.	*Z* value	*p* value	95% CI	I^2 (%)	*Q*	*p* value
CA-DYS	49	CA: 2840 DYS: 1772	1.114	0.080	13.921	<0.001	[0.957, 1.271]	78.58	223.228	<0.001
RL-DYS	18	RL: 423 DYS: 406	−0.086	0.128	−0.673	0.500	[−0.337, 0.164]	66.75	46.404	<0.001

[a]. A positive effect size indicates that CA controls performed better than DYS. A negative effect size indicates that RL controls performed worse than DYS. *k* = number of samples; *n* = total sample size; *g* = estimated Hedge's *g* in random-effects model; I^2 = the proportion of total variation caused by real heterogeneity; *Q* = Hedge's *Q* test of homogeneity.

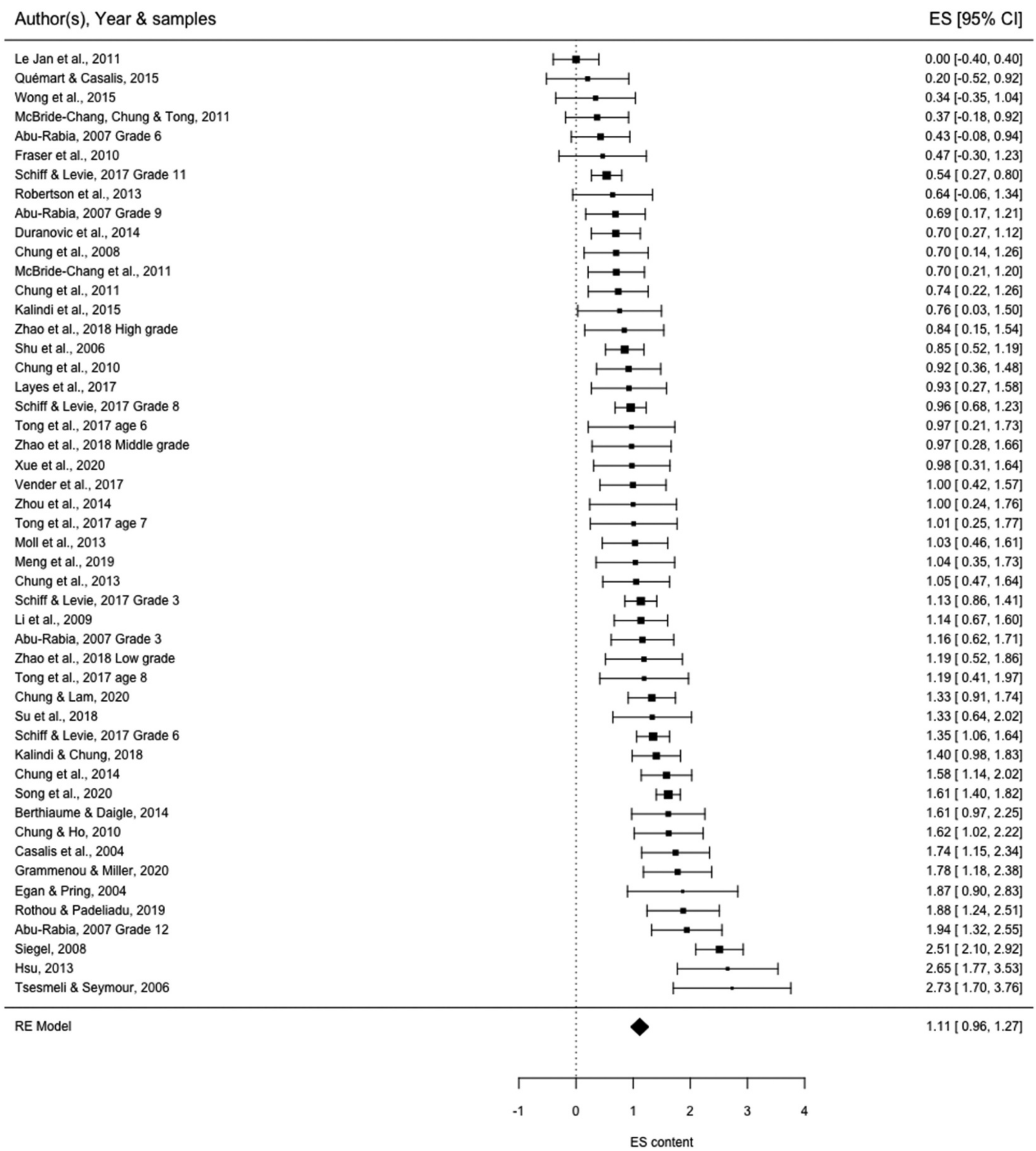

Figure 2. Forest plot: Strength of the standardized mean differences between CA and DYS groups.

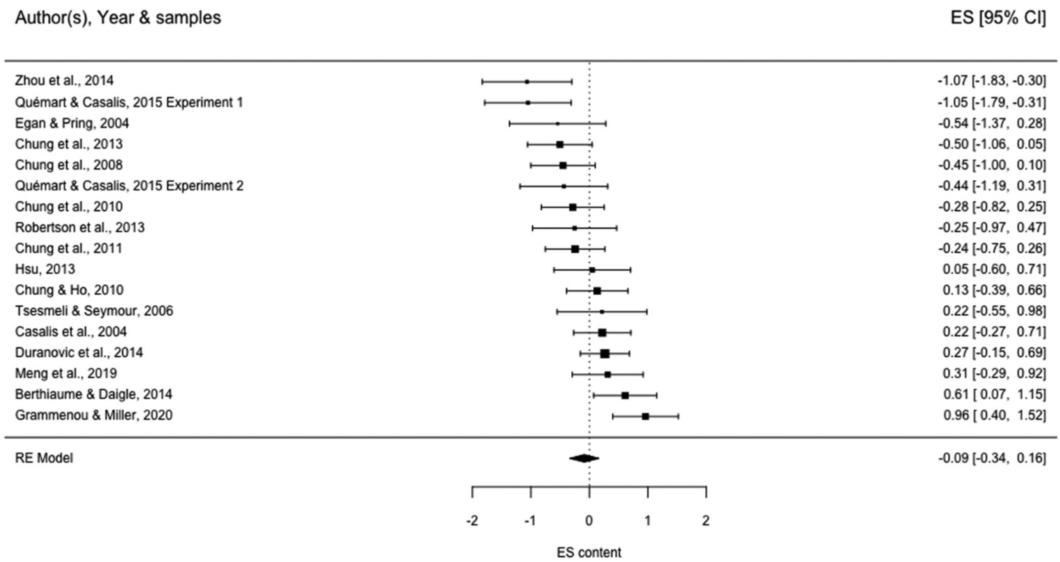

Figure 3. Forest plot: Strength of the standardized mean differences between RL and DYS groups.

Moderator analyses

The results of the moderator analyses are presented in Table 2 for the CA-DYS group comparisons and in Table 3 for the RL-DYS group comparisons. The only significant moderator in the CA-DYS comparisons was age. More specifically, the effect size was larger in studies with older children ($\beta = 0.095$, $p = .026$). As shown in Table 3, none of the other moderators explained the variability in the effect sizes in the RL-DYS comparisons.

Publication bias

The estimated overall effect size for the CA-DYS comparison was reliable. The estimated number size ($n = 15{,}649$) in the Fail-Safe N analysis suggested the robustness of the significant result(see Table 4). Next, we performed the Rank Correlation and Egger's Regression tests. As suggested by the Rank Correlation test, the Kendall's taus for the CA-DYS comparison (tau = 0.083, $p = .405$) and the RL-DYS comparison (tau = −0.323, $p = .076$) were not significant. The Egger's Regression test was not significant for the CA-DYS comparison ($z = 1.051$, $p = .293$), but was significant for the RL-DYS comparison ($z = -2.189$, $p = .028$). Finally, the "trim and fill" analyses showed that there were no missing studies in either side of the mean effect size (see funnel plots in Figure 4) for both the CA-DYS and the RL-DYS comparison. Viewed together, the results of these analyses suggest that there was no publication bias.

Discussion

Group differences

The main purpose of this meta-analysis was to examine if children with dyslexia experience deficits in MA when compared to both chronological-age and reading-level matched controls, and the possible effect of different moderators (e.g., age, type of MA task, language, semantic knowledge). Our findings revealed first a large effect size ($g = 1.11$; Cohen, 1988) when children with dyslexia were compared to their CA controls, which suggests that the MA skills of children with dyslexia are much lower than those of their same–age peers. The size of the deficits is comparable to that reported for rapid

Table 2. Results of moderator analyses in morphological awareness for the CA and DYS groups.

Moderator variable	Number of effect sizes (*k*)	g^a	*p* value	95% CI	Difference in *g* (highest-lowest category)	Significance test (*Q* or *β*)	*p* value
1. Mean age	34/34	–	–	–		0.095	0.026
2. MA task content					0.252	1.370	0.504
Compounding	26	1.086	<0.001	[0.869, 1.302]			
Derivation	10	1.314	<0.001	[0.964, 1.663]			
Inflection	9	1.062	<0.001	[0.712, 1.412]			
3. Type of MA task I					0.160	0.687	0.876
Production	29	1.237	<0.001	[1.013, 1.461]			
judgment	20	1.146	<0.001	[0.872, 1.420]			
Blending or Segmenting	4	1.100	0.001	[0.487, 1.713]			
Word analogy	13	1.077	<0.001	[0.720, 1.433]			
Type of MA task II (in Chinese)					0.111	0.358	0.835
Compounding	18	1.161	<0.001	[0.919, 1.402]			
Homophone	2	1.110	0.001	[0.409, 1.810]			
Homograph	13	1.049	<0.001	[0.776, 1.323]			
4. Language					0.657	4.783	0.310
English	6	1.517	<0.001	[1.034, 2.001]			
Chinese	26	1.086	<0.001	[0.868, 1.303]			
Arabic/Hebrew	9	1.003	<0.001	[0.657, 1.348]			
Finnish/Greek	3	1.405	<0.001	[0.778, 2.032]			
French	4	0.860	0.002	[0.308, 1.412]			
5. Modality of input/ output					0.317	1.383	0.500
Visual	12	1.254	<0.001	[0.933, 1.575]			
Oral	38	1.112	<0.001	[0.936, 1.288]			
Both	6	0.936	<0.001	[0.508, 1.364]			
6. Semantic knowledge					0.317	2.318	0.313
Vocabulary not assessed	26	1.213	<0.001	[0.994, 1.431]			
Matched on vocabulary	9	0.896	<0.001	[0.548, 1.243]			
Assessed but not matched	14	1.092	<0.001	[0.792, 1.391]			
7. Selection Criteria					0.093	0.309	0.577
Following screening	20	1.061	<0.001	[0.809, 1.314]			
Former diagnosis	28	1.155	<0.001	[0.943, 1.367]			

[a]. A positive effect size indicates that CA controls performed better than DYS. *k* = number of effect sizes; *g* = estimated Hedge's *g* for subsets of studies belonging to different categories of the moderator variable; *Q* = significant *Q* test value for categorical variables; *β* = regression coefficient in meta regressions for continues variables.

automatized naming (Araújo & Faísca, 2019) and orthographic knowledge (Georgiou et al., 2021), but somewhat lower to that of phonological awareness (Melby-Lervåg et al., 2012). When our findings are viewed together with those of previous meta-analyses (e.g., Georgiou et al., 2021; Kudo et al., 2015; Melby-Lervåg et al., 2012; Parrila et al., 2020), they suggest that individuals with dyslexia experience profound deficits in all three levels of the Triangle Model of Reading (Seidenberg & McClelland, 1989), and their "binding agent" (morphology; Kirby & Bowers, 2017). They also challenge the idea that children with dyslexia have relatively intact MA skills, which they can use to bypass their weaknesses in phonological processing. It is fair to say that the deficits in MA are not as large as deficits in phonological awareness, but they are still large.

It is important to stress that in contrast to the analyses with CA–matched controls, the DYS group did not differ significantly from the RL–matched controls (g = −0.08). RL–match designs are commonly used to test assumptions of causality following the logic that if the poor readers perform poorer than their RL-matched controls on task A assessing construct B, then construct B is a potential

Table 3. Results of moderator analyses in morphological awareness for the RL and DYS groups.

Moderator variable	Number of effect sizes (*k*)	g^a	*p* value	95% CI	Difference in *g* (highest-lowest category)	Significance test (*Q* or *β*)	*p* value
1. Mean age (RL/DYS)	17/17	–	–	–	–	0.011/ -0.031	0.879/ 0.638
2. MA task content					0.455	3.210	0.200
Compounding	8	−0.239	0.173	[−0.583, 0.104]			
Derivation	7	0.216	0.254	[−0.155, 0.588]			
Inflection	3	−0.134	0.651	[−0.716, 0.448]			
3. Type of MA task I					0.811	4.101	0.250
Production	13	−0.092	0.517	[−0.373, 0.188]			
judgment	7	0.037	0.838	[−0.322, 0.396]			
Blending or Segmenting	2	0.534	0.106	[0.115, 1.184]			
Word analogy	4	−0.277	0.293	[−0.796, 0.240]			
Type of MA task II (in Chinese)					0.481	2.520	0.283
Compounding	6	−0.361	0.009	[−0.633, −0.089]			
Homophone	1	0.119	0.748	[−0.613, 0.852]			
Homograph	4	−0.087	0.588	[−0.406, 0.230]			
4. Language					0.159	0.289	0.865
English	3	−0.185	0.551	[−0.797, 0.426]			
Chinese	8	−0.238	0.157	[−0.568, 0.092]			
French	5	−0.078	0.748	[−0.557, 0.400]			
5. Modality of input/output					0.428	1.784	0.409
Visual	3	0.228	0.439	[−0.350, 0.807]			
Oral	16	−0.195	0.131	[−0.449, 0.058]			
Both	3	−0.199	0.465	[−0.737, 0.337]			
6. Semantic knowledge					0.436	2.968	0.226
Vocabulary not assessed	9	0.112	0.505	[−0.219, 0.445]			
Matched on vocabulary	4	−0.323	0.199	[−0.817, 0.170]			
Assessed but not matched	5	−0.314	0.247	[−0.846, 0.218]			
7. Selection Criteria					0.267	1.079	0.298
Following screening	7	−0.241	0.219	[−0.625, 0.143]			
Former diagnosis	11	0.026	0.874	[−0.300, 0.352]			

[a]. A positive effect size indicates that RL controls performed better than DYS. A negative effect size indicates that RL controls performed worse than DYS. *k* = number of effect sizes; *g* = estimated Hedge's *g* for subsets of studies belonging to different categories of the moderator variable; *Q* = significant *Q* test value for categorical variables; *β* = regression coefficient in meta regressions for continues variables.

cause for dyslexia (e.g., Bradley & Bryant, 1978; Bryant & Goswami, 1986; see also Parrila et al., 2020, for a discussion on the RL design). Following this line of reasoning, our findings suggest that MA is not a diagnostic marker of dyslexia (root cause), but rather a secondary problem resulting from deficits in word-level reading.

Moderator effects

In regard to the second goal of our meta-analysis, with one exception (i.e., age), none of the moderators explained a significant amount of the observed variability in the effect sizes. There might be several explanations for these non-significant findings. First, it is possible that "type of MA task" is not as important as task difficulty (Deacon et al., 2008). For instance, in the "judgment" group, selecting "farmer" as the answer to "*My uncle is a* ______ *(farmer, farming)*" is relatively easy, but selecting the right answer with pseudowords is more difficult, e.g., "*My brother is a chocolate*

Table 4. Publication bias analyses.

Comparison	Fail-Safe N	Egger's Method		Rank Correlation Test		Trim and Fill Procedure	
		z	*p*	Kendall's tau	*p*	Imputed	Corrected effect sizes
CA-DYS	15649	1.051	0.293	0.083	0.405	0	1.114
RL-DYS	0	−2.189	0.028	−0.323	0.076	0	−0.086

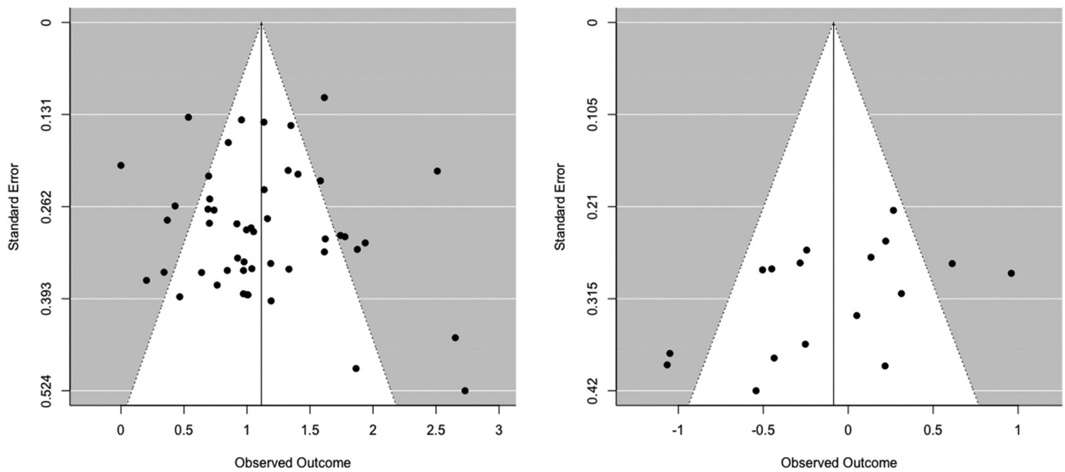

Figure 4. Funnel plots for the CA-DYS (left) and RL-DYS (right) comparisons.

________ *(cloveler, clovelly)*." Some of the more challenging tasks (on which both groups may do poorly) may call for a more explicit and abstract knowledge of morphology, rather than just choosing what "sounds right." In the same vein, analogy tasks are thought to be difficult because they also require analogical reasoning; when we use them with young children, we may find lower scores for both groups. Again, it may be task difficulty that is important: more difficult or very easy tasks (for the age group) may not show differences, but moderately difficult tasks may have more chance of showing differences. To investigate this would require many more studies to examine specific tests at defined age levels or studies with items selected for equal difficulty in different tasks.

In regard to language, the absence of a significant effect suggests that the MA deficits are likely universal. The universality of the findings is consistent with the binding agent theory (Kirby & Bowers, 2017); children will bind phonology, orthography, and semantics, regardless of language. Nevertheless, the absence of a significant effect of language might be due to the small number of languages included here or the lack of multiple studies within a given language that prevented us from running our analyses with enough power.

Although modality of input was not significant, it does show considerable differences in the effect sizes (1.25 for visual, 1.11 for oral, and 0.98 for tasks presented both orally and visually). Had we had a larger sample of effect sizes in the "visual" and "both" categories, this moderator may have reached statistical significance. The fact that the effect size in visual MA tasks was the largest is not surprising given that children with dyslexia experience difficulties in word reading and there was reading involved in the visually-presented MA tasks.

The significant effect of age was not surprising – as typically developing children get older, they read more, and that exposure is critical for establishing mental representations of morphemes (and words). The children who experience reading difficulties start off behind and fall further and further behind. The absence of RL-DYS effects is consistent with this, RL being a proxy for how much reading they have done (Bryant & Goswami, 1986).

Theoretical interpretations and educational implications

In integrating our results, we suggest that the following explains what we have found. Early phonological awareness (PA) could influence oral morphological awareness (MA), because sounds need to be distinguished to hear different morphemes (Cunningham & Carroll, 2015). Children to be later identified as having dyslexia would have lower PA and consequential lower MA compared to their CA controls, though not to their RL controls. This initially low MA would weaken later oral and written MA, as shown in the overall analysis of differences in the CA comparisons. At the same time, initial weak PA and oral MA in children with dyslexia would weaken their reading skills, leading to less reading and less automatization. These, in turn, would weaken the use of larger units (whether syllables or morphemes) and thus to these children doing less reading, contributing to even lower reading skill over time in comparison to their CA controls. In contrast, the children with dyslexia may be at roughly the same level in MA as their RL controls, and make as much progress in MA as the RL controls do in following years. This explanation is clearly post hoc and will require considerable detailed investigation to be tested.

If the theoretical interpretation presented in the preceding paragraph is accurate, it may be helpful to provide children with dyslexia instruction in MA in addition to PA and phonics. There is considerable research demonstrating that PA and phonics instruction is beneficial, but perhaps not universally successful (e.g., Galuschka et al., 2014). There is also a growing body of research on MA instruction which shows it has positive effects on reading development, including for those with reading difficulties (Bowers et al., 2010; Goodwin & Ahn, 2013). The existence of differences in MA under current instructional conditions does not imply that children with dyslexia cannot benefit from instruction in MA. Furthermore, the lack of differences in the RL control comparisons suggests that the lower MA performance of children with dyslexia is potentially a developmental lag, one which may be susceptible to instructional improvement. Our sense is that oral morphological instruction could start early, as shown by Lyster et al. (2016), in conjunction with appropriate phonological instruction. As formal reading instruction begins, morphological instruction could shift to focusing on written language, just as phonological instruction shifts from oral phonological awareness to written phonics. Detailed longitudinal studies are required to assess these instructional implications.

Limitations

Some limitations of the present study are worth mentioning. First, our sample size was relatively small and this may partly explain why we were not able to detect significant moderator effects. Relatedly, the small sample size did not allow us to test for possible interaction effects (e.g., type of MA task X language). Second, we did not examine how well the groups were matched on the reading tasks as this was beyond the scope of this meta-analysis. Parrila et al. (2020) found in a recent meta-analysis that even though dyslexics were matched to their controls on one reading task (i.e., the task used to select them), they differed in other reading tasks. Imperfect matching may have significant implications when searching for core deficits in dyslexia. Third, our meta-analysis included children and youth with dyslexia and our findings may not generalize to adults with dyslexia. We debated whether to include adults in this meta-analysis but we chose not to because there are different types of adult dyslexics (e.g., compensated dyslexics, high functioning dyslexics, persistent dyslexics) identified in different ways and this could have introduced error in our calculations. Fourth, we acknowledge that there is a debate regarding the best way of identifying children with dyslexia (Elliott, 2020). With the exception of two studies that described their participants as reading disabled, the rest of the studies in our meta-analysis described their participants as dyslexics. Some of these studies indicated that their sample was selected on the basis of a formal diagnosis of dyslexia, but they did not provide details on how this diagnosis was made. Unfortunately, we cannot evaluate the selection process in these studies. Finally, it is possible that some children with dyslexia in our studies also had developmental language disorders (see McArthur et al., 2000, on the comorbidity of developmental language disorders and dyslexia). Unfortunately, most studies did not include any information on that and we cannot estimate the effect of comorbidity on the observed effects.

Conclusion

Our findings add to a growing body of meta-analyses in dyslexia (e.g., Araújo & Faísca, 2019; Georgiou et al., 2021; Kudo et al., 2015; Melby-Lervåg et al., 2012; Parrila et al., 2020; Peng et al., 2017) by showing that children with dyslexia experience significant difficulties in MA when compared to their CA controls and that the group differences in MA increase with age.

Given that the overall effect size in MA is large (g = 1.11; Cohen, 1988) and similar to that reported for other key correlates of reading (e.g., phonological awareness, rapid naming, orthographic knowledge), it might be useful to include MA tasks in screening batteries together with more traditional phonological ones. That may help distinguish between pure phonological types of dyslexia and broader language problems that are more closely related to developmental language disorders.

Notes

1. We acknowledge that some production tasks require syntactic skills and therefore could be considered measures of morpho-syntax (see, Goodwin et al., 2022). To exclude the possibility that their inclusion in our meta-analysis has influenced the results, we reran our analyses without the studies that used these measures. The results remained essentially the same (g = 1.108 for the CA-DYS comparison and g = −0.057 for the RL-DYS comparison).
2. We also reran our analyses using robust variance estimation (RVE) meta-regression that includes all effect sizes for each study instead of an average of effect sizes and the results were the same (see Results section).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The study was partly funded by the National Social Science Fund of China (CHA180266) to Dr. Kan Guo.

ORCID

George K. Georgiou http://orcid.org/0000-0002-9081-992X
Kyriakoula M. Rothou http://orcid.org/0000-0002-3080-8006
John R. Kirby http://orcid.org/0000-0002-8874-1641

References

Abu-Rabia, S., Share, D., & Mansour, M. S. (2003). Word recognition and basic cognitive processes among reading-disabled and normal readers in Arabic. *Reading and Writing: An Interdisciplinary Journal*, *16*(5), 423–442. https://doi.org/10.1023/A:1024237415143

Apel, K. (2014). A comprehensive definition of morphological awareness. Implications for assessment. *Topics in Language Disorders*, *34*(3), 197–209. https://doi.org/10.1097/TLD.0000000000000019

Araújo, S., & Faísca, L. (2019). A meta-analytic review of naming-speed deficits in developmental dyslexia. *Scientific Studies of Reading*, *23*(5), 1–20. https://doi.org/10.1080/10888438.2019.1572758

Bar-Kochva, I., Korinth, S., & Hasselhorn, M. (2020). Effects of a morpheme-based training procedure on the literacy skills of readers with a reading disability. *Applied Psycholinguistics*, *41*(5), 1061–1082. https://doi.org/10.1017/S0142716420000120

Berthiaume, R., & Daigle, D. (2014). Are dyslexic children sensitive to the morphological structure of words when they read? The case of dyslexic readers of French. *Dyslexia*, *20*(3), 241–260. https://doi.org/10.1002/dys.1476

Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta-analysis* (1st) ed.). John Wiley & Sons. https://doi.org/10.1002/9780470743386
Borleffs, E., Maassen, B. A. M., Lyytinen, H., & Zwarts, F. (2017). Measuring orthographic transparency and morphological-syllabic complexity in alphabetic orthographies: A narrative review. *Reading and Writing*, *30*(8), 1617–1638. https://doi.org/10.1007/s11145-017-9741-5
Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, *80*(2), 144–179. https://doi.org/10.3102/0034654309359353
Bradley, L., & Bryant, P. E. (1978). Difficulties in auditory organization as a possible cause of reading backwardness. *Nature*, *271*(5647), 746–747. https://doi.org/10.1038/271746a0
Bryant, P., & Goswami, U. (1986). Strengths and weaknesses of the reading level design: A comment on Backman, Mamen, and Ferguson. *Psychological Bulletin*, *100*(1), 101–103. https://doi.org/10.1037/0033-2909.100.1.101
Caglar-Ryeng, Ö. (2010). *Morphological awareness skills of Norwegian adolescent dyslexics acquiring English as a second language*. [Unpublished master's thesis]. University of Tromsø.
Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189–209). Erlbaum.
Carroll, J. M., & Breadmore, H. L. (2018). Not all phonological awareness deficits are created equal: Evidence from a comparison between children with Otitis Media and poor readers. *Developmental Science*, *21*(3), 1–12. https://doi.org/10.1111/desc.12588
Casalis, S., Colé, P., & Sopo, D. (2004). Morphological awareness in developmental dyslexia. *Annals of Dyslexia*, *54*(1), 114–138. https://doi.org/10.1007/s11881-004-0006-z
Cavalli, E., Duncan, L. G., Elbro, C., El Ahmadi, A., & Cole, P. (2017). Phonemic-morphemic dissociation in university students with dyslexia: An index of reading compensation? *Annals of Dyslexia*, *67*(1), 63–84. https://doi.org/10.1007/s11881-016-0138-y
Chung, K. K. H., Ho, C. S.-H., Chan, D. W., Tsang, S.-M., & Lee, S.-H. (2010). Cognitive profiles of Chinese adolescents with dyslexia. *Dyslexia*, *16*(1), 2–23. https://doi.org/10.1002/dys.392
Chung, K. K. H., Ho, C. S.-H., Chan, D. W., Tsang, S.-M., & Lee, S. H. (2011). Cognitive skills and literacy performance of Chinese adolescents with and without dyslexia. *Reading and Writing*, *24*(7), 835–859. https://doi.org/10.1007/s11145-010-9227-1
Chung, K. K. H., Ho, C. S.-H., Chan, D. W., Tsang, S.-M., & Lee, S.-H. (2013). Contributions of syntactic awareness to reading in Chinese-speaking adolescent readers with and without dyslexia. *Dyslexia*, *19*(1), 11–36. https://doi.org/10.1002/dys.1448
Chung, K. K. H., Lo, J. C. M., Ho, C. S.-H., Xiao, X., & Chan, D. W. (2014). Syntactic and discourse skills in Chinese adolescent readers with dyslexia: A profiling study. *Annals of Dyslexia*, *64*(3), 222–247. https://doi.org/10.1007/s11881-014-0095-2
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Lawrence Erlbaum Associates.
Cunningham, A., & Carroll, J. (2015). Early predictors of phonological and morphological awareness and the link with reading: Evidence from children with different patterns of early deficit. *Applied Psycholinguistics*, *36*(3), 509–531. https://doi.org/10.1017/S0142716413000295
Deacon, S. H., Benere, J., & Pasquarella, A. (2013). Reciprocal relationship: Children's morphological awareness and their reading accuracy across grades 2 to 3. *Developmental Psychology*, *49*(6), 1113–1126. https://doi.org/10.1037/a0029474
Deacon, S. H., Parrila, R., & Kirby, J. R. (2008). A review of the evidence on morphological processing in dyslexics and poor readers: A strength or weakness? In F. Manis, A. Fawcett, G. Reid, & L. Siegel (Eds.), *The Sage Handbook of Dyslexia* (pp. 212–237). Sage.
Deacon, S. H., Tong, X., & Mimeau, C. (2016). Morphological and semantic processing in developmental dyslexia across languages: A theoretical and empirical review. In C. Perfetti, K. Pugh, & L. Verhoeven (Eds.), *Dyslexia across languages and writing systems: A handbook* (pp. 327–349). Cambridge University Press.
Diamanti, V., Goulandris, N., Stuart, M., & Campbell, R. (2014). Spelling of derivational and inflectional suffixes by Greek-speaking children with and without dyslexia. *Reading and Writing: An Interdisciplinary Journal*, *27*(2), 337–358. https://doi.org/10.1007/s11145-013-9447-2
Duncan, L. G., Casalis, S., & Colé, P. (2009). Early metalinguistic awareness of derivational morphology: Observations from a comparison of English and French. *Applied Psycholinguistics*, *30*(3), 405–440. https://doi.org/10.1017/S0142716409090213
Duranovic, M., Tinjak, S., & Turbic-Hadzagic, A. (2014). Morphological knowledge in children with dyslexia. *Journal of Psycholinguistic Research*, *43*(6), 699–713. https://doi.org/10.1007/s10936-013-9274-2
Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x
Elbro, C., & Arnbak, E. (1996). The role of morpheme recognition and morphological awareness in dyslexia. *Annals of Dyslexia*, *46*(1), 209–240. https://doi.org/10.1007/BF02648177
Elliott, J. G. (2020). Its time to be scientific about dyslexia. *Reading Research Quarterly*, *55*(S1), S61–S75. https://doi.org/10.1002/rrq.333

Elliott, J. G., & Grigorenko, E. L. (2014). *The dyslexia debate*. Cambridge University Press. https://doi.org/10.1017/CBO9781139017824

Galuschka, K., Ise, E., Krick, K., & Schulte-Körne, G. (2014). Effectiveness of treatment approaches for children and adolescents with reading disabilities: A meta-analysis of randomized controlled trials. *PLOS ONE*, *9*(2), Article e89900. https://doi.org/10.1371/journal.pone.0089900

Georgiou, G. K., Martinez, D., Vieira, A. P. A., & Guo, K. (2021). Is orthographic knowledge a strength or a weakness in individuals with dyslexia? Evidence from a meta-analysis. *Annals of Dyslexia*, *71*(1), 5–27. https://doi.org/10.1007/s11881-021-00220-6

Giazitzidou, S., & Padeliadu, S. (2022). Contribution of morphological awareness to reading fluency of children with and without dyslexia: evidence from a transparent orthography. *Annals of Dyslexia*, *72*(3), 509–531. https://doi.org/10.1007/s11881-022-00267-z

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257–285. https://doi.org/10.1080/10888438.2012.689791

Goodwin, A. P., Petscher, Y., & Reynolds, D. (2022). Unraveling adolescent language & reading comprehension: The monster's data. *Scientific Studies of Reading*, *26*(4), 305–326. https://doi.org/10.1080/10888438.2021.1989437

Hedges, L. V., & Olkin, I. (2014). *Statistical methods for meta-analysis*. Academic Press.

Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology*, *77*(1), 30–60. https://doi.org/10.1006/jecp.1999.2553

Kalindi, S. C., & Chung, K. K. H. (2018). The impact of morphological awareness on word reading and dictation in Chinese early adolescent readers with and without dyslexia. *Frontiers in Psychology*, *9*, 1–14. https://doi.org/10.3389/fpsyg.2018.00511

Kirby, J. R., & Bowers, P. N. (2017). Morphological instruction and literacy: Binding phonological, orthographic, and semantic features of words. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 437–461). John Benjamins Publishing Company.

Kirby, J. R., & Bowers, P. N. (2018). The effects of morphological instruction on vocabulary learning, reading, and spelling. In R. Berthiaume, D. Daigle, & A. Desrochers (Eds.), *Morphological processing and literacy development: Current issues and research* (pp. 217–243). Routledge.

Kirby, J. R., Deacon, H. S., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading & Writing*, *25*(2), 389–410. https://doi.org/10.1007/s11145-010-9276-5

Kotzer, M., Kirby, J. R., & Heggie, L. (2021). Morphological awareness predicts reading comprehension in adults. *Reading Psychology*, *42*(3), 302–322. https://doi.org/10.1080/02702711.2021.1888362

Kruk, R. S., & Bergman, K. (2013). The reciprocal relations between morphological processes and reading. *Journal of Experimental Child Psychology*, *114*(1), 10–34. https://doi.org/10.1016/j.jecp.2012.09.014

Kudo, M. F., Lussier, C. M., & Swanson, H. L. (2015). Reading disabilities in children: A selective meta-analysis of the cognitive literature. *Research in Developmental Disabilities*, *40*, 51–62. https://doi.org/10.1016/j.ridd.2015.01.002

Kuo, L., & Anderson, R. C. (2006). Morphological awareness and learning to read: A cross-language perspective. *Educational Psychologist*, *41*(3), 161–180. https://doi.org/10.1207/s15326985ep4103_3

Law, J. M., Wouters, J., & Ghesquière, P. (2015). Morphological awareness and its role in compensation in adults with dyslexia. *Dyslexia*, *21*(3), 254–272. https://doi.org/10.1002/dys.1495

Layes, S., Lalonde, R., & Rebaï, M. (2017). Study on morphological awareness and rapid automatized naming through word reading and comprehension in normal and disabled reading Arabic-speaking children. *Reading & Writing Quarterly*, *33*(2), 123–140. https://doi.org/10.1080/10573569.2015.1105763

Levesque, K. C., Breadmore, H. L., & Deacon, S. H. (2021). How morphology impacts reading and spelling: Advancing the role of morphology in models of literacy development. *Journal of Research in Reading*, *44*(1), 10–26. https://doi.org/10.1111/1467-9817.12313

Lin, L., Chu, H., Murad, M. H., Hong, C., Qu, Z., Cole, S. R., & Chen, Y. (2018). Empirical comparison of publication bias tests in meta-analysis. *Journal of General Internal Medicine*, *33*(8), 1260–1267. https://doi.org/10.1007/s11606-018-4425-7

Liu, P. D., & McBride-Chang, C. (2010). What is morphological awareness? Tapping lexical compounding awareness in Chinese third graders. *Journal of Educational Psychology*, *102*(1), 62–73. https://doi.org/10.1037/a0016933

Lyster, S.-A. H., Lervåg, A. O., & Hulme, C. (2016). Preschool morphological training produces long-term improvements in reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, *29*(6), 1269–1288. https://doi.org/10.1007/s11145-016-9636-x

Maynard, C., Brissaud, C., & Armand, F. (2018). The acquisition of inflectional morphology in children. In R. Berthiaume, D. Daigle, & A. Desrochers (Eds.), *Morphological processing and literacy development: Current issues and research* (pp. 217–243). Routledge.

McArthur, G. M., Hogben, J. H., Edwards, V. T., Heath, S. M., & Mengler, E. D. (2000). On the "specifics" of specific reading disability and specific language impairment. *Journal of Child Psychology and Psychiatry*, *41*(7), 869–874. https://doi.org/10.1111/1469-7610.00674

Melby-Lervåg, M., Lyster, S.-A., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352. https://doi.org/10.1037/a0026744

Meng, Z.-L., Wydell, T. N., & Bi, H.-Y. (2019). Visual□motor integration and reading Chinese in children with/without dyslexia. *Reading and Writing*, *32*(2), 493–510. https://doi.org/10.1007/s11145-018-9876-z
Parrila, R., Dudley, D., Song, S., & Georgiou, G. (2020). A meta-analysis of reading-level match dyslexia studies in consistent alphabetic orthographies. *Annals of Dyslexia*, *70*(1), 1–26. https://doi.org/10.1007/s11881-019-00187-5
Peng, P., Wang, C., Tao, S., & Sun, C. (2017). The deficit profiles of Chinese children with reading difficulties: A meta-analysis. *Educational Psychology Review*, *29*(3), 513–564. https://doi.org/10.1007/s10648-016-9366-2
Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383. https://doi.org/10.1080/10888430701530730
Plaut, D. C., McCelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115. https://doi.org/10.1037/0033-295X.103.1.56
Quémart, P., & Casalis, S. (2015). Visual processing of derivational morphology in children with developmental dyslexia: Insights from masked priming. *Applied Psycholinguistics*, *36*(2), 345–376. https://doi.org/10.1017/S014271641300026X
Robertson, E. K., Joanisse, M. F., Desroches, A. S., & Terry, A. (2013). Past-tense morphology and phonological deficits in children with dyslexia and children with language impairment. *Journal of Learning Disabilities*, *46*(3), 230–240. https://doi.org/10.1177/0022219412449430
Rothou, K. M., & Padeliadu, S. (2019). Morphological processing influences on dyslexia in Greek-speaking children. *Annals of Dyslexia*, *69*(3), 261–278. https://doi.org/10.1007/s11881-019-00184-8
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568. https://doi.org/10.1037/0033-295X.96.4.523
Siegel, L. S. (2008). Morphological awareness skills of English language learners and children with dyslexia. *Topics in Language Disorders*, *28*(1), 15–27. https://doi.org/10.1097/01.adt.0000311413.75804.60
Singson, M., Mahony, D., & Mann, V. (2000). The relation between reading ability and morphological skills: Evidence from derivation suffixes. *Reading and Writing*, *12*(3/4), 219–252. https://doi.org/10.1023/A:1008196330239
Snowling, M., Hulme, C., & Nation, K. (2020). Defining and understanding dyslexia: Past, present and future. *Oxford Review of Education*, *46*(4), 501–513. https://doi.org/10.1080/03054985.2020.1765756
Spencer, M., Muse, A., Wagner, R. K., Foorman, B., Petscher, Y., Schatschneider, C., Tighe, E. L., & Bishop, M. D. (2015). Examining the underlying dimensions of morphological awareness and vocabulary knowledge. *Reading and Writing*, *28*(7), 959–988. https://doi.org/10.1007/s11145-015-9557-0
Stafura, J. Z., & Perfetti, C. A. (2017). Integrating word processing with text comprehension: Theoretical frameworks and empirical examples. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 9–32). John Benjamins Publishing Company.
Suárez-Coalla, P., Martinez-Garcia, C., & Cuetos, F. (2017). Morpheme-based reading and writing in Spanish children with dyslexia. *Frontiers in Psychology*, *8*, 1952. https://doi.org/10.3389/fpsyg.2017.01952
Sümer Dodur, H. M. (2021). Inflectional morphological awareness, word reading and reading comprehension of Turkish students with learning disabilities. *International Online Journal of Education and Teaching (IOJET)*, *8*(3), 1543–1559. https://eric.ed.gov/?id=EJ1308237
Torppa, M., Lyytinen, P., Erskine, J., Eklund, K., & Lyytinen, H. (2010). Language development, literacy skills, and predictive connections to reading in Finnish children with and without familial risk for dyslexia. *Journal of Learning Disabilities*, *43*(4), 308–321. https://doi.org/10.1177/0022219410369096
Vender, M., Mantione, F., Savazzi, S., Delfitto, D., & Melloni, C. (2017). Inflectional morphology and dyslexia: Italian children's performance in a nonword pluralization task. *Annals of Dyslexia*, *67*(3), 401–426. https://doi.org/10.1007/s11881-017-0152-8
Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03
Wong, A. M.-Y., Kidd, J. C., Ho, C. S.-H., & Au, T. K.-F. (2010). Characterizing the overlap between SLI and dyslexia in Chinese: The role of phonology and beyond. *Scientific Studies of Reading*, *14*(1), 30–57. https://doi.org/10.1080/10888430903242043
Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3