

*Morphological instruction and
reading development in young L2 readers:
A scoping review of causal relationships*

Sihui (Echo) Ke

University of Kentucky, Lexington, USA

<https://orcid.org/0000-0003-2665-0637>

sihui.ke@uky.edu

Dongbo Zhang

University of Exeter, UK

<https://orcid.org/0000-0002-4175-2052>

d.zhang4@exeter.ac.uk

Abstract

This scoping review explores the causal relationship between morphological instruction and reading development in young L2 learners by synthesizing 12 primary studies published between 2004 and 2019 ($N = 1,535$). These studies focused on reading English as the target language and involved participants between kindergarten and Grade 12 from four countries (China, Egypt, Singapore, and the USA). Findings suggested that (a) morphological instruction led to consistent and positive gains in L2 children's morphological awareness and vocabulary knowledge, and the effect sizes (Cohen's d s) ranged from small to large; and (b) the relationship between morphological instruction and other outcomes such as phonological awareness, word reading accuracy, word reading fluency, spelling, and reading comprehension was inconclusive. Notably, transfer effects of L2 English morphological instruction on novel word learning in English or on reading development in an additional language were only examined and observed in four primary studies. Discussion was provided regarding future instructional and research design.

Keywords: causal effect; morphology; second language reading; literacy; scoping review

1. Introduction

English is a morphophonemic language. When there is inconsistency in sound-grapheme mapping, a morpheme is still preserved in orthographic units (e.g., *heal-health, courage-courageous, cats-dogs*; Frost, 2012, p. 269). It is thus not surprising that an increasing number of cross-sectional and longitudinal studies found significant correlations between morphological awareness (i.e., a learner's sensitivity to word-internal morphological structure) and English reading sub-skill development, including development in English as a second language (L2; e.g., Hayashi & Murphy, 2013; McBride-Chang et al., 2012; Saiegh-Haddad & Geva, 2008; Zhang et al., 2014).

Recently, emerging research has also started to test the causal relationship between morphology and reading development through morphological instruction. In this context, the question surrounding causal inference is not whether morphological knowledge is associated with reading outcomes, but whether changing or manipulating morphological knowledge via intervention might alter reading outcomes (see also an explanation of causal inference in developmental psychology in Foster, 2010). Morphological instruction draws learners' attention to intraword morphological structure and supports learning unfamiliar words based on familiar word parts. A central component of morphological instruction is word problem solving (Goodwin & Perkins, 2015). For example, *robotceptionist* might be unknown to both native and non-native speakers of English. Teachers can guide students to break down the novel word into *robot* and *ceptionist*, and students may then be able to infer the unknown word meaning as a robot that serves as a receptionist. Evidence has been provided that morphological instruction is beneficial for English reading development in both L1 (first language) and L2 English-speaking children (e.g., Carlo et al., 2004; Kiefer & Lesaux, 2012) and that the positive effect is more pronounced in L2 English reading (see a meta-analysis by Goodwin & Ahn, 2013). Kirby and Bowers (2017) posited that the question for research in this line should no longer be *whether* morphological instruction benefits reading development but *how* it can lead to positive gains in reading outcomes or achievement.

This synthesis study, in the form of a scoping review, thus explored the ways in which morphological instruction contributes to English reading development in young L2 learners. To achieve this goal, we selected 12 studies from existing literature following a systematic approach and evaluated the primary evidence both quantitatively and qualitatively. A scoping review maps the literature from a particular topic or research area and provides an opportunity to identify key concepts, gaps in the research, and types as well as sources of evidence to inform practice, policymaking, and research (Arksey & O'Malley, 2005;

Pham et al., 2014). We examined the implementation of morphological instruction in the selected studies, explored the extent to which morphological instruction influences the acquisition of a range of L2 reading-related outcomes (phonological awareness, morphological awareness, vocabulary, word decoding accuracy and fluency, spelling, and reading comprehension) by summarizing the effect sizes (Cohen's *ds*), and evaluated study and instructional designs following systematic coding schemes. The findings of this synthesis study will provide implications for future research about the causal relationship between morphology and L2 reading development as well as real-world classroom instruction.

2. Background: Findings from previous reviews of morphological instruction

Recent reviews have discussed the causal relationship between morphological instruction and English literacy development, including critical reviews (Henbest & Apel, 2017; Kirby & Bowers, 2017; Nagy et al., 2014), systematic syntheses (Brandes & McMaster, 2017; Carlisle, 2010), and quantitative meta-analyses (Bowers et al., 2010; Goodwin & Ahn, 2010, 2013; Reed, 2008).

In a meta-analysis of 30 independent studies and 92 standardized mean differences, Goodwin and Ahn (2013) assessed the overall effect of morphological instruction and examined possible moderator effects. Their findings indicated that children who received morphological instruction performed significantly better on diverse measures of literacy achievement than comparison groups; the overall effect size was medium ($d = 0.32$). Yet, effect sizes varied across the literacy outcomes, ranging from significant and moderate for five outcomes (i.e., phonological awareness, morphological awareness, vocabulary, word decoding, and spelling) to non-significant for reading fluency or reading comprehension. The authors further identified significant moderating effects of school level, type of experimental design, and type of literary measures. Specifically, larger effect sizes were found in studies of younger children than those of older children; likewise, quasi-experimental studies that adopted researcher-designed measures generated larger effect sizes than did experimental studies that adopted standardized measures. No significant moderating effects were found for instructional features (e.g., stand-alone morphological instruction versus integrated instruction including morphology in a comprehensive curriculum, length, and learner types). Instructional strategies (e.g., teaching affixes versus bases, promoting problem-solving or not), however, were not included in the analysis.

Lack of attention to instructional strategies was later addressed in Kirby and Bowers' (2017) critical review, which sought to answer *how* morphology has been taught and *to what extent*, *when*, and *for whom* morphological instruction is beneficial for literacy development (in English). The authors observed great

variation in how morphological instruction was implemented in prior studies. Major instructional strategies analyzed in the review included the isolation/integration mode; teaching oral/written morphology, including affix/base items; focusing on orthographic changes; and promoting problem solving and scientific inquiries. Kirby and Bowers summarized four patterns regarding the overall effects of morphological instruction: (a) compared to regular classroom instruction, morphological instruction had positive effects; (b) effects of morphological instruction were roughly equal to those of alternative treatments; (c) effects of morphological instruction were positive for meaning-related as well as form-related outcomes; and (d) effects were the strongest at sub-lexical levels (e.g., phonological awareness), weaker at lexical levels (e.g., vocabulary and word decoding), and the weakest at supra-lexical levels (e.g., sentence or passage-level reading comprehension). Kirby and Bowers (2017) also proposed a set of instructional principles and hypotheses to be tested. One hypothesis was about the transfer¹ effect of morphological instruction on word reading, spelling of unknown words, and reading comprehension at the sentence/paragraph level.

Similar questions have been posed in two other critical reviews (Carlisle, 2010; Nagy et al., 2014). Nagy et al. (2014), for example, pointed out that previous research mainly focused on morphology as a tool for inferring the meaning of new words; yet emerging evidence has shown that morphological awareness is also related to reading development at the word form level (e.g., word decoding and spelling). Regarding future inquiries concerning transfer effects of morphological instruction, Carlisle (2010, p. 481) further raised the following questions: “What aspects of programs contribute to significant effects on measures of ‘transfer of learning’ to new words and passages? Why is it that there are so few significant effects of morphological awareness on performance of reading comprehension measures?” Those reviews called for more research on cross-language comparisons of morphological instruction and the extent to which the relationship between morphological awareness and reading development is language-specific.

In summary, despite a strong association revealed in the literature between morphological awareness and L2 English reading development, the existing

¹ *Transfer* in this review is an overarching concept that refers to the influence of earlier learning in one situation on later learning in another situation (Seel, 2012). It has also been referred to as *learning transfer*, *transfer of learning*, *transfer of practice*, or *transfer of training* in educational literature (see a review in Larsen-Freeman, 2013). One oft-examined specific type of transfer of learning in L2 research is restricted to learners’ application of knowledge or skills developed in one language to situations involving another language, which was reviewed in this review (e.g., Zhang, 2016; Zhang et al. 2010). Another type of transfer effect we examined was learners’ application of morphological awareness in English to learning novel words in English (e.g., Davidson & O’Connor, 2019; Deng, 2016).

empirical evidence is mainly based on observation studies with correlational data, not interventional studies of causal connections. In addition, while some recent reviews provided evidence on a causal relationship in that morphological instruction was effective for English reading development, those reviews usually did not specifically target any particular learner group (e.g., Goodwin & Ahn, 2013; Kirby & Bowers, 2017). While larger gains seemed to emerge for less able English learners in those reviews, there was little immediate attention to L2 English readers. To this end, this scoping review synthesized the current body of primary studies and examined *how* morphological instruction contributes to L2 English reading development. It focused on young L2 English learners and was guided by three research questions (RQs):

RQ1: How has morphological instruction been implemented in the primary studies?

RQ2: To what extent has morphological instruction benefited the development of different reading-related skills (i.e., phonological awareness [PA], morphological awareness [MA], vocabulary, word decoding accuracy and fluency, spelling, and reading comprehension) in young L2 English readers?

RQ3: How has the causal effect of morphological instruction been examined?

3. Method

We followed the procedures of conducting a scoping review in the social sciences (Arksey & O'Malley, 2005; Pham et al., 2014). A literature search was conducted in January 2020 with key words (morphological instruction, morphological awareness instruction, morphology instruction, teaching morphology, morphological awareness) entered in PROQUEST Databases, PsycInfo, Google Scholar, and Web of Science². In addition, a manual search was conducted by extracting references from previous reviews of L1 and L2 morphological instruction (Bowers et al., 2010; Brandes & McMaster, 2017; Carlisle, 2010; Goodwin & Ahn, 2013; Henbest & Apel, 2017; Kirby & Bower, 2017; Nagy et al., 2014; Reed, 2008). Studies included in this review are those that (a) were reported in either published articles or unpublished dissertations; (b) reported morphological instruction in young L2 learners (defined as those below 18 years old and learning an additional language); and (c) included descriptive and/or inferential statistics for pre- and post-instruction testing results. Unpublished dissertation

² For all of the four databases, the search was conducted with the default settings. The initial search resulted in a total of 661 entries.

studies were included to avoid publication bias (following Oswald & Plonsky, 2010). Studies that focused solely on monolingual children or children with learning disabilities were excluded. Repetitive samples were also excluded.

As a result, 12 studies between 2004 and 2019 ($N = 1,535$) were identified. They consisted of 10 published journal articles and two unpublished dissertations. The primary studies focused on reading English as the target language and involved participants between kindergarten and Grade 12 from four countries (China, Egypt, Singapore, and the USA); yet the dominant education setting was the USA. Based on the target reading-related outcomes, the selected studies can be further categorized into seven groups: phonological awareness (PA), morphological awareness (MA), vocabulary knowledge, word decoding accuracy, word decoding fluency, spelling, and reading comprehension, as shown in Table 3 in the next section. It is noteworthy that only two of the studies considered cross-language transfer effects of morphological instruction and observed gains in morphological awareness in another language (Malay in Zhang, 2016; Chinese for children of high L2 English proficiency in Zhang et al., 2010) as a result of English morphological instruction. The other ten studies mainly examined the effect of L2 English morphological instruction on L2 English reading development (i.e., intra-lingual effect of morphological instruction). An exception might be Carlo et al. (2004), which included L1 Spanish in the instructional phase. More details can be found in the results section.

For the purpose of answering the three research questions, we followed these procedures: (a) morphological instruction design was coded after Goodwin and Ahn (2013) and evaluated (see also Kirby & Bowers, 2017), (b) effect sizes were coded and summarized, and (c) reading-related outcome measures were also coded after Goodwin and Ahn (2013) and compared across primary studies. The first author of this paper conducted coding twice and double-checked coding until the intra-coder agreement rate was 100%.

4. Results

4.1. The implementation of morphological instruction in the primary studies

This section addresses how morphological instruction has been implemented in the primary studies. To answer this question, we first summarized the characteristics of instructional programs across studies and then examined the specific strategies in relation to morphological instruction. Due to space limitations, we could not describe all the details of instructional implementation for each study in this review. An example of how to implement explicit morphological instruction into regular classrooms with young L2 learners can be found in Lesaux et al.

(2010), which reported a large-scale, mixed-methods study of the implementation and effectiveness of an academic vocabulary program designed for use in mainstream middle school classrooms with high proportions of language minority learners in the USA (including 21 classes in seven middle schools with 346 language minority learners and 130 native English speakers).

The design of instructional program by study is shown in Table 1, including seven major features: (a) learner profiles (grade level and language background), (b) randomization of treatment (experimental versus quasi-experimental), (c) instructional time (sessions and minutes in total), (d) the scope of intervention (stand-alone morphological instruction or morphological instruction as part of a more comprehensive instruction), (e) person/people implementing the intervention (researchers versus teachers), (f) control/comparison condition(s) (regular classroom or alternative treatments), and (g) fidelity and feasibility of the instructional program. Fidelity is often defined as the determination of how well an intervention is implemented in comparison with the original program design during an efficacy and/or effectiveness study; feasibility refers to how likely an intervention will be implemented with fidelity in the classroom (O'Donnell, 2008).

Table 1 Characteristics of instructional program by study

Author(s) & year	Country	Grade	Learner background (L2 only/mixed)	Research design (E/QS)	Length (weeks/sessions/minutes in total)	Scope (SA/P)	Instructor (T/R)	Control/comparison condition (UC/AT)	Fidelity	Feasibility
Badawi (2019)	Egypt	7	L2 only	E	NA (two units)	SA	NA	UC	NA	Yes
Carlo et al. (2014)	USA	5	Mixed	QS	60 sessions in 15 weeks, 2250 minutes	P	T	UC	Yes	Yes
Crosson & Moore (2017)	USA	6-12	L2 only	QS	NA (6 weeks)	P	T	AT	Yes	Yes
Davidson & O'Connor (2019)	USA	4-5	L2 only	QS	10 sessions, 150 minutes	SA	R	AT	Yes	Yes
Deng (2016)	USA	3-5	L2 only	E	16 sessions, 480 minutes	P	R	AT	Yes	Yes
Filippini (2007)	USA	1	Mixed	E	29 sessions, 394 minutes	P	R	AT	Yes	NA
Goodwin (2016)	USA	5-6	Mixed	E	4 sessions, 120 minutes	P	R	AT	Yes	NA
Kieffer & Lesaux (2012)	USA	6	Mixed	QS	72 sessions, 3780 minutes	P	T	UC	Yes	Yes
Lesaux et al. (2010)	USA	6	Mixed	QS	18 weeks, 8 day cycle, about 810 minutes	P	T	UC	Yes	Yes
Zhang (2016)	Singapore	4	Mixed	QS	8 sessions, 400 minutes	P	T	UC	Yes	NA
Zhang, et al. (2010)	China	5	L2 only	E	1 session, 45 minutes	SA	T	UC & AT	No	Yes
Zoski & Erickson (2017)	USA	K	Mixed	QS	6 weeks, 720 minutes	SA	R	AT	Yes	Yes

Note. E = experimental, QS = quasi-experimental, SA = stand-alone, P = part of comprehensive instruction, T = teacher, R = researcher, UC = usual classroom, AT = alternative treatment, NA = not available

Our analysis suggested that there were six major trends: (a) most of the studies focused on older children in Grade 3 or above; only two studies included young children in kindergarten or Grade 1 (Filippini, 2007; Zoski & Erickson, 2017); also, the majority of research was based on Spanish-speaking L2 English learners in the US; (b) there were more quasi-experimental studies than experimental studies with treatment randomization; (c) eight out of 12 studies implemented morphological instruction as part of a more comprehensive curriculum rather than providing it as a stand-alone treatment; (d) the majority of primary studies reported fidelity and feasibility of morphological instruction; (e) there was great variation in the length of instructional time, ranging from one 45-minute session in

total (Zhang et al., 2010) to 72 sessions totaling 3,780 minutes (Kieffer & Lesaux, 2012); and (f) there was no notable trend with regard to the following three features. First, in terms of learners' language backgrounds in the treatment group, half of previous research studies included L2 English learners only, and the other half mixed L2 and L1 English learners. Second, the instruction was delivered by either teachers or researchers. Last, for the control/comparison condition(s), about half of the research adopted regular classroom settings, that is, "business-as-usual," whereas the other half opted for alternative treatments such as phonological awareness and word decoding training or academic vocabulary learning.

As to the specific characteristics of morphological instruction design, we coded five categories, as shown in Table 2: (a) promoting morphological analysis versus morphological synthesis (*analysis* refers to instruction that guided learners to combine smaller word parts to produce words), (b) providing morphological instruction in oral modality versus written modality, (c) attending to spelling changes, (d) including word formation rules (inflection, derivation, compounding), and (e) including focal word parts (affixes and bases).

Table 2 Features of morphological instruction

Author(s) & year	Analysis vs. synthesis	Oral and written morphology	Word parts: affixes & bases	Word formation rules	Spelling changes	Morphological foils
Badawi (2019)	Both	NA	Bases	Inflection and derivation	NA	NA
Carlo et al. (2004)	Analysis	Both	Bases	Derivation (& some monomorphemic words)	Yes	NA
Crosson & Moore (2017)	Analysis	Both	Bases	Derivation	Yes	Yes
Davidson & O'Connor (2019)	Analysis	Both	Both	Derivation	Yes	NA
Deng (2016)	Analysis	Written	Both	Derivation	NA	NA
Filippini (2007)	Both	Both	Both	Inflection	Yes	NA
Goodwin (2016)	Synthesis	Written	Both	Inflection and derivation	Yes	Yes
Kieffer & Lesaux (2012)	Analysis	Written	Affixes	Derivation	NA	NA
Lesaux et al. (2010)	Analysis	Both	Affixes	Derivation	Yes	NA
Zhang (2016)	Both	Both	Affixes	Derivation	Yes	NA
Zhang et al. (2010)	Both	Both	Bases	Compounding	NA	NA
Zoski & Erickson (2017)	Analysis	Both	Affixes	Inflection and derivation	NA	NA

Note. *Morphological foils* refer to word pairs that are not morphologically related such as *ear* and *earth*; NA = not available

Table 2 shows an emerging trend in previous studies: Researchers focused largely on morphological analysis only (in seven out of 12 studies), included both oral and written modalities (in eight studies), and attended to spelling changes (in seven studies). Another trend is that more than half of the studies (seven out of 12) focused on derived words, the dominant word formation rule in academic English words. There was, nonetheless, one study that targeted inflectional words with participants who were Ggrade 1 students in the USA (i.e., Filippini, 2007); other three studies included both inflectional and derived words; and one study examined compounding with L2 English learners in China (i.e., Zhang et al., 2010). With regard to the reoccurring word parts used in instructional materials, one-third of the studies included affixes only, one-third included bases only, and the rest included both affixes and bases.

To sum up, the analysis of both general program designs and specific features of morphological instruction suggested that most of the evidence was gathered from studies of a quasi-experimental design that implemented morphological instruction as part of a comprehensive curriculum with children in Grade 3 and above. The aim of the explicit morphological instruction in the majority of the studies was to promote morphological analysis of affixed words with/without spelling changes. Half of the studies focused on derived words only; three included both inflected and derived words; and the rest included inflected or compounded words only. Most of the studies reported fidelity and feasibility of the instruction. In what follows, we further examine the impact of morphological instruction on L2 reading development in young learners.

4.2. The impact of morphological instruction on English reading subskills development in young L2 learners

The 12 primary studies can be categorized into three tracks, the details of which are provided in Table 3: (a) Track 1 included seven studies that reported Cohen's *d* for pre- and post-testing comparisons as a result of morphological instruction in group(s); (b) Track 2 consisted of three studies where the research design was similar to that of Track 1, yet different effect sizes were reported (Blake's modified gain ratio in Badawi, 2019; eta-squared (η^2) in Carlo et al., 2014; and Hedges' *g* in Goodwin, 2016); (c) Track 3 was comprised of two studies that adopted a multi-baseline single case design (i.e., Davidson & O'Connor, 2019; Deng, 2016). It is noted that Cohen's *d* and Hedges' *g* are interpreted in a similar way: A small effect = 0.2; a medium effect = 0.5; a large effect = 0.8 (Cohen, 1988; Hedges, 1981). As to η^2 , according to Miles and Shevlin (2001), the benchmarks for small, medium, and large effects are 0.01, 0.06, and 0.14, respectively.

Table 3 Effect sizes (Cohen's *ds*) by independent sample

Author(s) & year	Measurement type	L2 only/mixed	Treatment sample	Control sample	Grade level	Effect size	PA	MA	Vocabulary	Decoding accuracy	Fluency	Spelling	Reading comprehension
Crosson & Moore (2017)	R/S	L2 only	27	NA	11-12	Cohen's <i>d</i>	NA	NA	2.21	NA	NA	NA	NA
Crosson & Moore (2017)	R/S	L2 only	25	NA	6-8	Cohen's <i>d</i>	NA	NA	0.57	NA	NA	NA	NA
Crosson & Moore (2017)	R/S	L2 only	30	NA	9-10	Cohen's <i>d</i>	NA	NA	1.46	NA	NA	NA	NA
Filippini (2007)	R/S	Mixed	25	18	1	Cohen's <i>d</i>	0.44	0.44	0.64	0.3	0.3	NA	Non-significant
Kieffer & Lesaux (2012)	R	Mixed	299	183	6	Cohen's <i>d</i>	NA	0.16	NA	NA	NA	NA	NA
Lesaux et al. (2010)	R/S	Mixed	296	180	6	Cohen's <i>d</i>	NA	0.2	0.39	NA	NA	NA	0.15
Zhang (2016)	R	Mixed	64	45	3-4	Cohen's <i>d</i>	NA	30-50	NA	0.4	NA	NA	NA
Zhang, et al. (2010)	R	L2 only	43	83	5	Cohen's <i>d</i>	NA	0.53	NA	NA	NA	NA	NA
Zoski & Erickson (2017)	R/S	Mixed	17	NA	Kindergarten	Cohen's <i>d</i>	NA	3.96	NA	NA	NA	Non-significant	NA

Note: R = researcher-designed, S = standardized, NA = not available, PA = phonological awareness, MA = morphological awareness

The seven studies in Track 1, which reported Cohen's d for pre- and post-testing results, yielded nine independent samples of 826 participants who received morphological instruction treatment. The participants were kindergarten to Grade 12 students from three countries (China, Singapore, and the USA). Seven reading-related outcomes were reported (as illustrated in Table 3), including PA ($k = 1$), MA ($k = 6$), vocabulary knowledge ($k = 5$), word reading accuracy ($k = 2$), word reading fluency ($k = 1$), spelling ($k = 1$), and reading comprehension ($k = 2$). The effect sizes were medium for PA, small to large for MA, small to large for vocabulary knowledge, small for decoding accuracy and fluency, non-significant for spelling, and small or non-significant for reading comprehension.

Track 2 samples ($k = 3$) included 328 participants in Grades 5 to 7 from two countries (Egypt in Badawi, 2019; and the USA in Carlo et al., 2014 and Goodwin et al., 2016). Badawi (2019) focused on two reading-related outcomes (i.e., MA and reading comprehension), set the acceptable range for effect sizes (Blake's modified gain ratio) between 1.20 and 2.00, and found a small effect of morphological instruction on MA (effect size = 1.28) and a minimal effect on reading comprehension (effect size = 1.06). Carlo et al. (2004) used η^2 to assess the difference in the pre- and post-test results for two outcomes (i.e., vocabulary knowledge and reading comprehension). The authors observed a large effect size for vocabulary knowledge (0.34) and a medium effect size for reading comprehension (0.08). Lastly, the findings of Goodwin (2016) suggested that the intervention versus comparison instruction was moderately effective at supporting vocabulary knowledge (g s were 0.41 and 0.47 for two different vocabulary tasks, respectively) and highly effective for MA ($g = 0.69$), and that non-significant differences were found for word reading fluency and reading comprehension.

Track 3 included two primary studies that had small sample pools and adopted multi-baseline single-case within-subject design (i.e., Davidson & O'Connor, 2019; Deng, 2016). Based on the two studies, a total of 12 participants (independent samples) from Grades 3 to 5 received morphological treatment. Davidson and O'Connor (2019) measured the changes in vocabulary knowledge and observed large effect sizes (Cohen's d s) across participants (ranging from 1.83 to 1.96). Deng (2016) measured gains in both vocabulary knowledge and reading comprehension. For both outcomes, the effect sizes ranged from small to large, 0.34 to 4.76 for vocabulary knowledge and 0.27 to 1.68 for reading comprehension.

In sum, the 12 primary studies have examined the impact of morphological instruction on gains in seven reading-related outcomes, including PA, MA, vocabulary knowledge, word decoding accuracy, word reading fluency, spelling, and reading comprehension. There was notably more evidence for two outcomes: MA and vocabulary knowledge. It seems that, as a result of morphological instruction, there were consistent and positive gains in L2 children's MA and

vocabulary; the effect sizes varied from small to large. However, there were no conclusive findings with PA, reading accuracy and fluency, spelling, and reading comprehension because of the limited sample size ($k_s \leq 2$).

4.3. Trends in the measurement of instructional effects

As stated earlier, about half of the primary studies compared the effect in a morphological instruction treatment group against that in a regular classroom group (e.g., Carlo et al., 2004) or a student group that received alternative treatments such as phonological awareness and decoding training (e.g., Filippini, 2007) or academic vocabulary learning (e.g., Crosson & Moore, 2017). In this section, we further examine how previous studies measured gains for inferring the effect of morphological instruction in light of four issues, as shown in Table 4: (a) reporting within-subject differences (e.g., pre- and post-test differences within the treatment group only) or between-subject differences (e.g., comparing gains between the treatment group and the control group) or mixed; (b) measuring reading-related outcomes based on standardized tests or researcher-designed instruments or mixed; (c) investigating transfer effects in word learning by including learned or novel word items in the tests; and (c) exploring cross-language transfer effects (e.g., whether morphological instruction in English can facilitate the development of reading subskills in another language). Accordingly, there are four major findings.

Table 4 Research design in the selected studies

Author(s) & year	Within-/ between- subject	Standardized (S) / researcher-de- signed (R)	Transfer to novel word reading/learning	Cross-language transfer
Badawi (2019)	Between	R	NA	NA
Carlo et al. (2004)	Between	R	NA	NA
Crosson & Moore (2017)	Within	R/S	NA	NA
Davidson & O'Connor (2019)	Within	R	Yes	NA
Deng (2016)	Within	R	Yes	NA
Filippini (2007)	Both	R/S	NA	NA
Goodwin (2016)	Between	R/S	NA	NA
Kieffer & Lesaux (2012)	Between	R	Yes	NA
Lesaux et al. (2010)	Between	R/S	Yes	NA
Zhang (2016)	Both	R	NA (not explicitly)	Yes (English to Malay)
Zhang, et al. (2010)	Between	R	Yes	Yes (English to Chinese and Chinese to English)
Zoski & Erickson (2017)	Both	R/S	NA	NA

Note. NA = not available

First, our analysis indicated that half of the studies adopted a between-subject design, three other studies used mixed designs, and another three used a within-subject design (Crosson & Moore, 2017; Davidson & O'Connor, 2019; Deng, 2016). In Crosson and Moore's (2017) study, participants experienced two

different interventions, that is, a morphology-focused academic vocabulary intervention versus an academic vocabulary intervention only, which were counterbalanced. Whereas the majority of the studies examined treatment effects in participant groups, Davidson and O'Connor (2019) as well as Deng (2016) focused on individual learners (total $N < 10$). Both studies followed a multi-baseline, single-case design that compared each individual participant's performance across three phases: baseline, intervention, and maintenance. Participants' responses were gathered through testing sessions during the baseline phase before morphological instruction, and there were multiple baselines (three, six, and nine sessions) across individual participants. Both studies tested participants' vocabulary knowledge, including both learned and novel words. Effect size was calculated in the percentage of non-overlapping data (PND, Campbell, 2013). PND scores over 90% indicate high effectiveness; scores between 70% and 90% are considered moderately effective; scores below 70% are questionable. A second finding was that seven of the 12 studies relied on researcher-designed outcome measures only, whereas the other five studies incorporated both researcher-designed and standardized outcome measures.

A third finding was related to the use of novel word items in outcome measures. Only five of the 12 studies explicitly reported that both learned and novel word items were included in pre- and post-tests. Lastly, only two studies examined cross-language transfer effects, namely, effects of morphological instruction in English on the development of reading-related skills in the other language of L2 readers or bilingual children (i.e., Zhang, 2016; Zhang et al., 2010). Specifically, Zhang (2016) implemented instruction on English derivational morphology in Grade 4 English-Malay bilingual children in Singapore and found that the instruction not only led to significant gains in English abilities but also improved children's Malay derivational awareness. In Zhang et al.'s (2010) study on Chinese-speaking learners of English as a foreign language in China, fifth graders received morphological instruction with a focus on compounding in either Chinese or English, while other children did not receive any treatment. The authors found that in the English compounding treatment group, participants with high L2 English reading proficiency transferred English compound awareness to Chinese compound awareness.

In summary, in order to examine the causal effect of morphology on L2 English reading development, the majority of previous studies adopted a between-subject design and compared the pre- and post-instruction testing performance between an English morphological intervention group and a control group. Few studies examined the transfer effects of morphological instruction, namely, learners' application of morphological training in English to new word reading/learning in English or to reading tasks in another language. It should

also be noted that researchers in previous studies sometimes considered the influence of learner-related factors when they assessed instructional effectiveness (e.g., grade level in Crosson & Moore, 2017; L2 English speaking versus L1 English speaking in Kieffer & Lesaux, 2012; and L2 English proficiency in Zhang et al., 2010 reviewed above). The findings from these studies suggest that children in upper grade levels and those who speak a language other than English at home might benefit more from English morphological instruction and that there is a possibility for children to transfer L2 English morphological awareness to their L1 when their L2 English proficiency reaches a certain level.

5. Discussion: Evidence regarding the causal relationship between morphology and L2 English reading development

This scoping review focused on the causal evidence of the impact of morphological instruction on English reading development in a specific learner group (i.e., young L2 learners). To answer the research questions, most of the existing evidence was gathered from studies that implemented morphological instruction as part of a comprehensive curriculum for children in Grades 3 or above in the USA. While there were consistent and significant findings about the positive impact of morphological instruction on learners' MA and vocabulary knowledge, the findings were inconclusive regarding the outcome of PA, word decoding accuracy and fluency, spelling, and reading comprehension, which was mainly due to the limited number of studies available in the literature. Finally, the selected studies primarily adopted a between-subject, quasi-experimental design and compared the pre- and post-instruction testing performance between an English morphological intervention group and a control group. The research measurement instruments were mostly researcher-designed or a mixture of researcher-designed and standardized tests. Notably, very few studies examined the transfer effects of morphological instruction, including gains in reading tasks of novel word items and changes in reading ability in another language (exceptions are discussed later).

Perhaps because our review focuses on L2 learners younger than 18 years old, the findings are both consistent with and divergent from those in previous systematic reviews where mixed learner groups were involved or little attention was paid to any specific learner group (e.g., Carlisle, 2010; Goodwin & Ahn, 2013; Kirby & Bowers, 2017; Nagy et al., 2014). Specifically, consistent with previous reviews, we found positive gains in MA and vocabulary knowledge as a result of morphological instruction, yet the effects on reading comprehension were inconsistent. Likewise, there was insufficient evidence for PA, word reading accuracy and fluency, and spelling as outcomes. On the one hand, the results of our analysis

echo the calls in Carlisle (2010) and Nagy et al. (2014) for more research that targets word forms and reading comprehension as reading-related outcomes. On the other hand, our results could not validate Kirby and Bowers's (2017) proposal that the effect size of morphological instruction is the largest at the sublexical level, followed by the lexical level, and the smallest at the supralexic level.

Another notable finding of this review, which seems to deviate from those of previous reviews, concerns the question of *when* morphological instruction should be implemented. Goodwin and Ahn's (2013) meta-analysis suggested that learners below Grade 3 benefited more from morphological instruction than those in upper grade levels. Yet, this finding did not seem to be the case in our review. Cross and Moore (2017), for example, compared the performance among three grade groups (Grades 6-8, 9-10, 11-12) and found that the largest instructional effects were observed in the oldest group (Grades 11-12). In the present review, the samples were mainly of learners in Grade 3 and higher; only two examined younger children. Crosson and Moore (2017) was the only one of the 12 selected studies that directly examined an age/grade effect. In this regard, our review evidence may not be conclusive and future research should pay more attention to younger (L2) learners. Yet, a greater morphological instructional effect for older, as opposed to younger, L2 learners may be reasonable in that learners may need to achieve an adequate English proficiency (e.g., oral vocabulary and comprehension) to experience the maximum benefits of English morphological instruction. In this respect, Goodwin and Ahn (2013), which did not have an L2 focus, and the present review may not be at all contradictory, but rather show the complex interplay of factors – in the context of this discussion, learners' language backgrounds and English proficiency – that needs to be taken into consideration for morphological instruction.

Lastly, it is worthwhile to note that although only a few studies have examined the transfer effects of morphological instruction (Davidson & O'Connor, 2019; Deng, 2016; Zhang, 2016; Zhang et al., 2010), the research designs adopted by those studies provide much to inform future research. The concept of transfer, though often defined and approached in diverse ways, has received a lot of attention in research literature about reading. On the one hand, morphological instruction should aim to develop a capacity in learners to attend to morphological patterning and apply insights of the patterning in word learning and other literacy activities. Achievement gains should thus not be narrowly restricted to what was taught (e.g., knowledge of affixes). In other words, transfer of taught skills is essential. In this respect, Davidson and O'Connor (2019) as well as Deng (2016), which examined the transfer effect of morphological instruction to learning novel word items and formed the first track of transfer studies reviewed in this paper, shed light on future research on morphology and (L2) reading. In addition,

their 3-phase (pre-instruction, post-instruction, and maintenance), within-subject case design with multiple baselines, in comparisons to the 2-phase (pre- and post-instruction), between-subject design used in the majority of selected studies, seems particularly helpful to address transferred morphological learning and long-term literacy development.

On the other hand, in light of the fact that young L2 readers are often concurrent learners of literacy skills in two or more languages (e.g., English and their L1), a natural issue to consider is whether morphological instruction in one language would benefit reading development in the other language, that is, cross-language transfer of instructional effects. As noted at the beginning of this paper, evidence supporting transfer of reading subskills in L2 or bilingual reading was almost exclusively based on cross-language correlational associations. In this respect, Zhang (2016) and Zhang et al. (2010), which constituted the second track of the transfer studies reviewed in this paper, have expanded our understanding of the impacts of morphological instruction in different bi-/multilingual and instructional contexts (e.g., English as the school subject and the medium of instruction in Singapore, a multilingual society, or English as foreign language in China). More importantly, these two studies both implemented morphological instruction in a treatment group and administered literacy tests in two languages in both the treatment and control groups. Subsequently, both between- and within-group testing score differences can be compared (though Zhang et al. 2010 only reported between-group comparisons results). This design can inform future research that aims to generate causal evidence for the cross-language transfer of reading subskills. Finally, as Carlisle (2010) pointed out, there is a need for cross-linguistic comparisons to explore the language-universal vis-à-vis language-specific effects of morphology in reading development. Based on the positive transfer effects observed in Zhang's (2016) and Zhang et al.'s (2010) studies, it seems that the contribution of morphology to reading development is language universal. Zhang (2016) implemented derivation awareness instruction whereas Zhang et al. (2010) focused on compounding. It is unclear whether the inclusion of different word formation rules (inflection, derivation, compounding) might alter the transfer effects of morphological instruction.

6. Conclusions, limitations, implications and a research agenda

This scoping review synthesized 12 primary studies pertinent to the relationship between morphological instruction and the development of a range of reading-related outcomes in young L2 learners. The evidence was based on studies published between 2004 and 2019 ($N = 1,535$), which focused on reading English as the target language and involved participants between kindergarten and Grade

12 from four countries (China, Egypt, Singapore, and the USA). It can be tentatively concluded that explicit morphological instruction has a positive impact on morphological awareness and vocabulary knowledge in L2 learners in Grade 3 and above. However, there is insufficient evidence to conclude whether morphological instruction is more or less effective for other important outcomes, including phonological awareness, word reading accuracy and fluency, spelling, and reading comprehension; or whether morphological instruction is equally or more beneficial for younger children. Emerging evidence has suggested that the effects of morphological instruction delivered in English is transferrable to novel word learning in English and to reading development in another language. However, because of the relatively small independent samples and vast variation in the primary studies, we could not conduct a meta-analysis to test moderating effects of construct-, learner-, linguistic-, instruction- and assessment-related factors. Another limitation was that the literature search did not combine other key words such as *vocabulary instruction/learning* and *reading development*, which could have affected the sample pool of this review. Also, the coding was conducted by the first author and only intra-coder coding agreement was reported. More systematic reviews with a more rigorous literature search and inclusion/exclusion criteria, as well as inter-coding reliability, are needed.

A few implications can be drawn for pedagogical practice. According to the majority of the selected studies reviewed above, it is feasible for teachers to implement explicit morphological instruction in the regular curriculum for L2 learners in Grades 3 or above (for a more concrete design, see Goodwin et al., 2012). Educators and learners can anticipate positive gains in morphological awareness and vocabulary learning. It has also been recommended that the instruction should not just focus on *analyzing* word-internal structure but also engage students in problem-solving or inquiry-based activities to *produce* novel complex words (Kirby & Bowers, 2017).

To improve the scientific understanding of the causal relationship between morphology and L2 reading development in young learners, researchers might consider the following agenda: (a) measure transfer effects; these effects should be tapped by including both learned and novel word items in testing instruments, measuring changes in an additional language other than the language targeted in morphological instruction, and administering immediate and delayed post-tests; (b) adopt a cross-linguistic perspective; it is necessary for researchers to include learners of less commonly examined (nonalphabetic) language backgrounds and consider different word formation rules (inflection, derivation and compounding) in the future; (c) explore morphological instruction in younger children; although it is often held that refined morphological awareness will not emerge until the upper grade levels in English-speaking children (e.g.,

Berninger et al., 2010), recent studies have suggested that it is feasible and effective to implement morphological instruction for younger children (Apel et al., 2013; Devonshire et al., 2013); yet, it is still unclear as to whether there is any long-term benefit to L2 learners' reading development by implementing morphological instruction at the kindergarten level and Grades 1 to 2 when children typically transition from speaking to *learning to read*; and (d) expand the scope to different linguistic and educational contexts; so far, the majority of evidence is based on Spanish-speaking L2 English learners in the USA, and the findings might not be readily generalizable to other language and educational contexts.

References

* indicates a study selected in the scoping review

- Apel, K., Brimo, D., Diehm, E., & Apel, L. (2013). Morphological awareness intervention with kindergartners and first-and second-grade students from low socioeconomic status homes: A feasibility study. *Language, Speech, and Hearing Services in Schools, 44*(2), 161-173.
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19-32.
- *Badawi, M. F. A. (2019). The effect of explicit English morphology instruction on EFL secondary school students' morphological awareness and reading comprehension. *English Language Teaching, 12*(4), 166-178.
- Berninger, V. W., Abbott, R. D., Nagy, W., & Carlisle, J. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research, 39*(2), 141-163.
- 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.
- Brandes, D. R., & McMaster, K. L. (2017). A review of morphological analysis strategies on vocabulary outcomes with ELLs. *Insights into Learning Disabilities, 14*(1), 53-72.
- Campbell, J. M. (2013). Commentary on PND at 25. *Remedial and Special Education, 34*(1), 20-25.
- Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly, 45*(4), 464-487.
- *Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., Lively, T. J., & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188-215.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- *Crosson, A. C., & Moore, D. (2017). When to take up roots: The effects of morphology instruction for middle school and high school English learners. *Reading Psychology, 38*(3), 262-288.
- *Davidson, S. J., & O'Connor, R. E. (2019). An intervention using morphology to derive word meanings for English language learners. *Journal of Applied Behavior Analysis, 52*(2), 394-407.
- *Deng, Q. (2016). *The effect of a self-regulated vocabulary intervention on word knowledge, reading comprehension, and self-regulated learning for elementary English language learners* (Doctoral dissertation, University of Nebraska). Retrieved from ProQuest Dissertations Publishing Publication (No. 10102763).

- Devonshire, V., Morris, P., & Fluck, M. (2013). Spelling and reading development: The effect of teaching children multiple levels of representation in their orthography. *Learning and Instruction, 25*, 85-94.
- *Filippini, A. L. (2007). *Effects of a vocabulary-added instructional intervention for at-risk English learners: Is efficient reading instruction more effective?* (Doctoral dissertation, University of California). Retrieved from ProQuest Dissertations Publishing Publication (No. 3274436).
- Foster, E. M. (2010). Causal inference and developmental psychology. *Developmental Psychology, 46*(6), 1454-1480.
- Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences, 35*(5), 263-279.
- *Goodwin, A. P. (2016). Effectiveness of word solving: Integrating morphological problem-solving within comprehension instruction for middle school students. *Reading and Writing, 29*(1), 91-116.
- Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia, 60*(2), 183-208.
- 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.
- Goodwin, A., Lipsky, M., & Ahn, S. (2012). Word detectives: Using units of meaning to support literacy. *The Reading Teacher, 65*(7), 461-470.
- Goodwin, A. P., & Perkins, J. (2015). Word detectives: Morphological instruction that supports academic language. *The Reading Teacher, 68*(7), 510-523.
- Hayashi, Y., & Murphy, V. A. (2013). On the nature of morphological awareness in Japanese-English bilingual children: A cross-linguistic perspective. *Bilingualism: Language and Cognition, 16*(1), 49-67.
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6*(2), 107-128.
- Henbest, V. S., & Apel, K. (2017). Effective word reading instruction: What does the evidence tell us? *Communication Disorders Quarterly, 39*(1), 303-311.
- *Kieffer, M. J., & Lesaux, N. K. (2012). Effects of academic language instruction on relational and syntactic aspects of morphological awareness for sixth graders from linguistically diverse backgrounds. *The Elementary School Journal, 112*(3), 519-545.
- Kirby, R. J., & Bowers, P. N. (2017). Morphological instruction and literacy: Binding phonological, orthographic, and semantic features of words. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 437-461). John Benjamins.
- Larsen-Freeman, D. (2013). Transfer of learning transformed. *Language Learning, 63*, 107-129.
- *Lesaux, N. K., Kieffer, M. J., Faller, S. E., & Kelley, J. G. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for

- linguistically diverse students in urban middle schools. *Reading Research Quarterly*, 45(2), 196-228.
- McBride-Chang, C., Liu, P. D., Wong, T., Wong, A., & Shu, H. (2012). Specific reading difficulties in Chinese, English, or both: Longitudinal markers of phonological awareness, morphological awareness, and RAN in Hong Kong Chinese children. *Journal of Learning Disabilities*, 45(6), 503-514.
- Miles, J. & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. Sage.
- Nagy, W. E., Carlisle, J. F., & Goodwin, A. P. (2014). Morphological knowledge and literacy acquisition. *Journal of Learning Disabilities*, 47(1), 3-12.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research*, 78(1), 33-84.
- Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in second language research: Choices and challenges. *Annual Review of Applied Linguistics*, 30, 85-110.
- Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods*, 5(4), 371-385.
- Reed, D. K. (2008). A synthesis of morphology interventions and effects on reading outcomes for students in grades K–12. *Learning Disabilities Research & Practice*, 23(1), 36-49.
- Saiegh-Haddad, E., & Geva, E. (2008). Morphological awareness, phonological awareness, and reading in English–Arabic bilingual children. *Reading and Writing*, 21(5), 481-504.
- Seel, N.M. (2012) Transfer of learning. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning*. Springer. Retrieved from: <https://doi.org/10.1007/978-1-4419-1428-6>
- *Zhang, D. (2016). Morphology in Malay-English biliteracy acquisition: An intervention study. *International Journal of Bilingual Education and Bilingualism*, 19(5), 546-562.
- Zhang, D., Koda, K., & Sun, X. (2014). Morphological awareness in biliteracy acquisition: A study of young Chinese EFL readers. *International Journal of Bilingualism*, 18(6), 570-585.
- *Zhang, J., Anderson, R. C., Li, H., Dong, Q., Wu, X., & Zhang, Y. (2010). Cross-language transfer of insight into the structure of compound words. *Reading and Writing*, 23(3-4), 311-336.
- *Zoski, J. L., & Erickson, K. A. (2017). Multicomponent linguistic awareness intervention for at-risk kindergarteners. *Communication Disorders Quarterly*, 38(3), 161-171.