Applied Psycholinguistics (2024), **45**, 745–765
doi:10.1017/S0142716424000213

ORIGINAL ARTICLE

The development of morphological awareness and vocabulary: What influences what?

Germán Grande[1], Vassiliki Diamanti[2], Athanassios Protopapas[2,3], Monica Melby-Lervåg[2,3] and Arne Lervåg[1,3]

[1]Department of Education, University of Oslo, Oslo, Norway, [2]Department of Special Needs Education, University of Oslo, Oslo, Norway and [3]CREATE – Centre for Research on Equality in Education, University of Oslo, Oslo, Norway
Corresponding author: Arne Lervåg; Email: a.o.lervag@create.uio.no

(Received 10 May 2023; revised 22 December 2023; accepted 17 May 2024; first published online 24 September 2024)

Abstract
The awareness of words' morphological structure has been thought to allow generalizing meaning to other, similarly constructed words. Conversely, a large vocabulary is thought to facilitate the recognition of words' morphological regularities, thereby contributing to morphological awareness. For this reason, morphological awareness and vocabulary have been suggested to be reciprocally associated across development. We followed 242 (girls $=$ 119) Norwegian preschoolers ($M_{\text{age}} = 5.5$ years) from preschool through Grade 2 and examined the cross-lagged relations between morphological awareness (inflections and derivations) and vocabulary (receptive and expressive). Our results confirm that the traditional cross-lagged panel model shows significant cross-lagged relations between morphological awareness and vocabulary, as previous studies have shown. However, no cross-lagged relations were found when we accounted for longitudinal measured stability through a cross-lagged panel model with lag-2 paths or unmeasured stability through the random intercept cross-lagged panel model. We found that approximately 50% of the variation in morphology and vocabulary was due to highly stable and invariant factors across grades. We discuss how the significant cross-lagged relations found in previous studies could have been due to their not accounting for the right type of stability when using longitudinal panel data.

Keywords: cross-lagged panel model; morphological awareness; random intercept cross-lagged panel model; vocabulary

The development of morphological awareness and vocabulary: What influences what?

An important aspect of children's morphological skill development is often referred to as gaining *morphological awareness*, and it involves the ability to reflect on and manipulate the morphemic structure of words (Carlisle, 1995). Thus, morphological awareness implies an explicit understanding of how morphemes can alter words' function through the process of inflection (e.g., from *close* to *closed*) or a change in meaning through derivation (e.g., *enclosed*; Anglin, 1993; Kuo & Anderson, 2006). For instance, the word *enclosed* consists of three meaning units called *morphemes*: *en-* (which means to bring something into a condition), *close* (which means to block against entry or passage), and *-ed* (which indicates the past tense of a regular verb). Morphological awareness is considered a causal driver in vocabulary development because awareness of the recurrent parts of words enables one to deduce the meanings of new words with similar morphological constructions (Carlisle, 2007; e.g., *enhanced* and *encountered* or *closeness* and *disclosure*). However, the opposite has also been suggested—namely that vocabulary serves as a foundation for morphology because a rich vocabulary can enable children to appreciate words' morphological regularities (e.g., the suffix *-ing*: reading, opening, gathering), thereby enhancing morphological awareness (Sparks & Deacon, 2013).

This hypothesized reciprocal relation is expected to grow stronger through schooling because academic words increasingly build on multiple morphemes (Berninger et al., 2010; Nagy & Anderson, 1984) and because children are expected to incorporate morphological information in order to create lexical representations, especially when reading (Levesque et al., 2021; Perfetti, 2007). For this reason, frequent exposure to new words can enhance the awareness of morphological regularities, increasing opportunities to understand new word meanings.

Cross-lagged effects between morphological awareness and vocabulary

Based on theories of a reciprocal relationship between morphology and vocabulary, Sparks and Deacon (2013) studied the cross-lagged relationship between morphological awareness and receptive vocabulary among English-speaking children from Grade 2 to Grade 3. A unidirectional relationship was found in which Grade 2 morphological awareness had a significant cross-lagged effect on Grade 3 receptive vocabulary (i.e., vocabulary regressed on earlier morphology) when accounting for morphology's lagged effect (i.e., the autoregressor) and covariates such as phonological skills, reading, and nonverbal reasoning.

The same cross-lagged association was studied among Mandarin-, Cantonese-, and Korean-speaking children. As McBride-Chang et al. (2008) explained, compounding morphology (e.g., *sun* + *shine* = sunshine) is highly transparent and effective in word formation using these languages' character-based orthographies. Their results showed bidirectional cross-lagged effects between the awareness of compound morphology and expressive vocabulary from kindergarten to Grade 1 in the three languages after controlling for nonverbal reasoning and phonological skills. Interestingly, another study of Mandarin speakers found no cross-lagged relations between the awareness of compound morphology and a measure combining receptive and expressive vocabulary from Grade 1 to Grade 2 (Dulay et al., 2021). Finally, Kieffer and Lesaux (2012) examined the growth trajectories of awareness of derivational morphology and receptive vocabulary from Grade 4 to Grade 7 in Spanish-speaking students learning English. They found that growth trajectories correlated after controlling for word reading and phonological

https://doi.org/10.1017/S0142716424000213 Published online by Cambridge University Press

awareness. However, the Grade 4 skills for one process did not predict the other process's later growth.

Thus, previous studies have presented mixed evidence of the cross-lagged effects between morphological awareness and vocabulary. A reason for this mixed evidence could be that studies have either combined measures of inflectional and derivational morphology (Sparks & Deacon, 2013) or focused on word compounding (Dulay et al., 2021; McBride-Chang et al., 2008) or derivational morphology (Kieffer & Lesaux, 2012) alone. Comparably, vocabulary has been measured through tasks that tap receptive vocabulary (Kieffer & Lesaux, 2012; Sparks & Deacon, 2013), expressive vocabulary (McBride-Chang et al., 2008), or a combination of both vocabularies (Dulay et al., 2021). These combined factors may explain the inconsistent findings among character-based orthographies (Dulay et al., 2021; McBride-Chang et al., 2008) and alphabetic orthographies (Kieffer & Lesaux, 2012; Sparks & Deacon, 2013). Indeed, meta-analytic evidence has indicated that the correlations between morphology and vocabulary tasks can vary, depending on whether tasks assess inflectional or derivational morphology and expressive or receptive vocabulary (Lee et al., 2022). Thus, accounting for such differences may be necessary when studying bidirectional associations early in development.

Challenges in interpreting longitudinal bidirectional associations

A challenge in interpreting these studies is that cross-lagged relations have typically been estimated using the traditional cross-lagged panel model (CLPM; e.g., Dulay et al., 2021; McBride-Chang et al., 2008; Sparks & Deacon, 2013). The CLPM accounts for cross-lagged effects (i.e., regressions across domains) and lagged effects (i.e., autoregressors) but has been shown to overestimate cross-lagged effects when stable individual differences over time are ignored (Hamaker et al., 2015; Lucas, 2023; Lüdtke & Robitzsch, 2022; Usami et al., 2019). Further, since the CLPM assumes that all confounders are measured and controlled for (except those controlled for by autoregressors), the usefulness of this kind of model in testing causal theories about reciprocal relationships has been contested (Hamaker et al., 2015; Lucas, 2023; Lüdtke & Robitzsch, 2022; Usami et al., 2019). This criticism also applies to the reciprocal relationship posited between morphological awareness and vocabulary.

To address such criticism, different modeling strategies have been proposed. One strategy is the cross-lagged panel model with lag-2 effects (CLPM2; Lüdtke & Robitzsch, 2022), which controls for delayed time effects, not just effects between consecutive time points as in the traditional CLPM. The CLPM2 relies on the idea that controlling for all previous time exposure provides a more rigorous covariate control and can reduce the potential for unmeasured confounding (for more details, see VanderWeele, 2021; VanderWeele et al., 2020). Still, as with the CLPM, the CLPM2 requires the measurement and control of all potential confounders to interpret cross-lagged relations as causal. In contrast, the random intercept cross-lagged panel model (RI-CLPM) separates the stable variation that is shared between persons over time (i.e., time-invariant stability) from within-person changes (Hamaker et al., 2015). In this sense, the cross-lagged relations within participants can be causally interpreted to a greater extent because all the unobserved confounding that is stable over time (e.g., genes, nonverbal abilities, socioeconomic status) can be parametrically controlled (Hamaker et al., 2015; Mulder & Hamaker, 2021; Usami et al., 2019). Empirical studies of language and mathematics skills have demonstrated that accounting for differences in stability between and within persons results in different cross-lagged relations (Brinchmann et al., 2018) or even opposite-direction effects than those estimated by the CLPM (Bailey et al., 2020).

The current study

In this study, we aim to expand the work on associations between morphological awareness and vocabulary by examining the cross-lagged effects of inflectional and derivational on expressive versus receptive vocabulary. Further, unlike previous studies (Dulay et al., 2021; McBride-Chang et al., 2008; Sparks & Deacon, 2013), we controlled for measurement error since uncorrected measurement error introduces biases (Cole & Preacher, 2014) when studying structural relations. These biases can be particularly problematic if the extent of measurement error differs between two processes when studying reciprocal relationships.

Studying Norwegian morphology can be informative since Norwegian shares similarities with other alphabetic orthographies. It has a complex syllabic structure and a semi-transparent orthography (Seymour et al., 2003) that often requires attention to silent letters (e.g., utlen*d*ing, where the d is not pronounced but signals the length of the preceding vowel) and complex graphemes (e.g., produk*sj*on, with the two-letter grapheme sj denoting /ʃ/). Norwegian inflectional morphology is slightly more complex than English, as it marks noun and adjective gender and number, verb tense, and noun definiteness. Norwegian derivational morphology is more similar to English in that it is extensive and allows for productive and consistent meaning generalization.

Our study spanned the critical period from preschool through early elementary grades. Norwegian children start formal literacy instruction at the beginning of Grade 1, having received mainly informal language stimulation throughout preschool. Therefore, if we found significant stable differences in morphological awareness and vocabulary between children—as have been found for other skills (Bailey et al., 2020; Brinchmann et al., 2018)—then we would need to ascribe this stability to common factors that affect the development of both skills before reading instruction begins. Previous studies have not captured this distinction, having typically assessed children after they had received at least some literacy instruction.

Regarding cross-lagged relations, we hypothesized that vocabulary could more substantially affect derivational morphology, than the opposite relation, because of the complexity of the derivational domain and its longer developmental trajectory (Anglin, 1993; Lee et al., 2022). Moreover, we hypothesized that the faster developmental trajectory of inflectional morphology (Lee et al., 2022; Ragnarsdóttir et al., 1999) could likely affect vocabulary because vocabulary continues to grow throughout a child's school years. Therefore, we aim to answer the following research questions:

Can we find cross-lagged relations between morphological awareness and vocabulary when estimating with the more rigorous CLPM2 and the RI-CLPM instead of the classic CLPM used in former studies?

https://doi.org/10.1017/S0142716424000213 Published online by Cambridge University Press

To what degree can the relation between morphological awareness and vocabulary be explained by stable trait-like factors?

Method

Participants

Two hundred and forty-two Norwegian-speaking children (119 girls) with a mean age of 5.5 years ($SD = 3.51$ months, range = 59–72 months) were assessed once per year—in preschool (T1), Grade 1 (T2), and Grade 2 (T3). Children were recruited from 61 preschools (and subsequently 34 schools) outside of Oslo, in southeastern Norway, in municipalities approximating the national average educational level and socioeconomic status (Statistics Norway, 2021a, 2021b). There was an average of 4.28 and 7.62 children per preschool and school, respectively. Written consent was obtained from parents and oral assent from the children themselves. Two hundred and fifty-nine children were part of the original sample. Children with a first language other than Norwegian or a severe learning or developmental disorder diagnosis, such as sensory impairment, autism, or intellectual disability, were excluded. In addition, five children were excluded because either the child themself or the parents wanted to withdraw from the study. Our project was approved by the Norwegian Agency for Shared Services in Education and Research (2024).

Design and procedure

Trained research assistants assessed children individually at their respective preschools and schools. The present study comprised well-known standardized vocabulary tests that had been adapted to Norwegian and research-developed tasks assessing awareness of inflectional and derivational morphology. These tasks belonged to a larger test battery assessing cognitive, (pre)reading, and (pre)math skills (not reported here).

Measures

Morphological awareness measures

Inflectional Production Task. The inflectional production task (adapted from Diamanti et al., 2018; Diamanti et al., 2017) involved asking children to produce an inflection of a target pseudoword from a booklet of 16 picture items. These items comprised pseudoverbs, pseudonouns, and pseudoadjectives. For instance, when presenting a picture of a turtle drawing on a piece of paper, the examiner said, “The turtle colors the *paper” (åmtet /ˈomte/, a definite neuter singular pseudonoun in Norwegian meant to refer to the sheet of paper and constructed to match the phonological structure of the real word /arke/ meaning sheet of paper) and then showed a second picture in which the turtle drew on several pieces of paper. In the subsequent sentence, “The turtle colors . . . ”, the child was required to say “the *papers” (åmtene /ˈomtene/) in order to reflect the change from the singular to the plural.

Derivational Production Task. In the derivation production task (adapted from Diamanti et al., 2018; Diamanti et al., 2017), the child was asked to provide a derivation of a target word (i.e., a real word). For instance, when presenting a picture of a cat with lots of hair, the examiner said, "The cat has a lot of hair (hår /hoːr/); the cat is very ... " and the correct reply would be "hairy" (hårete /hoːrete/). We included four trial items and 14 picture items. Most items required participants to derive nouns, adjectives, and adverbs from verbs or verbs from nouns.

Vocabulary measures

Receptive Vocabulary. We measured receptive vocabulary using the Norwegian translation of the British Vocabulary Scale-II (BPVS; Dunn et al., 1997), which has been adapted and normed for a Norwegian sample (Lyster et al., 2010). The test comprised 144 items distributed across 12 blocks of increasing difficulty.

Expressive Vocabulary. We measured expressive vocabulary among preschool children using an adapted version of the word definition subscale of the Wechsler Preschool and Primary Scale of Intelligence, 4th edition (WPPSI-IV; Wechsler, 2012). For participants in Grades 1 and 2, the expressive vocabulary scale from WIPPSI-IV was extended by adding the final 13 items from the vocabulary scale of the Norwegian adaptation of WISC-IV (Wechsler, 2003).

Control measures

Phonological Awareness. To measure phonological awareness (PA) in preschool, a phoneme isolation task comprising 24 items was used. The child was asked to isolate the initial or final phoneme from a word by articulating the target phoneme (e.g., "What is the last sound in lam?" correct answer: "/ˈm/"). Practice items were given before each block of 12 initial phonemes and 12 final phonemes. The task was discontinued after six consecutive incorrect responses.

Working Memory. To measure preschool working memory (WM), we used the backward digit span task from the Working Memory Test Battery for Children (Pickering & Gathercole, 2001). In this task, the administrator spoke a sequence of digits to the child, asked them to recall the digits in reverse order, and scored the answers as *correct* or *incorrect*. We used the sum of correct responses to measure WM.

Nonverbal Reasoning. We measured preschool children's nonverbal ability (NVIQ) with the matrix task from the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-IV; Wechsler, 2012). The test required the child to visually identify a missing piece from a matrix and choose the piece that completed the matrix from an array of possible choices. The test comprised 26 items of increasing complexity and was administered without time constraints.

Analyses

All the longitudinal models were estimated using the Structural Equation Modeling (SEM) framework in Mplus Version 8.3 (Muthén & Muthén, 1998–2017). The psych package (Revelle, 2022) in R was used for descriptive statistics and

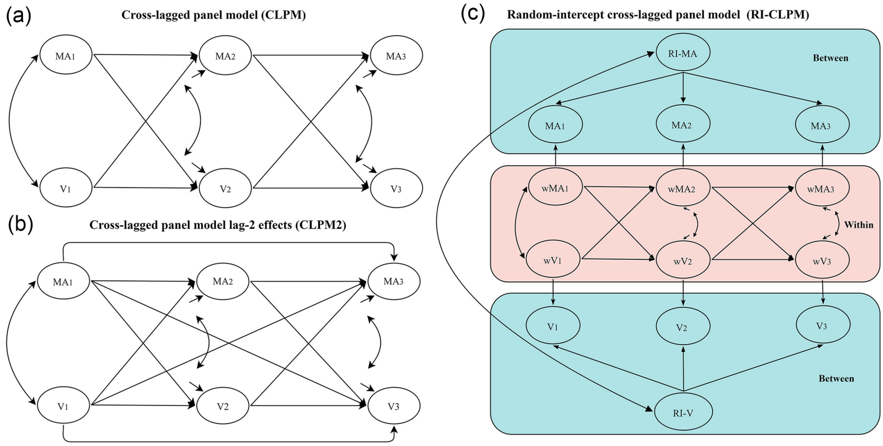

Figure 1. Three model types to estimate cross-lagged relations with varying degrees of control of confounding.

Note. MA = morphological awareness; V = vocabulary; RI = random intercept; the ellipses represent latent variables, the double-headed arrows represent covariances, and the single-headed arrows represent regression paths.

estimations of internal consistency. Data and analysis scripts and outputs (Mplus and R files) are available at https://osf.io/wz6pv/.

Design effects (Peugh, 2010) for the outcome variables ranged between 1.00–1.18 and were thus clearly below the critical threshold of 2.0, suggested by Muthén (1991) and Muthen and Satorra (1995) with the exception of expressive vocabulary in Grade 1 (design effect = 2.03). To account for clustering effects, we adjusted the standard errors of the models involving expressive vocabulary using the Huber-White sandwich estimator (Type = complex) in Mplus. The ordinary maximum likelihood estimator (ML) was otherwise used.

To reduce measurement biases (Cole & Preacher, 2014), we corrected measurement error by fixing each indicator's error variance to the inverse of its estimated reliability multiplied by its variance. This approach is considered optimal when only one indicator is available per construct (Brown, 2015, pp. 122–127), as in the present study.

To test the cross-lagged relationship between morphological awareness and vocabulary, we compared the results of three types of models. The first model was the traditional CLPM, shown in Figure 1(a), which estimated the cross-lagged effect of morphological awareness on vocabulary when controlling for lagged vocabulary (i.e., the autoregressor) and vocabulary's effect on morphological awareness when controlling for lagged morphological awareness. The second model was the CLPM2, shown in Figure 1(b). The CLPM is nested within the CLPM2, but the CLPM2 differs by also accounting for the second-order lagged and cross-lagged effects (e.g., T1 morphology and vocabulary → T3 morphology and vocabulary; Lüdtke & Robitzsch, 2022). The third model, the RI-CLPM, is shown in Figure 1(c) (Hamaker et al., 2015; Mulder & Hamaker, 2021). The CLPM is also nested within the RI-CLPM, but the RI-CLPM differs in that it includes latent variables that are referred to as random intercepts and reflect stable trait-like deviation from

https://doi.org/10.1017/S0142716424000213 Published online by Cambridge University Press

the group means. By including these random intercepts in the model, we are able to separate the stable trait-like variation that they reflect (i.e., between-person variation) from within-person fluctuations over time (i.e., within-person variation). The within-person variation is the temporal deviations from expected scores (expected from the group means and the stable deviations from the group mean) of each individual at each time point (Hamaker et al. 2015; Usami et al, 2019). Thus, in this model, the random intercept captures all stable unmeasured variables that affect the observed variables. Accordingly, it adjusts the within-person lagged and cross-lagged regressions for unmeasured confounders that do not change across time. It should also be mentioned that it is possible, when having four or more time points, to add a growth factor (slope) in addition to the random intercept to this model, turning the model into what has been referred to as a latent curve model with structured residuals (LCM-SR; Curran et al., 2014). This can be important in cases where there is individual variation in growth, which might confound the lagged and cross-lagged relations (Falkenström et al., 2023). However, in the current dataset, simple latent growth models showed that there was no need for adding such a factor, as variation around the growth rate was not significant in any of the four variables.

We then tested if the less restricted CLPM2 and RI-CLPM fit the data better than the more restricted CLPM to see if second-order effects or stable trait factors might confound the cross-lagged effects of the CLPM. The CLPM and RI-CLPM were compared with a chi-bar-square difference test (for details, see Hamaker et al., 2015; Kuiper, 2018; Stoel et al., 2006), and CLPM and CLPM2 were compared with Satorra-Bentler chi-square (Satorra & Bentler, 2010) and chi-square difference tests for the models with and without design effects, respectively.

Further, before addressing our research question about whether cross-lagged relations between morphological awareness and vocabulary can be found in the more rigorous CLPM2 and the RI-CLPM instead of the classic CLPM, we first tested if the lagged and cross-lagged regressions changed across time between constituent time points in the three models across time. Regression parameters that did not change across time were fixed to be equal in the final models.

Additionally, since CLPM and CLPM2 require all potential confounders to be measured and controlled for (except those controlled for through autoregressors), for comparison purposes, we re-estimated these models while controlling morphological awareness and vocabulary (T1–T3) for preschool WM, NVIQ, and PA (see Tables 5s and 6s in the online supplement).

In sum, to test the hypothesized causal relations between morphological awareness and vocabulary, we used the following four models: *inflectional morphology-expressive vocabulary*, *inflectional morphology-receptive vocabulary*, *derivational morphology-expressive vocabulary*, and *derivational morphology-receptive vocabulary*.

Missing data

After completing the data collection in Grade 1, 97.8% of data was present for the 242 participants, followed by 93.1% of the data after data collection in Grade 2. Due to a national lockdown in 2020 caused by COVID-19 pandemic, examiners could not complete some assessments in their assigned schools leading to variable

https://doi.org/10.1017/S0142716424000213 Published online by Cambridge University Press

completion rates in Grade 2. We performed a Kruskal Wallis test through the finalfit package (Harrison et al., 2023) in R to examine if children with missing patterns in morphological awareness and vocabulary in Grade 2 had different mean performance in variables in earlier grades or in terms of gender or the families' socioeconomic status. We found no relations between the missing patterns in Grade 2 and other variables. An exception was inflectional morphology, for which children with missing values in Grade 2 inflectional morphology had significantly different means in Preschool inflectional morphology (not missing, $M = 7.2$ $SD = 4.1$; missing, $M = 5.5$ $SD = 3.7$ $p = .034$). The missing data were handled by using Full Information Maximum Likelihood (FIML) under the assumption of missing at random (MAR). For more details, see online supplement.

Results

Descriptive

Means, standard deviations, maximum scores, and reliability on the three measurement occasions for all measures are shown in Table 1. Correlations between the observed measures are presented in Table 2. The concurrent correlations between inflectional and derivational morphology were significant across time points (ranging from $r = .41$ to .49). The longitudinal correlations within tasks were stable across the three grades for inflectional morphology ($r =$ .49–.73), derivational morphology ($r =$ .28–.39), expressive vocabulary ($r =$ .32–.39), and receptive vocabulary ($r =$ .41–.49), and all were significant ($p < .001$).

Model comparisons

The RI-CLPM fit the data better than the CLPM for all the relations—inflectional morphology and expressive vocabulary ($\Delta\overline{\chi}^2$ (3) $= 22.339$ $p < .001$), inflectional morphology and receptive vocabulary ($\Delta\overline{\chi}^2$ (3) $= 9.502$ $p = .014$), derivational morphology and expressive vocabulary ($\Delta\overline{\chi}^2$ (3) $= 24.242$ $p < .001$) and derivational morphology and receptive vocabulary ($\Delta\overline{\chi}^2$ (3) $= 15.682$ $p = .001$). This suggests that there are stable trait-like variations in the variables. Similarly, since the chi-square for all the CLPMs was significant (inflectional morphology and expressive vocabulary: $\chi^2 = 22.256$ (4), $p < .001$; inflectional morphology and receptive vocabulary: $\chi^2 = 9.808$ (4), $p < .043$; derivational morphology and expressive vocabulary : $\chi^2 = 23.689$ (4), $p < .001$; derivational morphology and receptive vocabulary: $\chi^2 = 15.813$ (4), $p = .003$) it also follows that the saturated CLPM2s were significantly better than the CLPMs. This suggests that the first-order effects do not fully cover all the lagged and/or cross-lagged effects in the data. For more details about the model comparisons, see online supplement Tables S1–S4.

Next, we found that there were no significant changes across time in the lagged and cross-lagged effects for any of the CLPM2s. For the CLPM and RI-CLPMs, there were no significant changes across time for the lagged and the cross-lagged regressions in the models testing the relations between vocabulary and derivational morphology. In the CLPM and RI-CLPMs, testing relations between vocabulary and

Table 1. Descriptive statistics and reliability of the observed scores

Measure	Preschool (T1)				First Grade (T2)				Second Grade (T3)			
	M (SD)	Min–Max	ω_t	N	*M (SD)*	Min–Max	ω_t	N	*M (SD)*	Min–Max	ω_t	N[1]
Inflectional morphology	7.00 (4.10)	0–16	.86	241	10.72 (3.78)	1–16	.89	223	10.59 (3.83)	1–16	.87	211
Derivational morphology	4.62 (2.27)	0–11	.67	241	5.57 (2.32)	0–12	.68	237	7.25 (2.34)	1–13	.65	231
Expressive vocabulary	24.11 (7.01)	5–43	.89	236	27.15 (6.54)	9–43	.81	236	32.13 (6.18)	13–49	.80	171
Receptive vocabulary	63.97 (11.90)	37–91	.94	237	77.30 (11.58)	53–112	.93	237	86.96 (12.21)	48–113	.91	181
Phonological awareness	11.40 (6.60)	2–24	.95	241								
Nonverbal reasoning	11.67 (5.30)	1–22	.92	239								
Working memory	5.81 (3.49)	0–14	.87	236								

Note. ω_t = reliability estimated with McDonald's Omega total; [1] due to the COVID-19 pandemic, part of our assessments in T3 was interrupted by a national lockdown, resulting in fewer children participating in T3.

https://doi.org/10.1017/S0142716424000213 Published online by Cambridge University Press

Table 2. Estimated bivariate correlations among observed measures

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Inflectional morphology (T1)	—													
2. Inflectional morphology (T2)	.50***	—												
3. Inflectional morphology (T3)	.49***	.73***	—											
4. Derivational morphology (T1)	.49***	.33***	.29***	—										
5. Derivational morphology (T2)	.36***	.41***	.40***	.39***	—									
6. Derivational morphology (T3)	.28***	.35***	.42***	.31***	.28***	—								
7. Expressive vocabulary (T1)	.26***	.20**	.22***	.15*	.15*	.19**	—							
8. Expressive vocabulary (T2)	.27***	.36***	.30***	.13*	.22**	.18**	.35***	—						
9. Expressive vocabulary (T3)	.30***	.40***	.31***	.18**	.24**	.30***	.39***	.32***	—					
10. Receptive vocabulary (T1)	.40***	.38***	.34***	.31***	.25***	.18**	.41***	.30***	.32***	—				
11. Receptive vocabulary (T2)	.25***	.25***	.25***	.20**	.34***	.11	.34***	.35***	.26***	.45***	—			
12. Receptive vocabulary (T3)	.32***	.34***	.33***	.23**	.27***	.16*	.36***	.28**	.40***	.41***	.49***	—		
13. Phonological awareness (T1)	.36***	.30***	.38***	.23***	.12	27***	.23***	.14*	.15*	.27***	.16**	.17*	—	
14. Nonverbal reasoning (T1)	.21**	.07	.14*	.10	.14*	14*	19**	.09	.12	.17**	.13*	.23**	.11	—
15. Working memory (T1)	.37***	.24***	.23***	.22**	.13	24***	14*	.20**	.36***	.22***	.16*	.20**	.41***	.18**

Note. T1 = Preschool; T2 = Grade 1; T3 = Grade 2; *** = $p < .001$; ** = $p < .010$; * = $p < .05$.

inflectional morphology, the lagged and cross-lagged effects did not change across time, with the exception of the lagged regressions for inflectional morphology (for details about model selection, see Tables S1 to S4 in the online supplement). As a consequence of these comparisons, in subsequent analyses we fixed all the lagged and cross-lagged effects to be equal across time, with the exception of the lagged effects of inflection morphology.

Inflectional morphology and expressive vocabulary

As Tables 3 and 4 show, both the CLPM and the CLPM2 suggested a unidirectional relationship between inflectional morphology and expressive vocabulary whereby earlier inflectional morphology affected later expressive vocabulary. Additionally, lagged effects showed stability in both morphology and vocabulary. These results did not change substantially after covariate control (see Tables S5 and S6 in the online supplement). In contrast to these models, and as Table 5 depicts, no cross-lagged effects were found between inflectional morphology and expressive vocabulary in the RI-CLPM. A lagged effect in this model was only found for inflectional morphology between T2 and T3. Further, the RI-CLPM showed that both inflectional morphology and expressive vocabulary were characterized by a high proportion of stable between-person variation—as reflected in the standardized factor loadings of the random intercept of inflections (.69, .72, and .73) and the random intercept of expressive vocabulary (.60, .66, and .73)—and this stable between-person variance was highly correlated (r = .71; see Table 5). While the CLPM exhibited a less adequate fit with the data (χ^2 = 28.460 (8), p < .001, CFI = .922, SRMR = .068, RMSEA = .103 [90% CI = .064–.145]), both the CLMP2 (χ^2 = 8.865 (4), p = .065, CFI = .982, SRMR = .041, RMSEA = .071 [90% CI = .000 – .135]) and the RI-CLPM (χ^2 = 6.921 (4), p = .140, CFI = .989, SRMR = .027, RMSEA = .055 [90% CI = .000–.122]) fit the data very well.

Inflectional morphology and receptive vocabulary

As Table 3 shows, the CLPM suggested a bidirectional cross-lagged relationship between inflectional morphology and receptive vocabulary. In contrast, Tables 4 and 5 reveal that neither the CLPM2 nor the RI-CLPM suggested significant cross-lagged relationships. However, lagged effects were observed in the CLPM and CLPM2, which indicated stability in the constructs over time. We obtained similar results after controlling for covariates in the CLPM and CLPM2 (see Tables S5 and S6 in the online supplement). However, the RI-CLPM revealed that the stability occurred primarily in the between-person variance, where the random intercepts had strong standardized factor loadings for inflectional morphology (.71, .76, and .76) and receptive vocabulary (.68, .69, and .67), rather than in the constructs' within-person variance. Further, the correlation between the stable parts of morphology and vocabulary (i.e., the random intercepts) was high (r = .80). Among these models, the CLPM fit the data less well (χ^2 = 17.224 (7), p = .016, CFI = .973, SRMR = .080, RMSEA = .080 [90% CI = .031–.125]) than the CLPM2 (χ^2 = 7.597 (4), p = .107, CFI = .990, SRMR = .034, RMSEA = .061 [90% CI = .000–.127]) and the RI-CLPM (χ^2 = 1.963 (4), p = .743, CFI = 1.00,

https://doi.org/10.1017/S0142716424000213 Published online by Cambridge University Press

Table 3. Cross-lagged panel model (CLPM) with standardized parameter coefficients

Model	Inflectional morphology-expressive vocabulary			Inflectional morphology-receptive vocabulary			Derivational morphology-expressive vocabulary			Derivational morphology-receptive vocabulary		
Lagged effects	T1→T2	T2→T3	—	T1→T2	T2→T3	—	T1→T2	T2→T3	—	T1→T2	T2→T3	—
MA	.65***	.78***	—	.55***[†]	.79***[†]	—	.54***	.50***	—	.57***	.52***	—
V	.33**	.31***	—	.44***	.46***	—	.39***	.37***	—	.46***	.45***	—
Cross-lagged effects	T1→T2	T2→T3	—	T1→T2	T2→T3	—	T1→T2	T2→T3	—	T1→T2	T2→T3	—
MA → V	.28***	.32***	—	.16**	.15**	—	.16**	.17*	—	.12	.11	—
V → MA	.05	.05	—	.10*	.11*	—	.13*	.11*	—	.04	.04	—
Same-wave correlations	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
MA ↔ V	.30***	.28***	−.02	.45***	.05	.07	.19*	.18*	27*	.40***	.30**	.07

Note. MA = morphological awareness; V = vocabulary; T1 = Preschool; T2 = Grade 1; T3 = Grade 2; *** = $p < .001$; ** = $p < .010$; * = $p < .05$; all lagged and cross-lagged effects were restricted to be equal over time, except when stated otherwise; [†] freely estimated effect without equality constraints.

Table 4. Cross-lagged panel model with second-order effects (CLPM2) and standardized parameter coefficients

Model	Inflectional morphology-expressive vocabulary			Inflectional morphology-receptive vocabulary			Derivational morphology-expressive vocabulary			Derivational morphology-receptive vocabulary		
Lagged effects	T1→T2	T2→T3	T1→T3	T1→T2	T2→T3	T1→T3	T1→T2	T2→T3	T1→T3	T1→T2	T2→T3	T1→T3
MA	.62***	.69***	.17*†	.60***	.66***	.18*†	.47***	.42***	.16†	.50***	.43***	.16†
V	.27**	.24**	.29**†	.42***	.39***	.17*†	.34**	.30***	.28**†	.43***	.38***	.22**†
Cross-lagged effects	T1→T2	T2→T3	T1→T3	T1→T2	T2→T3	T1→T3	T1→T2	T2→T3	T1→T3	T1→T2	T2→T3	T1→T3
MA → V	.27***	.29**	.03†	.11	.10	.13†	.11	.11	.09†	.07	.06	.07†
V → MA	.01	.01	.05†	.08	.08	.06†	.10	.08	.09†	.03	.02	.02†
Same-wave correlations	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
MA ↔ V	.31***	.29***	−.07	.46***	.07	.06	.19**	.21**	.25	.40***	.33***	.07

Note. MA = morphological awareness; V = vocabulary; T1 = Preschool; T2 = Grade 1; T3 = Grade 2; *** = $p < .001$; ** = $p < .010$; * = $p < .05$; all lagged and cross-lagged effects were restricted to be equal over time, except when stated otherwise; † freely estimated effect without equality constraints.

Table 5. Random intercept cross-lagged panel model (RI-CLPM) with standardized parameter coefficients

Model	Inflectional morphology-expressive vocabulary			Inflectional morphology-receptive vocabulary			Derivational morphology-expressive vocabulary			Derivational morphology-receptive vocabulary		
Between-person												
Factor loadings[a]	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
RI–MA	.69***	.72***	73***	.71***	.76***	.76***	.70***	.69***	.67***	.73***	.71***	.68***
RI–V	.60***	.66***	.73***	.68***	.69***	.67***	.60***	.66***	.71***	.66***	.68***	.67***
RI correlation												
MA ↔ V		.71***			.80***			.61***			.56***	
Lagged effects	T1→T2		T2→T3	T1→T2		T2→T3	T1→T2		T2→T3	T1→T2		T2→T3
MA	.18[†]		.66**[†]	.06[†]		.56***[†]	.05		.05	.03		.03
V	−.03		−.03	.06		.06	−.03		−.03	.07		.07
Within-person												
Cross-lagged effects	T1→T2		T2→T3	T1→T2		T2→T3	T1→T2		T2→T3	T1→T2		T2→T3
MA → V	.09		.10	−.18		−.14	−.08		−.10	.01		.01
V → MA	−.14		−.13	−.07		−.07	−.09		−.07	−.05		−.05
Same-wave correlations	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
MA ↔ V	.03	.17	−.22	.14	−.33	−.02	−.11	.05	.21	.22	.29	−.03

Note. RI = random intercept; MA = morphological awareness; V = vocabulary; T1 = Preschool; T2 = Grade 1; T3 = Grade 2; *** = $p < .001$; ** = $p < .010$; * = $p < .05$; all lagged and cross-lagged effects were restricted to be equal over time, except when stated otherwise; [†] freely estimated effect without equality constraints; [a] in order to estimate random intercepts, all unstandardized factor loadings were fixed to 1.

SRMR = .024, RMSEA = .000 [90% CI = .000–.069]), which had an excellent fit to the data.

Derivational morphology and expressive vocabulary

As Table 3 demonstrates, the CLPM suggested a bidirectional relationship between derivational morphology and expressive vocabulary. However, this relationship was no longer significant after we controlled for the measured covariates (see Table S5 in the online supplement). Further, Tables 4 and 5 show that these cross-lagged effects were not present in the two models with more robust confounding control (i.e., the CLPM2 and the RI-CLPM). The lagged effects were significant in the CLPM and the CLPM2 but not the RI-CLPM. Further, the standardized factor loadings for the random intercepts of both derivational morphology (.70, 69, and .67) and expressive vocabulary (60, .66, and .71) were strong in the RI-CLPM. This finding reflected that a large part of the variation in derivational morphology and expressive vocabulary was stable between-person variance that did not change over time. The correlation between these stable parts of derivational morphology and expressive vocabulary was .61.

While the CLPM did not have a good fit to the data ($\chi^2 = 21.156$ (8), $p = .007$, CFI = .893, SRMR = .057, RMSEA = .082 [90% CI = .040–.126]), both the CLPM2 ($\chi^2 = 6.734$ (4), $p = .151$, CFI = .978, SRMR = .045, RMSEA = .053 [90% CI = .000–.121]) and the RI-CLPM ($\chi^2 = 5.177$ (5), $p = .395$, CFI = 1.00, SRMR = .032, RMSEA = .012 [90% CI = .000–.091]) fit the data very well.

Derivational morphology and receptive vocabulary

Between derivational morphology and receptive vocabulary, no cross-lagged relationships were found in any of the models (see Tables 3, 4, and 5 and Tables S5 and S6 in the online supplement). Again, the lagged effects were significant in the CLPM and the CLPM2 but not the RI-CLPM. Further, the random intercepts had strong standardized factor loadings for both derivational morphology (.73, .71, and .68) and receptive vocabulary (.66, .68, and .67) in the RI-CLPM. This finding shows that a large part of the variation in derivational morphology and receptive vocabulary is stable between-person variance that does not change over time. The correlation between these stable parts of derivational morphology and receptive vocabulary was .56. Finally, model fit statistics demonstrated that the CLPM fit the data less adequately ($\chi^2 = 22.139$ (8), $p = .005$, CFI = .935, SRMR = .072, RMSEA = .085 [90% CI = .044 – .129]) than the CLPM2 ($\chi^2 = 8.733$ (4), $p = .068$, CFI = .978, SRMR = .049, RMSEA = .070 [90% CI = .000–.134]) and the RI-CLPM ($\chi^2 = 5.300$ (5), $p = .380$, CFI = .999, SRMR = .048, RMSEA = .016 [90% CI = .000–.092]), which fit the data very well.

Discussion

The present study examined the cross-lagged relationship between morphological awareness and vocabulary. Its results show that the more covariate control the models embed, the less evidence they provide for cross-lagged relations between morphological awareness and vocabulary skills. In contrast to the two models with limited control for unmeasured covariates, the model intended to control for all unmeasured time-invariant covariates (the RI-CLPM) showed no cross-lagged effects between the two processes. In addition, approximately half of the variance in both morphological awareness and vocabulary (varying from 44% to 58%) could be characterized as stable trait-like variance that correlated highly across the two domains.

What influences what?

The proposed CLPM2 shows that controlling for all previous time exposure explains away the significant cross-lagged effects seen in the CLPM, whether controlling for measured covariates or not. One exception was the significant effects of inflectional morphology on expressive vocabulary between preschool and Grade 2 that were found with the CLPM2. Thus, this positive finding suggests that awareness of early inflectional morphology—not just knowledge of specific inflected words—can be longitudinally associated with vocabulary, as other studies have also suggested (Sparks & Deacon, 2013). Nevertheless, and despite our attempts to control for vocabulary effects in the inflectional task (i.e., by using pseudowords), we caution against considering the effect of inflectional morphology on expressive vocabulary as causal because the RL-CLPM did not confirm this effect. Instead, the RI-CLPM suggests that the relationship between morphological awareness and vocabulary is due to stable aspects that do not change over time between children. Our preferred interpretation propounds that the relatively high shared stability between morphological awareness and vocabulary must be ascribed to persistent genetic or environmental factors from preschool to Grade 2.

In this regard, studies have suggested a genetic influence on morphological awareness (Xie et al., 2022) and vocabulary (Hart et al., 2009; Olson et al., 2011), and this influence can vary as a function of environmental factors that may be specific to the context of each study. However, these studies have also suggested certain stability in gene–environment influences over time, which may be partially responsible for the stable differences between children (e.g., Brinchmann et al., 2018). Given this suggestion, the strong associations observed between morphological awareness and vocabulary (Kieffer & Lesaux, 2012; Lee et al., 2022) may plausibly reflect unmeasured stable aspects that, when controlled for (as in our RI-CLPMs), reveal no signs of cross-lagged relationships.

It is important to mention that the longitudinal stability between the children in morphological awareness and vocabulary does not mean that children's language skills do not change over time. While children's vocabulary and awareness of morphology continue to develop over time, they appear to do so at a similar pace, as reflected in the strong correlations between the random intercepts. This is to say that a child's relative position within a group—whether scoring higher or lower than another child—remains relatively stable across time although the child improves their score over time. However, being good at morphological awareness at a particular time point is not associated with being good at vocabulary skills, or vice versa, a year later when we control for this stability between the children.

Limitations and concluding remarks

Our study highlights the importance of controlling for between-person stability to better approximate real causal relations. However, our results also face some limitations. For instance, our raw score correlations and reliabilities for derivational morphology were not as strong as expected. In turn, this finding could reflect that awareness of derivations may still be underdeveloped among Norwegian-speaking students by the middle of Grade 2, as it is among speakers of other languages (Berninger et al., 2010; Carlisle, 2000; Diamanti et al., 2018). Thus, our derivational task may have captured less consistent performance than the inflectional morphology task. Our findings show no relation between derivational morphology and vocabulary at this stage (The only exception was found in the CLPM model with the least covariate control.) Still, some experimental evidence suggests that learning derivations have a positive effect on vocabulary and literacy skills during later stages of development (Torkildsen et al., 2021), warranting further study.

Our study sheds light on the relationship between morphological awareness and vocabulary before the onset of reading instruction and during the first two years of formal schooling. Our results reveal that common factors that are stable over time drive the development of children's language skills. Moreover, our analyses confirm earlier findings on this proposed relation (e.g., the CLPM) and show that early awareness of inflectional morphology remains significantly associated with later expressive vocabulary—even after the more robust covariate control found in the CLPM2, which has not received as much attention. However, the RI-CLPM suggests that this association is explained by stable common factors for both language skills that future research should address when assessing their cross-lagged relations.

Replication package. Data and analysis scripts and outputs (Mplus and R files) are available at https://osf.io/wz6pv/.

Supplementary material. To view supplementary material for this article, please visit https://doi.org/10.1017/S0142716424000213

Acknowledgements. Funded by The Research Council of Norway, Grant 283586, and from The Research Council of Norway, Centres of Excellence, Grant 331640.

References

Anglin, J. (1993). Vocabulary development: A morphological analysis. *In Monographs of the society for research in child development* (Vol. **58**, pp. 1–166).

Bailey, D. H., Oh, Y., Farkas, G., Morgan, P., & Hillemeier, M. (2020). Reciprocal effects of reading and mathematics? Beyond the cross-lagged panel model. *Developmental Psychology*, **56**(5), 912–921. https://doi.org/10.1037/dev0000902

Berninger, V. W., Abbott, R. D., Naggy, W., & Carlisle, J. F. (2010). Growth in phonological, orthographic, and morphological awareness in grades 1 to 6. *Journal of Psycholinguistic Research*, **39**(2), 141–163. https://doi.org/10.1007/s10936-009-9130-6

Brinchmann, E. I., Braeken, J., & Lyster, S. A. H. (2018). Is there a direct relation between the development of vocabulary and grammar? *Developmental Science*, **22**(1). https://doi.org/10.1111/desc.12709

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford Publications.

Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189–209). Lawrence Erlbaum Associates.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, **12**(3–4), 169–190. https://doi.org/10.1023/A:1008131926604

Carlisle, J. F. (2007). Fostering morphological processing, vocabulary development, and reading comprehension. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension.* Guilford Press.

Cole, D. A., & Preacher, K. J. (2014). Manifest variable path analysis: Potentially serious and misleading consequences due to uncorrected measurement error. *Psychological Methods*, **19**(2), 300–315. https://doi.org/10.1037/a0033805

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, **82**(5), 879–894. https://doi.org/10.1037/a0035297

Diamanti, V., Benaki, A., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., & Protopapas, A. (2018). Development of early morphological awareness in Greek: Epilinguistic versus metalinguistic and inflectional versus derivational awareness. *Applied Psycholinguistics*, **39**(3), 545–567. https://doi.org/10.1017/S0142716417000522

Diamanti, V., Mouzaki, A., Ralli, A., Antoniou, F., Papaioannou, S., & Protopapas, A. (2017). Preschool phonological and morphological awareness as longitudinal predictors of early reading and spelling development in Greek. *Frontiers in Psychology*, **8**, 2039. https://doi.org/10.3389/fpsyg.2017.02039

Dulay, K. M., Law, S. Y., McBride, C., & Ho, C. S.-H. (2021). Reciprocal effects of morphological awareness, vocabulary knowledge, and word reading: A cross-lagged panel analysis in Chinese. *Journal of Experimental Child Psychology*, **206**, 105100. https://doi.org/10.1016/j.jecp.2021.105100

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *British picture vocabulary scale* (2nd ed.). NFER-Nelson.

Falkenström, F., Solomonov, N. & Rubel, J. (2023). To detrend, or not to detrend, that is the question? The effects of detrending on cross-lagged effects in panel models. *Psychological Methods.* Advance online publication. https://doi.org/10.1037/met0000632

Hamaker, E. L., Kuiper, R. M., Grasman, R. P. P. P., & Harlow, L. L. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, **20**(1), 102–116. https://doi.org/10.1037/a0038889

Harrison, E., Drake, T., & Ots, R. (2023). finalfit: Quickly create elegant reggression results tables and plots when modelling. R Package version 1.0.1. Retrieved from: https://github.com/ewenharrison/finalfit

Hart, S. A., Petrill, S. A., DeThorne, L. S., Deater-Deckard, K., Thompson, L. A., Schatschneider, C., & Cutting, L. E. (2009). Environmental influences on the longitudinal covariance of expressive vocabulary: Measuring the home literacy environment in a genetically sensitive design. *Journal of Child Psychology and Psychiatry*, **50**(8), 911–919. https://doi.org/10.1111/j.1469-7610.2009.02074.x

Kieffer, M. J., & Lesaux, N. K. (2012). Development of morphological awareness and vocabulary knowledge in Spanish-speaking language minority learners: A parallel process latent growth curve model. *Applied Psycholinguistics*, **33**(1), 23–54. https://doi.org/10.1017/S0142716411000099

Kuiper, R. M. (2018). Chi-bar-square difference test [Shiny app software]. Retrieved from: https://www.uu.nl/staff/RMKuiper/Websites%20%2F%20Shiny%20apps

Kuo, L. J., & 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

Lee, J. W., Wolters, A., & Grace Kim, Y.-S. (2022). The relations of morphological awareness with language and literacy skills vary depending on orthographic depth and nature of morphological awareness. *Review of Educational Research*, **93**(4), 528–558. https://doi.org/10.3102/00346543221123816

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

Lucas, R. E. (2023). Why the cross-lagged panel model is almost never the right choice. *Advances in Methods and Practices in Psychological Science*, **6**(1). https://doi.org/10.1177/25152459231158378

Lüdtke, O., & Robitzsch, A. (2022). A comparison of different approaches for estimating cross-lagged effects from a causal inference perspective. *Structural Equation Modeling: A Multidisciplinary Journal*, **29**(6), 888–907. https://doi.org/10.1080/10705511.2022.2065278

Lyster, S.-A. H., Horn, E., & Rygvold, A.-L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av british picture vocabulary scale (BPVS II). *Spesialpedagogikk*, 74, 35–74. Retrieved from: https://www.utdanningsnytt.no/files/2019/08/21/Spesialpedagogikk%209%202010.pdf

McBride-Chang, C., Tardif, T., Cho, J.-R., Shu, H. U. A., Fletcher, P., Stokes, S. F., Wong, A., & Leung, K. (2008). What's in a word? Morphological awareness and vocabulary knowledge in three languages. *Applied Psycholinguistics*, **29**(3), 437–462. https://doi.org/10.1017/s014271640808020x

Mulder, J. D., & Hamaker, E. L. (2021). Three extensions of the random intercept cross-lagged panel model. *Structural Equation Modeling: A Multidisciplinary Journal*, **28**(4), 638–648. https://doi.org/10.1080/10705511.2020.1784738

Muthén, B. O. (1991). Multilevel factor analysis of class and student achievement components. *Journal of Educational Measurement*, **28**(4), 338–354.

Muthen, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, **25**, 267–316.

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly*, **19**(3), 304–330. https://doi.org/10.2307/747823

Norwegian Agency for Shared Services in Education and Research (2022). About Sikt- Norwegian Agency for Shared Services in Education and Research. Retrieved May 25 from About Sikt – Norwegian Agency for Shared Services in Education and Research.

Olson, R. K., Keenan, J. M., Byrne, B., Samuelsson, S., Coventry, W. L., Corley, R., Wadsworth, S. J., Willcutt, E. G., DeFries, J. C., Pennington, B. F., & Hulslander, J. (2011). Genetic and environmental influences on vocabulary and reading development. *Scientific Studies of Reading*, **15**(1), 26–46. https://doi.org/10.1080/10888438.2011.536128

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, **11**(4), 357–383. https://doi.org/10.1080/10888430701530730

Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, **48**(1), 85–112. https://doi.org/10.1016/j.jsp.2009.09.002

Pickering, S. J., & Gathercole, S. E. (2001). *Working memory test battery for children*. Psychological Corporation UK.

Ragnarsdóttir, H., Simonsen, H. G., & Plunket, K. (1999). Acquisition of past tense morphology in Icelandic and Norwegian children: An experimental study. *Journal of Child Language*, **26**(01), 577–618.

Revelle, W. (2022). psych: Procedures for Psychological, Psychometric, and Personality Research. Northwestern University, Evanston, Illinois. R package version 2.2.12, https://CRAN.R-project.org/package = psych.

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika*, **75**(2), 243.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, **94**(2), 143–174. https://doi.org/10.1348/000712603321661859

Sparks, E., & Deacon, S. H. (2013). Morphological awareness and vocabulary acquisition: A longitudinal examination of their relationship in English-speaking children. *Applied Psycholinguistics*, **36**(2), 299–321. https://doi.org/10.1017/s0142716413000246

Statistics Norway. (2021a). Educational attainment of the population. Retrieved from: https://www.ssb.no/en/utdanning/statistikker/utniv

Statistics Norway. (2021b). Number of employments and earnings. Retrieved from: https://www.ssb.no/statbank/table/11652/

Stoel, R. D., Garre, F. G., Dolan, C., & van den Wittenboer, G. (2006). On the likelihood ratio test in structural equation modeling when parameters are subject to boundary constraints. *Psychological Methods*, **11**, 439–455. https://doi.org/10.1037/1082-989X.11.4.439

Torkildsen, J. v. K., Bratlie, S. S., Kristensen, J. K., Gustafsson, J.-E., Lyster, S.-A. H., Snow, C., Hulme, C., Mononen, R.-M., Naess, K.-A. B., Lopez-Pedersen, A., Wie, O. B., & Hagtvet, B. (2021). App-based morphological training produces lasting effects on word knowledge in primary school children: A randomized controlled trial. *Journal of Educational Psychology*, **114**(4), 833. https://doi.org/10.1037/edu0000688

Usami, S., Murayama, K., & Hamaker, E. L. (2019). A unified framework of longitudinal models to examine reciprocal relations. *Psychological Methods*, **24**(5), 637–657. https://doi.org/10.1037/met0000210

VanderWeele, T. J. (2021). Causal inference with time-varying exposures. In T. L. Lash, T. J. VanderWeele, S. Haneuse, & K. J. Rothman (Eds.), Modern epidemiology (pp. 605–618). Wolters Kluwer.

VanderWeele, T. J., Mathur, M. B., & Chen, Y. (2020). Outcome-wide longitudinal designs for causal inference: A new template for empirical studies. *Statistical Science*, **35**(3), 437–466. https://doi.org/10.1214/19-STS728

Wechsler, D. (2003). *WISC-IV technical and interpretive manual.* Psychological Corporation.

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence—fourth edition*. Bloomington: Pearson.

Xie, Q., Zheng, M., Ho, C. S.-H., McBride, C., Fong, F. L. W., Wong, S. W. L., & Chow, B. W.-Y. (2022). Exploring the genetic and environmental etiologies of phonological awareness, morphological awareness, and vocabulary among Chinese–English bilingual children: The moderating role of second language instruction. *Behavior Genetics*, **52**(2), 108–122. https://doi.org/10.1007/s10519-021-10096-2

Cite this article: Grande, G., Diamanti, V., Protopapas, A., Melby-Lervåg, M., & Lervåg, A. (2024). The development of morphological awareness and vocabulary: What influences what? *Applied Psycholinguistics* **45**, 745–765. https://doi.org/10.1017/S0142716424000213