Morphological Awareness and Reading Comprehension: Are All Tasks The Same?

Émilie Cloutier, Prof., Université du Québec à Montréal

Robert Savage, Dean, York University

Pierre-Olivier Caron, Prof., Université de la TÉLUQ

Catherine Turcotte, Prof., Université du Québec à Montréal

Presented at the SSSR 33rd annual meeting

July, 2026



Presentation

- Introduction
- Theoretical background
- Method
- Results
- Discussion
- Conclusion

Introduction





Morphological awareness (MA) is an important predictor of literacy, as it has been associated with word reading, spelling, and reading comprehension (e.g., Deacon et al., 2015; Kieffer & Box, 2013; Tong et al., 2013 vs. Goodwin et al., 2012; Jarmulowicz et al., 2007; Proctor et al., 2011).



However, its unique contribution to reading comprehension remains inconsistent across studies.



One possible explanation for these inconsistent findings is the way MA is assessed.



Researchers use a wide variety of tasks, and there is no clear consensus about what constitutes morphological awareness or how it should be measured.

Morphological awareness is a multidimensional construct (Kuo & Anderson, 2006).

Because morphemes carry phonological, orthographic, semantic, and grammatical information, different tasks may measure different aspects of MA and engage different cognitive processes.



Task characteristics (e.g., oral vs. written responses, contextualized vs. decontextualized items, production vs. judgment tasks) may influence the cognitive and linguistic demands of MA measures, meaning that different tasks may not assess the same underlying abilities (Apel et al., 2013; Berninger & Winn, 2006; Liu et al., 2024).



These differences in task demands may explain why MA measures vary in their ability to predict reading comprehension.



Tasks that engage semantic and syntactic processing more closely aligned with comprehension may show stronger relations with reading comprehension than tasks relying primarily on orthographic or transcription processes.

Theoretical background



- MA involves reflecting on and manipulating morphemes, which convey semantic, phonological, orthographic, and grammatical information (Carlisle, 1995; Kuo & Anderson, 2006).

- The complexity of the relation between two morphologically related words depends on word characteristics.

(e.g. Factors such as root frequency, semantic transparency, and phonological or orthographic changes between base and derived words, Carlisle, 2003; Goodwin & Ahn, 2010; Berninger et al., 2010; Logan, 2010).

Types of task

Production tasks:	• In these tasks, participants must produce a word either by deriving it or by 'extracting the base usually' presented in a cloze sentence (e.g., Carlisle, 2000; Goodwin et al., 2012). (*Farm. My uncle is a_____ (farmer)*. e.g., Carlisle, 2000). • In an extract the base' task, sometimes referred to as a decomposition task, participants perform the opposite. Using a multimorphemic word, the extract the base word (*Farmer. My uncle works on a ____. (farm)*. e.g., Carlisle, 2000; Goodwin et al., 2012).
Judgement tasks:	• In studies using judgement tasks, participants are asked to judge if two words come from the same semantic family (e.g., *Does farm come from farmer?*). The response in these types of tasks is usually a yes or no answer. The tasks can use real words or pseudo-words.
Analogy tasks:	• In an analogy task, items are presented in an A:B, C:D format. (e.g. *Farm/ Farmer : Bake/ _____(Baker)*). These tasks are typically presented orally. Most of the time, no context clues are given, but some studies have used tasks containing sentences to complete the analogy (e.g., Nagy, 2006). These tasks can be performed with real words or pseudo-words.
Identification task:	• In these tasks, participants must identify morphemes in a giving word (e.g. Rinn**ing**, Apel et al., 2013).

Previous Studies



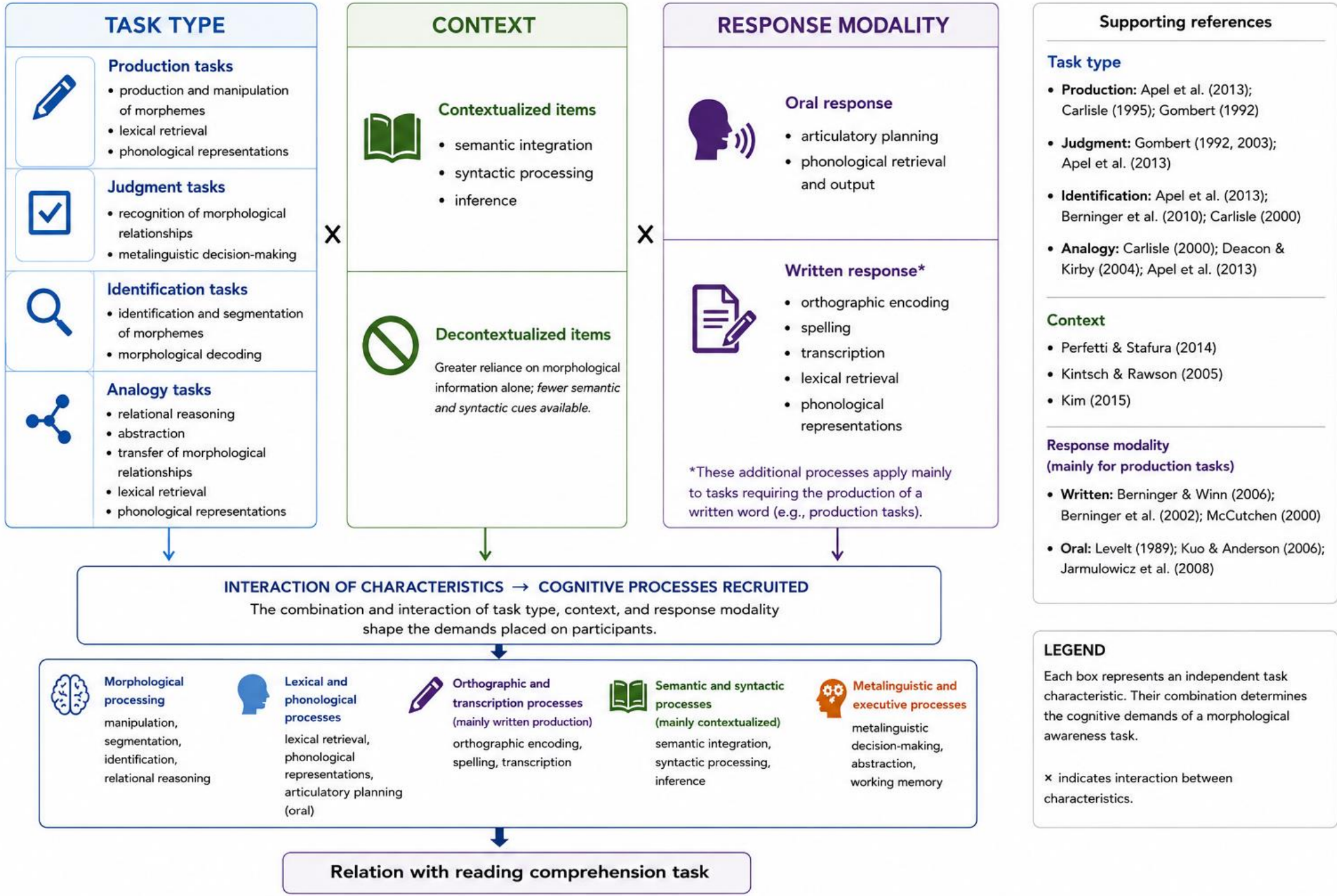
Only two studies have systematically examined task differences.



Apel et al. (2013) found that different task types explained MA differently and that only a production ("relative") task predicted reading comprehension.



Liu et al. (2024) showed that the type of morphemes included in the task (inflectional vs. derivational vs. compound) significantly moderated the relation between MA and reading comprehension, whereas response modality did not.



Supporting references

Task type

- **Production:** Apel et al. (2013); Carlisle (1995); Gombert (1992)
- **Judgment:** Gombert (1992, 2003); Apel et al. (2013)
- **Identification:** Apel et al. (2013); Berninger et al. (2010); Carlisle (2000)
- **Analogy:** Carlisle (2000); Deacon & Kirby (2004); Apel et al. (2013)

Context

- Perfetti & Stafura (2014)
- Kintsch & Rawson (2005)
- Kim (2015)

Response modality (mainly for production tasks)

- **Written:** Berninger & Winn (2006); Berninger et al. (2002); McCutchen (2000)
- **Oral:** Levelt (1989); Kuo & Anderson (2006); Jarmulowicz et al. (2008)

LEGEND

Each box represents an independent task characteristic. Their combination determines the cognitive demands of a morphological awareness task.

× indicates interaction between characteristics.

Reasearch objectives



1) explore the factorial structure within each task using a principal component analytic approach to assess item quality. This is the first preliminary step toward understanding morphological awareness as a construct and how task variations affect its assessment.



2) to examine which of these MA tasks and variations better predicts reading comprehension above and beyond both vocabulary and word reading for fifth graders to identify whether distinct measures of morphology are unique predictors of reading comprehension.

Method



Participants

- This study collected data from 126 Grade 5 students attending six English-language schools (facilitating access to the tasks and assessment measures) in the Greater Montreal area.
- The students' mean age at the time of data collection was 10 years and 3 months.

MA measures

Tasks	Written responses	Oral responses	Contextual cues	Items
Decomposition task *(Carlisle, 2000)*		✓	✓	Practice: a. Driver. Children are too young to _____. [drive] b. Improvement. My teacher wants my spelling to _____. [improve]
Derivation task *(Carlisle, 2000)*		✓	✓	Practice: a. Farm. My uncle is a _____. [farmer] b. Help. My sister is always _____. [helpful]
Extract the Base *(Goodwin et al., 2012)*	✓		✓	1. publicity The _____ was happy with the show. 2. sensitive Animals often _____ storms before they come.
Word analogy *(Deacon et al., 2014)*		✓		luck : lucky :: curl : _____ curly wide : width :: deep : _____ depth

MA measures

- For each task, several additional characteristics were also analyzed:
  - Root frequency, based on the Morpho_lex database
  - Affix frequency, based on the Morpho_lex database
  - Morphological transformation opacity



Table 18. Word characteristics in the Extract the Base

item	prompt word	answer	root frequency	affix frequecy	derivational or flexional	orthographic transformation	phonological transformation
1	*publicity*	*public*	212578	-ity 1647588	derivational	transparent	opaque
2	*sensitive*	*sense*	144057	-ive 1037354	derivational	transparent	transparent
3	*breathe*	*breath*	20693	-(breathe flexion) 28749	flexional	transparent	transparent
4	*musician*	*music*	163742	- ian 394113	derivational	transparent	opaque
5	*runner*	*run*	434649	- er 4569119	derivational	transparent	transparent
6	*fearful*	*fear*	46846	- ful 429561	derivational	transparent	transparent
7	*width*	*wide*	22369	-(wide flexion) 118804	flexional	transparent	transparent
8	*continuous*	*continue*	141133	- ious 450833	derivational	opaque	transparent
9	*bathe*	*bath*	18499	-(bathe flexion) 18499	flexional	transparent	opaque
10	*procedure*	*proceed*	64662	-ure 541479	derivational	opaque	transparent
11	*dangerous*	*danger*	43185	-ious 450833	derivational	transparent	transparent
12	*cleanliness*	*clean*	60401	-ly 3857999 -ness 181553	derivational	transparent	transparent

Other Measures

- **Comprehension**
- *WOODCOCK-JOHNSON PASSAGE COMPREHENSION SUBTEST* (Woodcock *et al*., 2001)
- **Vocabulary**
- *PEABODY PICTURE VOCABULARY TEST* (Dunn et Dunn, 1997)
- **Word identification**
- *WOODCOCK-JOHNSON LETTER-WORD IDENTIFICATION SUBTEST* (Woodcock *et al*., 2001)

Data Collection

	Group	Measures	Assessement time
First meeting	Sub-group (5 students)	*Extract the Base test* (Goodwin *et al.*, 2012) *Peabody Picture Vocabulary test* (Dunn et Dunn, 1997)	25 minutes
Second meeting	Individual	*Woodcock-Johnson Passage Comprehension subtest* (Woodcock *et al.*, 2001) *The Word Analogy task* (Deacon *et al.*, 2014) *Derivation test* (Carlisle, 2000) *Decomposition test* (Carlisle, 2000)	40 minutes

Results



Descriptive statistics

	Mean	SD	Skewness	Kurtosis
Passage Comprehension	92.77	8.34	-0.04	0.11
Vocabulary	98.99	11.77	-0.83	2.75
Letter-Word identification	113.83	12.50	-0.47	0.48
Decomposition	22.90	3.02	-0.96	1.21
Extract the base	23.18	5.50	-1.91	3.33
Derivation	17.84	5.09	-0.19	-0.54
Word Analogy	14.49	3.08	-1.06	2.99

Item difficulty

p values for items 1 to 14 of the **Decomposition task**

Items	1	2	3	4	5	6	7	8	9	10	11	12	13	14
p values	1.00	.98	.56	.60	.73	.95	.65	.86	.95	.90	.94	.98	.99	.99

p values for items 15 to 28 of the **Decomposition task**

Items	15	16	17	18	19	20	21	22	23	24	25	26	27	28
p values	.56	.20	.98	.98	.95	.45	.97	.22	.77	.96	.92	.97	.94	.95

p values for items 1 to 24 in the **Derivation task**

Items	1	2	3	4	5	6	7	8	9	10	12	13	14
p values	.55	.98	.47	.22	.69	.69	.61	.78	.93	.90	.40	.54	.29

p values for items 15 to 28 in the **Derivation task**

Items	15	16	17	18	19	20	21	22	23	24	25	26	27	28
p values	.40	.38	.54	.85	.36	.56	.98	.50	.81	.69	.68	.84	.48	.88

High difficulty ($p \geq .90$) Low difficulty ($p \leq .30$)

Item difficulty

p values for items 1 to 10 in the **Word Analogy task**

Items	1	2	3	4	5	6	7	8	9	10
p values	.86	.85	.60	.69	.58	.87	.87	.80	.84	.89

p values for items 11 to 20 in the **Word Analogy task**

Items	11	12	13	14	15	16	17	18	19	20
p values	.79	.86	.81	.77	**.17**	**.27**	**.90**	.55	.67	.85

p values for items 1 to 14 for the **Extract the Base task**

Items	1	2	3	4	5	6	7	8	10	11	12	13	14
p values	.79	.79	.73	**.95**	**.98**	**.94**	.83	**.91**	**.92**	.65	**.96**	**.97**	.78

p values for items 15 to 28 for the **Extract the Base task**

Items	15	16	17	18	19	20	21	22	23	24	25	26	27	28
p values	**.92**	.77	.53	.65	.82	.87	.83	.82	**.92**	.70	.79	.88	.73	.84

High difficulty ($p \geq .90$) Low difficulty ($p \leq .30$)

1) explore the factorial structure within each task using a principal component analytic approach to assess item quality

- To address this objective, a principal component analyse was conducted on the morphological awareness tasks
- The Minimum Average Partial (MAP) test was used to determine the number of factors to retain, as it is recommended for this purpose (Caron, 2019; Wayne F. Velicer, 1976).
- The MAP test indicated a one component solution for each task. Accordingly, a single component was retained for each task for use in the subsequent analyses.

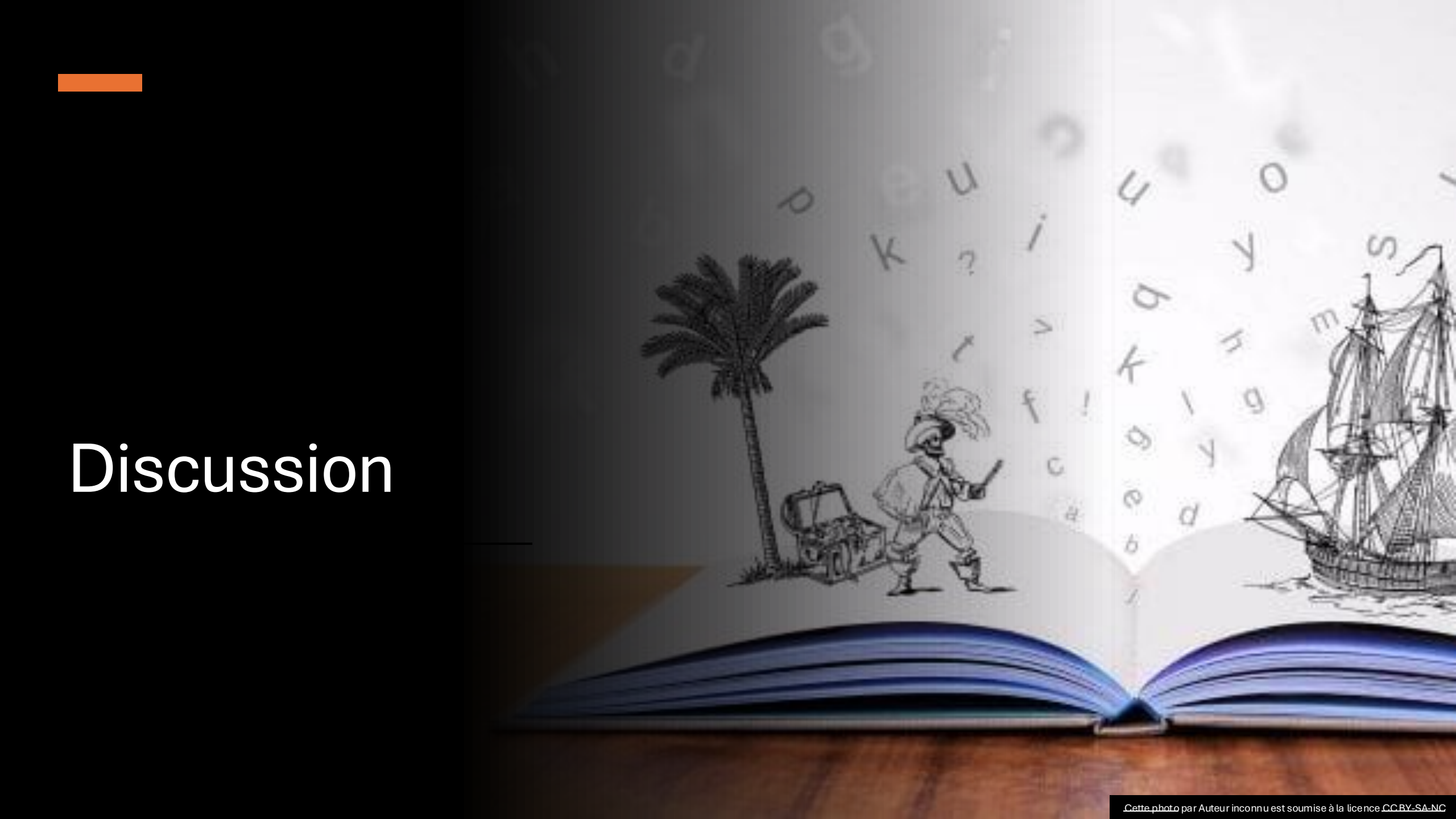
2) To examine which of these **MA tasks and variations** better predicts reading comprehension above and beyond both vocabulary and word reading for fifth graders to identify whether distinct measures of morphology are unique predictors of reading comprehension.

Results of the four linear regression analyses predicting passage reading comprehension while controlling for vocabulary and word reading (N = 126).

COMPONENTS	β	SE	β_{std}	P
1. Decomposition	2.047	.725	.250	**.006**
2. Extract the Base	.969	.831	.113	.246
3. Derivation	1.671	.728	.202	**.024**
4. Word Analogy	.471	.726	.057	.518

Significant result ($p < .05$)

Discussion



Cette photo par Auteur inconnu est soumise à la licence CC BY-SA-NC

Main findings



Morphological awareness (MA) tasks differed substantially in item difficulty and psychometric quality.



All four tasks measured a single underlying construct, but several poorly functioning items were identified and removed.



Only the Derivation and Decomposition tasks uniquely predicted reading comprehension, after controlling for vocabulary and word reading.

Why did some tasks predict reading comprehension?



Oral responses + sentence context were associated with stronger prediction of reading comprehension for our sample.

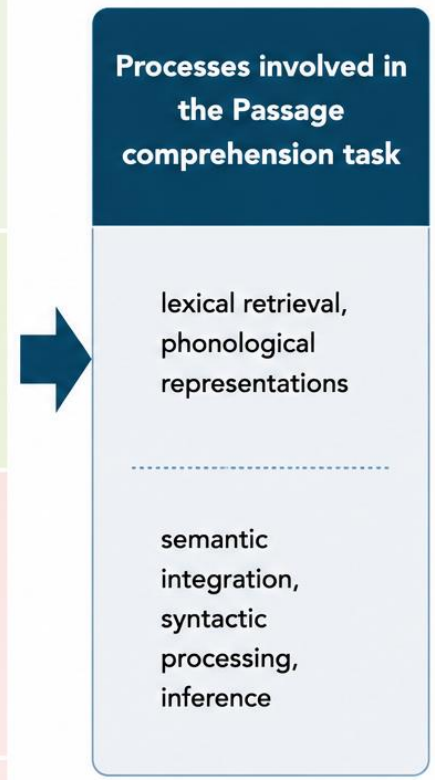


Written-response tasks (e.g., Extract the Base) may recruit additional spelling and transcription processes, reducing their overlap with reading comprehension.



Contextualized tasks likely engage semantic integration, lexical retrieval, and inference generation, processes that closely resemble those involved in reading comprehension.

Task	Written response	Oral response	Contextual clues	Type of task
Decomposition (Carlisle, 2000)	✗	✓ lexical retrieval, phonological representations	✓ semantic integration, syntactic processing, inference	**Production task** **Decomposition**
Derivation (Carlisle, 2000)	✗	✓ lexical retrieval, phonological representations	✓ semantic integration, syntactic processing, inference	**Production task** **Derivation**
Extract the Base (Goodwin et al., 2012)	✓ orthographic encoding, spelling, transcription, lexical retrieval, phonological representations	✗	✓ semantic integration, syntactic processing, inference	**Production task** **Extract the base**
Word analogy (Deacon et al., 2015)	✗	✓ lexical retrieval, phonological representations	✗	**Analogy reasoning**



Psychometric Implications



Item performance was influenced by linguistic characteristics such as:

- Root frequency
- Affix frequency
- Morphological transparency/opacity



However, task performance reflects more than morphological awareness alone, involving lexical, semantic, orthographic, and contextual processes.



Theoretical Implications

- Differences between MA tasks likely reflect differences in the cognitive processes they recruit, not just differences in morphological awareness.
- The contribution of morphological awareness to reading comprehension should therefore be interpreted in light of task characteristics.
- Findings support contemporary reading comprehension models in which successful comprehension depends on the interaction of multiple linguistic and cognitive processes

Practical Implications

- Not all MA tasks are equally informative for assessing reading comprehension. Researchers and practitioners should:
  - choose MA tasks according to their assessment goals,
  - prioritize tasks with stronger psychometric properties,
  - consider how task format (oral vs. written, contextualized vs. isolated) influences performance.

Limitations

- Cross-sectional correlational design: no causal conclusions can be drawn.
- Single reading comprehension measure (passage comprehension only), limiting generalizability.
- Extract the Base task was administered under different testing conditions (small groups vs. individual administration).
- Sample size was insufficient for confirmatory factor analysis (CFA) or structural equation modeling (SEM).
- Not all existing MA task variations were examined.

Conclusion



Take-home message

- Not all morphological awareness tasks measure the construct equally.
- Task design (for MA and RC) matters for both psychometric validity and prediction of reading comprehension.
- Assessment choices influence theoretical conclusions about the role of morphological awareness in reading comprehension.

Thank you!

References