The importance of vocabulary in longitudinal prediction of reading comprehension by word-level skills

Angeliki Mouzaki,[1] Georgios D. Sideridis,[1] Panagiotis G. Simos[1] & Athanassios Protopapas[2]

[1]University of Crete; [2]Institute for Language & Speech Processing, Greece

Background and objectives

Reading comprehension (RC) is thought to be related to (a) word-level reading skills and (b) oral language skills at the level of both word (vocabulary) and text (syntactic and text processing). In the context of the lexical quality hypothesis, measures of lexical skill assess the semantic network substrate necessary for the text comprehension. We have previously suggested that lexical skills, indexed by vocabulary measures, may mediate the effects of word-level reading skills to RC. Therefore, we expect vocabulary measures to be strong longitudinal predictors of RC, taking up RC variance predicted by reading skills.

Participants and measures

This is a longitudinal study of reading development originally including 587 children in Grades 2–4 from 17 schools in Crete and Attika, followed up through Grades 4–6. Data analyzed for this study come from the 2nd (Fall 2005) and 4th assessment waves (Fall 2006). Table below shows participants by grade and gender, excluding children of immigrants and children with combined WISC III Blocks + Vocabulary < 6. Measures included word and pseudoword reading accuracy and fluency, RAN, text-based knowledge of morphosyntax, passage reading efficiency, listening (Token) and reading comprehension (multiple-choice questions), spelling, WISC block design, expressive (WISC) and receptive (PPVT) vocabulary.

	Grade			
	2	3	4	Total
Boys	84	70	68	222
Girls	86	82	81	249
Total	170	152	149	471

Results

RC at Wave 4 was regressed onto Wave 2 measures, grouped into word-level print skills, language, general cognitive ability, vocabulary. Tables show change in proportion of RC4 variance accounted for after entering each variable group (below) and standardized coefficients and semi-partial correlations (right). Independent contributions to predicting RC variance one year later were: Vocabulary 8.3%, morphosyntactic knowledge 2.2%, word & pseudoword reading 1.8%.

W2 variables predicting RC4	Adj. R^2	R^2 chg	F chg	df1,df2	p
Age, gender	.006	.011	2.235	2,392	.108
WISC-III Blocks	.045	.041	17.057	1,391	.000
RAN, listening comprehension	.124	.083	18.605	2,389	.000
Word & pseudoword accuracy & fluency, spelling	.177	.063	5.992	5,384	.000
Morphosyntactic knowledge	.214	.042	7.064	3,381	.000
WISC-III Vocabulary, PPVT	.296	.083	23.265	2,379	.000

W2 variables predicting RC4	Adj. R^2	R^2 chg	F chg	df1,df2	p
Age, gender	.006	.011	2.235	2,392	.108
WISC-III Blocks	.045	.041	17.057	1,391	.000
RAN, listening comprehension	.124	.083	18.605	2,389	.000
WISC-III Vocabulary, PPVT	.256	.133	35.315	2,387	.000
Word & pseudoword accuracy & fluency, spelling	.279	.033	3.559	5,382	.004
Morphosyntactic knowledge	.296	.022	4.094	3,379	.007

W2 variables predicting RC4	Adj. R^2	R^2 chg	F chg	df1,df2	p
Age, gender	.006	.011	2.235	2,392	.108
WISC-III Blocks	.045	.041	17.057	1,391	.000
RAN, listening comprehension	.124	.083	18.605	2,389	.000
WISC-III Vocabulary, PPVT	.256	.133	35.315	2,387	.000
Morphosyntactic knowledge	.287	.036	6.685	3,384	.000
Word & pseudoword accuracy & fluency, spelling	.296	.018	2.036	5,379	.073

	β	t	p	r^*
Age	.03	.65	.51	.03
Gender	.10	2.32	.02	.10
WISC-III Blocks	−.01	−.20	.84	−.01
Listening comprehension	.09	1.89	.06	.08
RAN	−.03	−.56	.58	−.02
WISC-III Vocabulary	.10	1.92	.06	.08
PPVT	.30	5.66	.00	.24
Word reading accuracy	.12	1.94	.05	.08
Pseudoword reading accuracy	−.05	−.86	.39	−.04
Word reading fluency	−.10	−1.26	.21	−.05
Pseudoword reading fluency	.07	.99	.32	.04
Spelling	.10	1.48	.14	.06
Morphosyntax – Nouns	.10	2.00	.05	.08
Morphosyntax – Verbs	.12	2.49	.01	.11
Morphosyntax – Sentences	−.04	−.72	.47	−.03

*Semi-partial correlations with RC4 in the full regression model

Discussion

Vocabulary was by far the most significant longitudinal predictor of RC, followed by word reading accuracy. After partialling out vocabulary, reading measures contributed much less RC variance than what was contributed by vocabulary after partialling out word & pseudoword reading. Most longitudinally reliable RC variance can be attributed to lexical, print-independent skills. These results indicate that what print-dependent lexical knowledge would contribute to RC is largely taken up by variance from oral measures, mainly vocabulary. Thus vocabulary dominates decoding in accounting for RC variance. Similar results from latent variable model analysis of cross-sectional data suggest that vocabulary mediates decoding effects in comprehension (Protopapas, Sideridis, Mouzaki, & Simos, 2007; SSR 11(3): 165–197)

amouzaki@edc.uoc.gr, sideridis@psy.soc.uoc.gr, akis.simos@gmail.com, protopap@ilsp.gr