Learning to read and write in different languages: What's the difference?

Dr. Beth A. O'Brien
Principal Research Scientist
Deputy Director Centre for Research in Child Development
National Institute of Education, Nanyang Technological University, Singaore

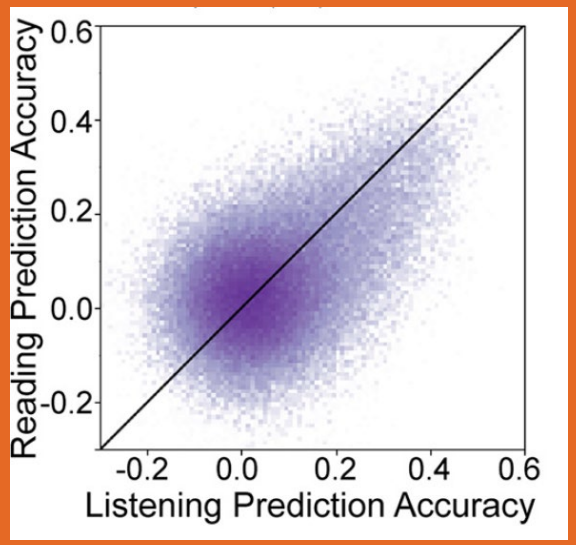
Language

&

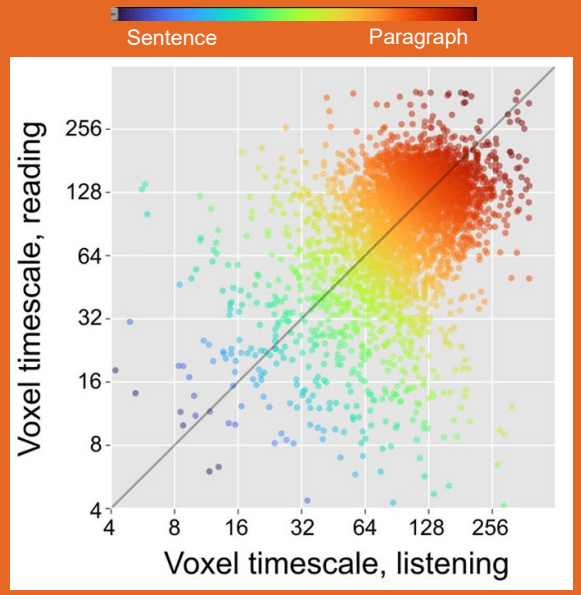
Literacy

Understanding and Communicating

Overlapping brain networks



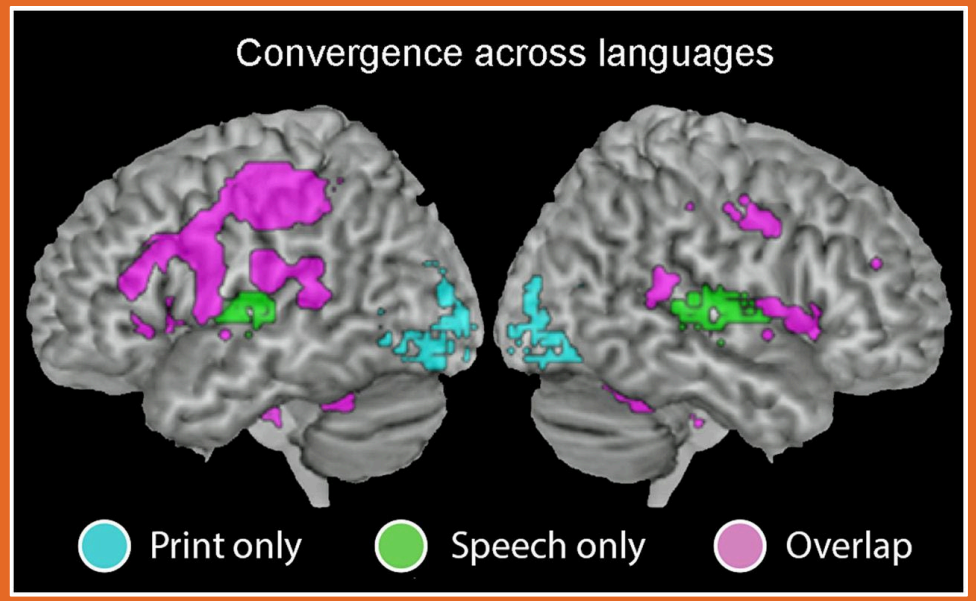
Deniz et al. (2019)



Chen et al. (2024)

Listening to & Reading stories

Semantic categorization task
Listening to or Reading words

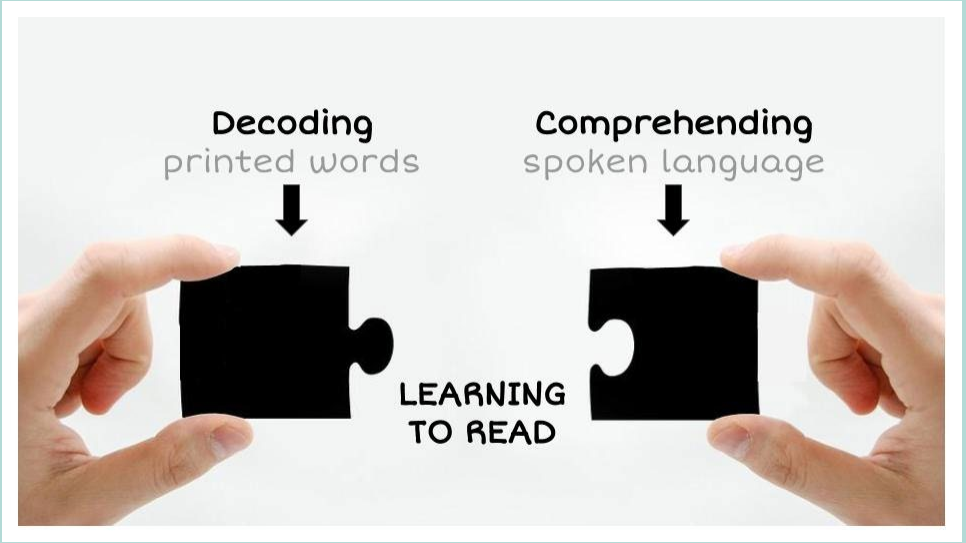


Rueckl et al. (2018) PNAS

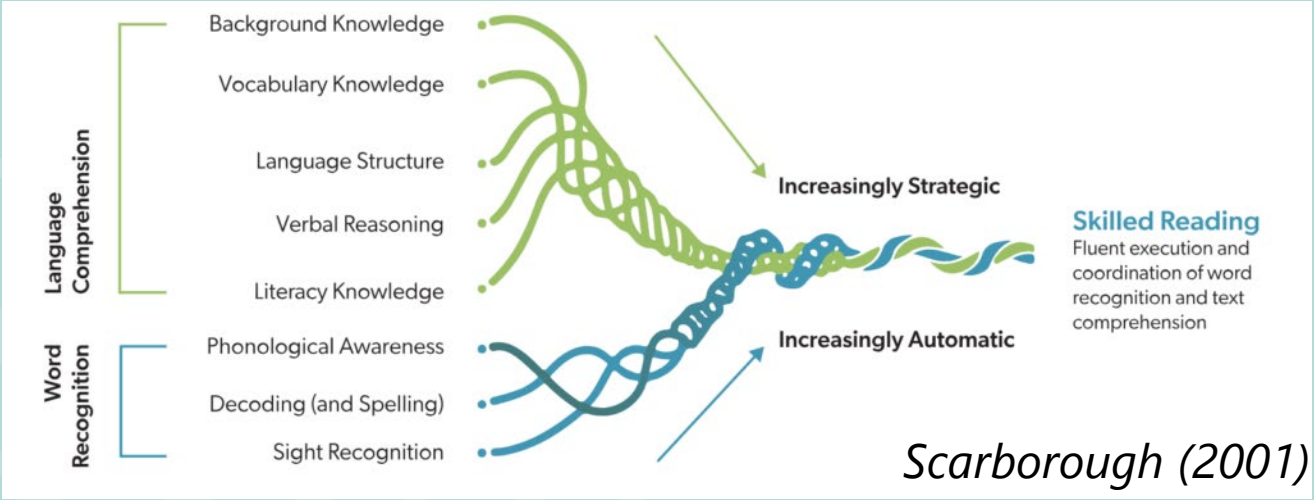
Reading Models

Simple View of Reading

RC = D x LC



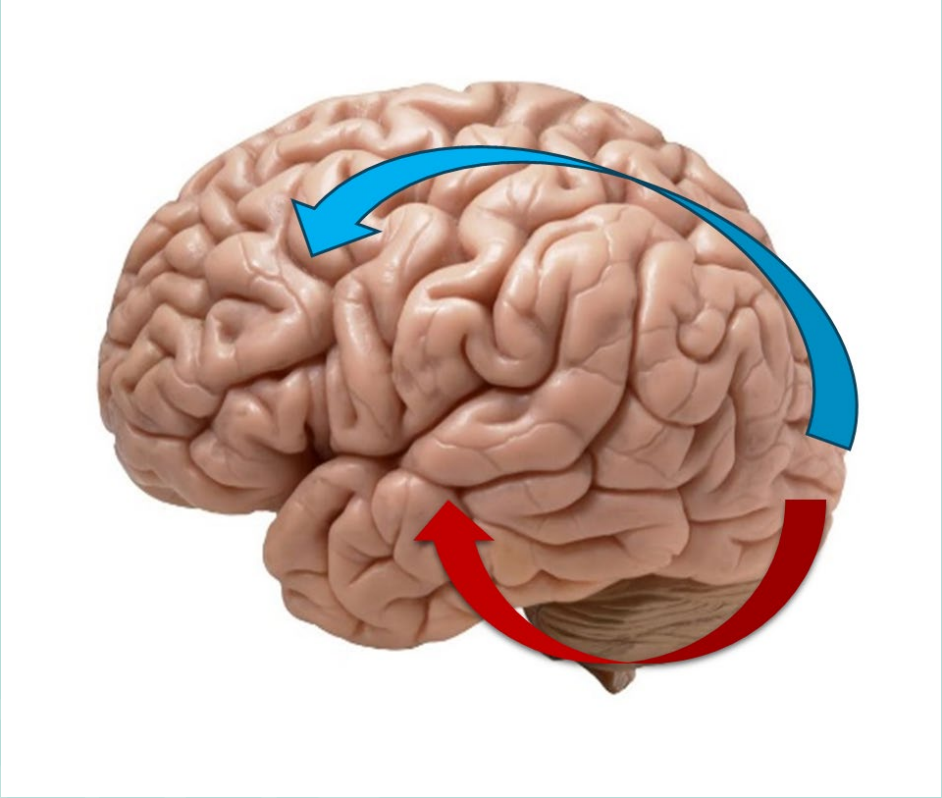
Gough & Tunmer (1986)
Hoover & Gough (1990)



Scarborough (2001)

Reading Models

Dual Route Model of reading words aloud



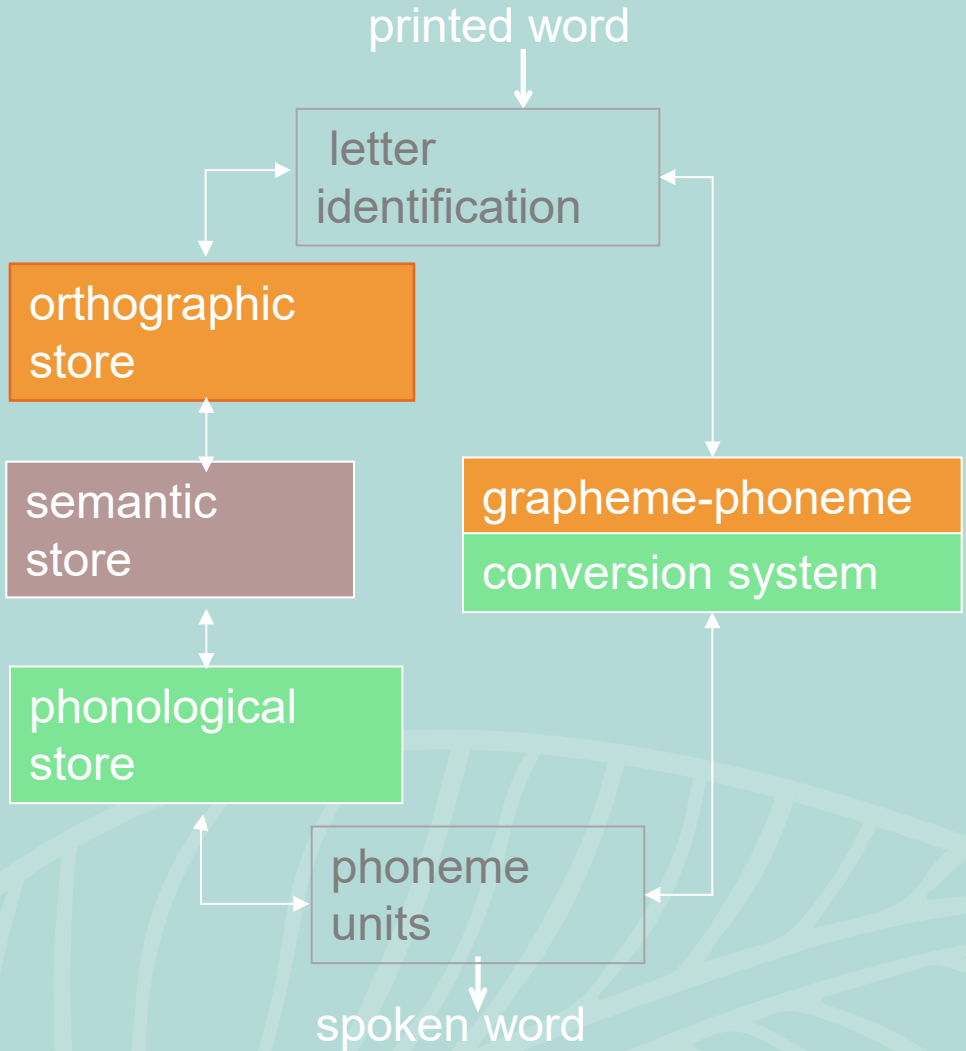
Lexical
Addressed phonology
(ventral)

Sublexical
Assembled phonology
(dorsal)

Coltheart et al. (1993)
Jobard et al. (2003)

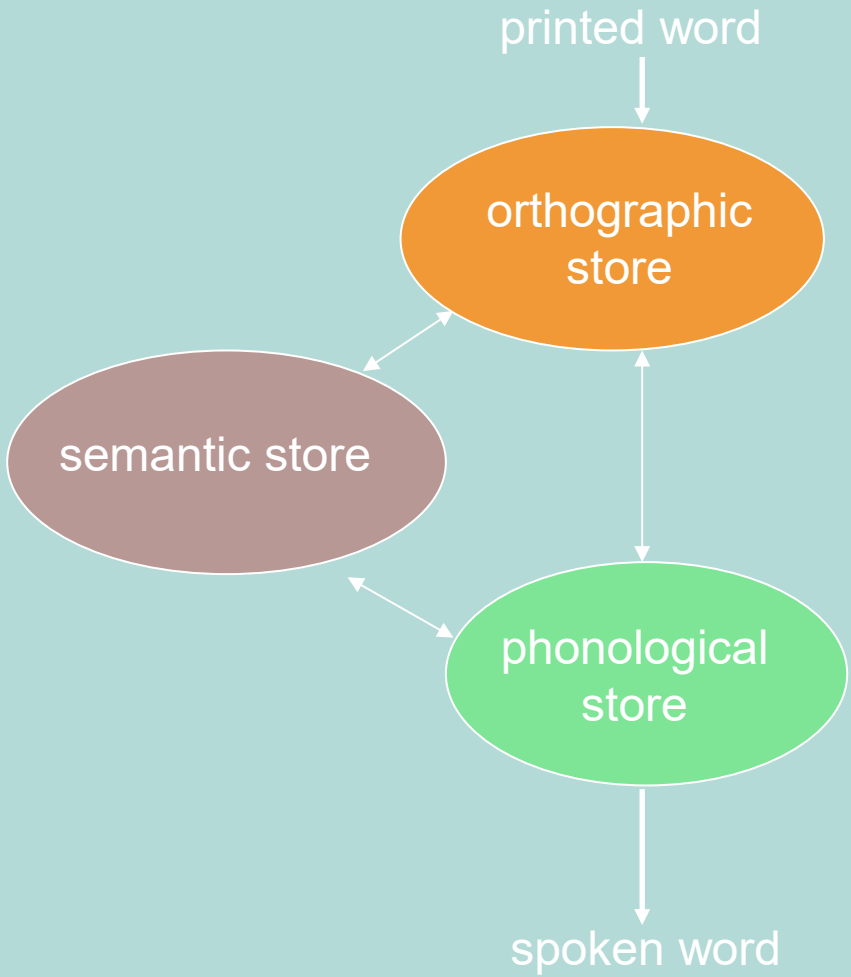
Reading Models

Dual Route Model



Castles & Coltheart (1993)

Triangle Model



Seidenberg and McClelland (1989)
Plaut et al. (1996)

Reading in different scripts

Universal Properties

?

Language Specific Properties

Reading models across languages

Orthographic Depth Hypothesis

Katz & Feldman (1983)

OPAQUE
Inconsistent GPC

TRANSPARENT
1:1 GPC

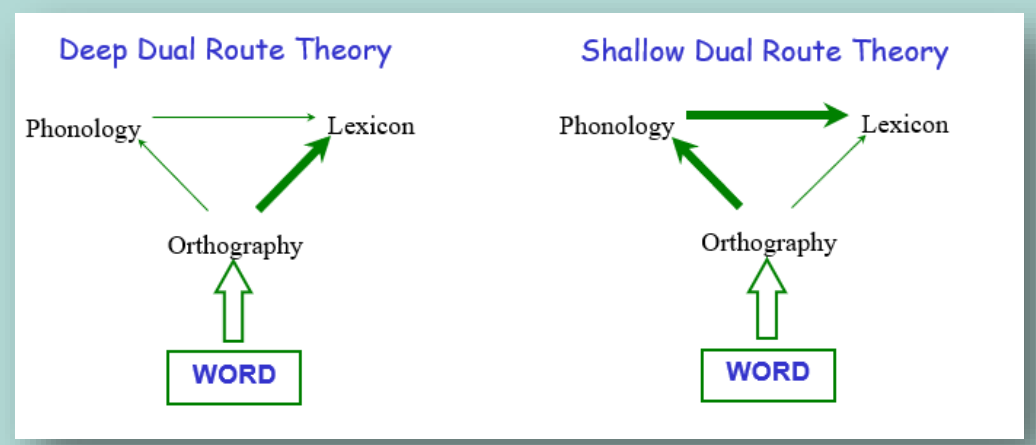
Kanji
Chinese
English
Italian
Tamil
Malay

Psycholinguistic Grainsize Theory

Ziegler & Goswami (2005)

The unit level of processing for reading is affected by:

- The consistency of Print-Speech,
- Availability of phonological units in the language,
- Granularity of Writing System

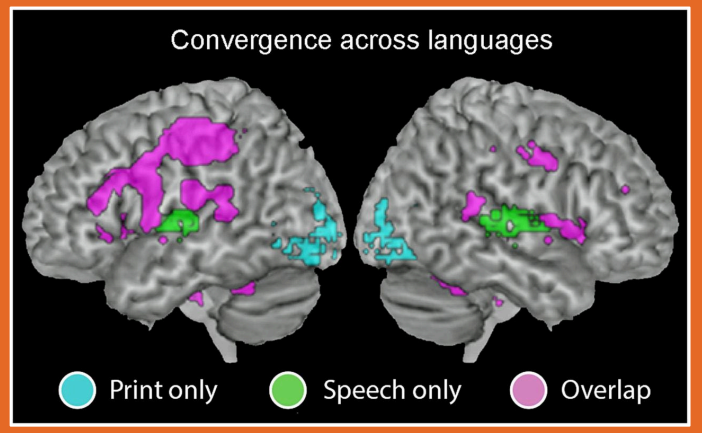


Lexical Constituency Model

Perfetti et al. (2005; 2013)

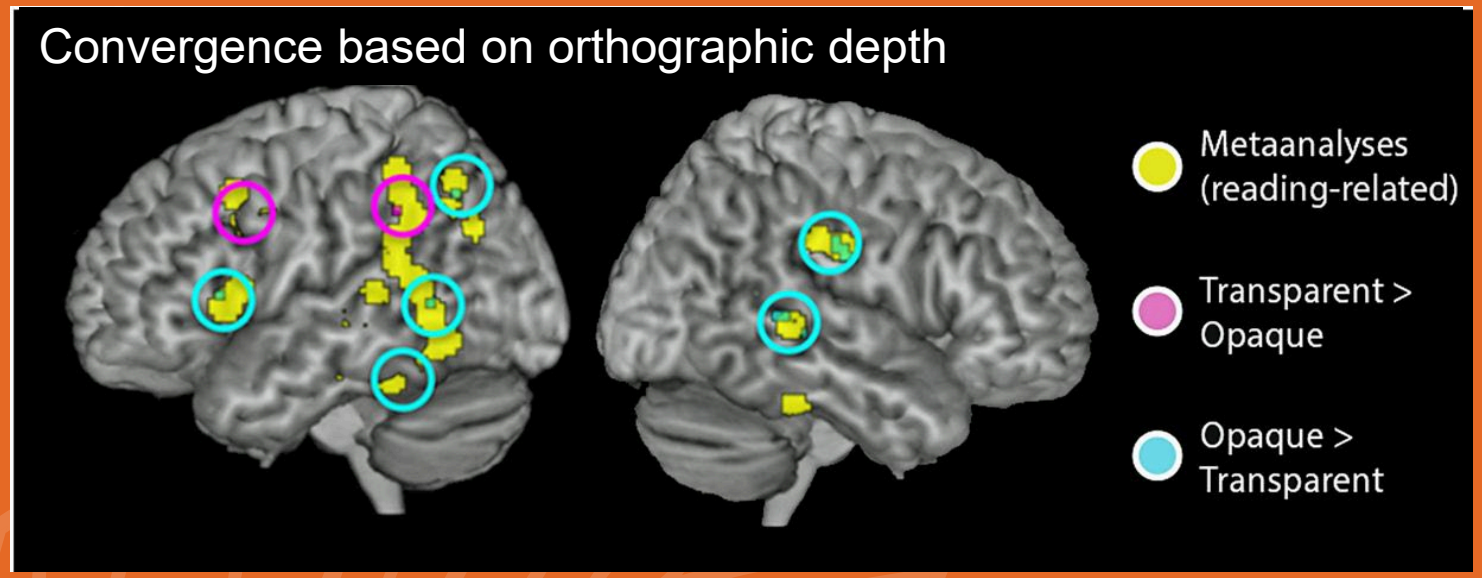
Readers use relevant Print-Speech units
which maximize efficiency

Overlapping brain networks



Semantic categorization task
Listening to or Reading words

Print-speech overlap



Spanish
English
Hebrew
Chinese

Rueckl et al. (2018) PNAS

Impact of script sets on the neural representations of reading

orthographic depth

OPAQUE — *TRANSPARENT*

Chinese

围棋

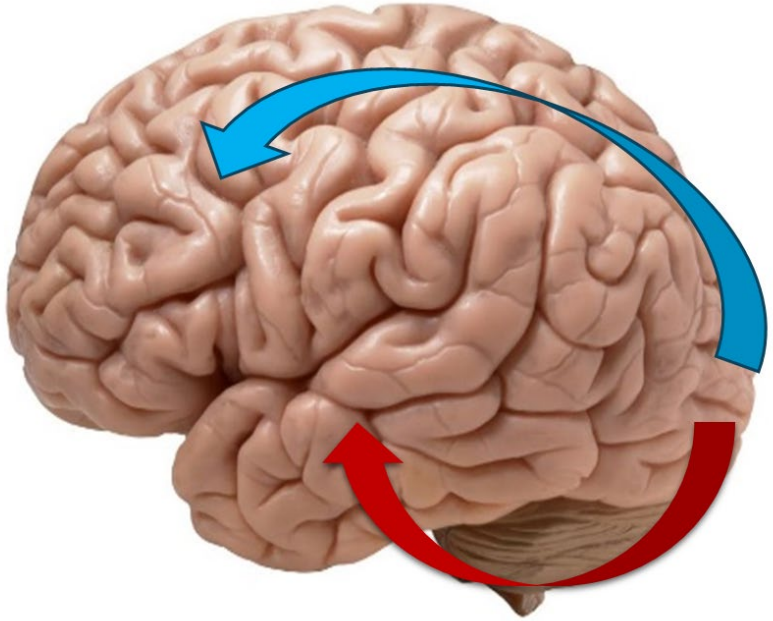
opaque relation of print to phonology

English

Tamil

கல்

close correspondence of print to phonology



Sublexical
(dorsal)

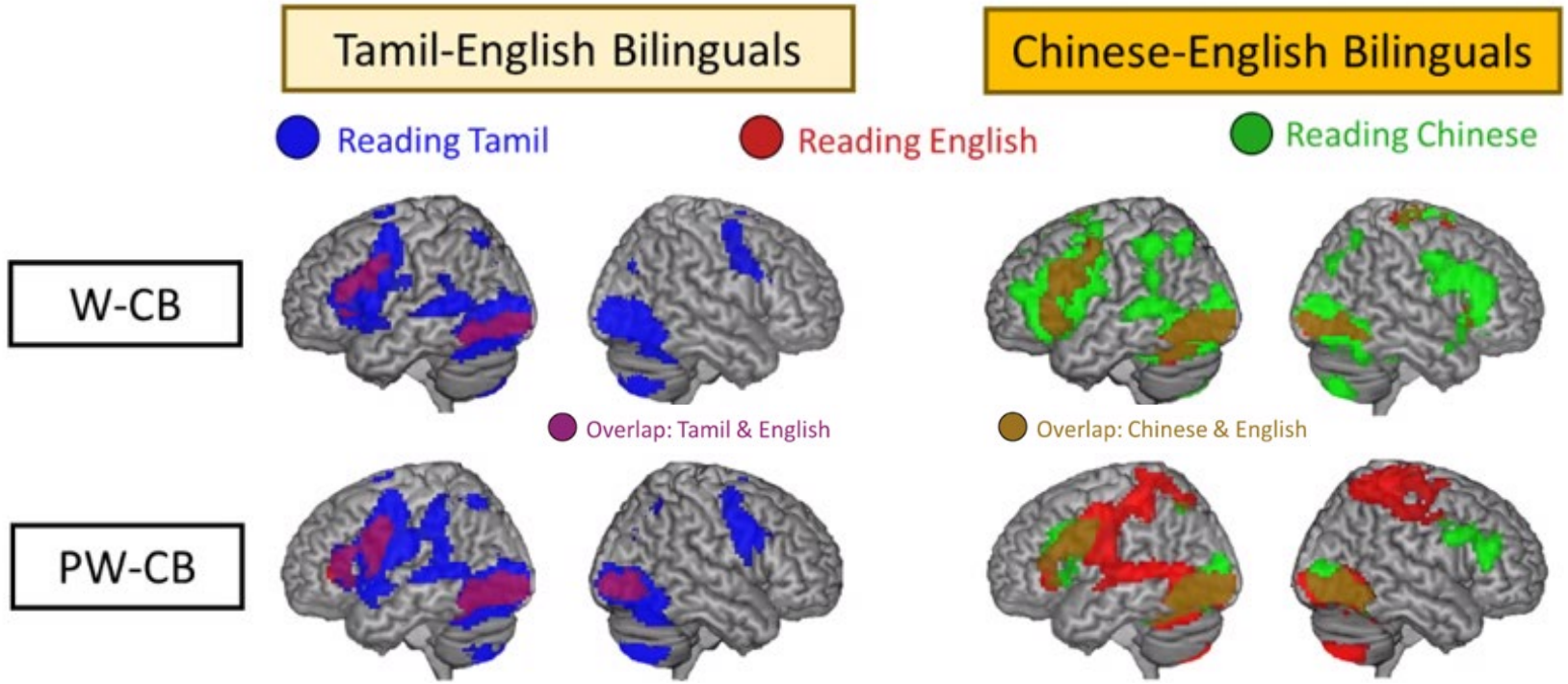
Lexical
(ventral)

Wu et al. (2020)

CRADLE Centre for Research and Development in Learning

Johns Hopkins Science of Learning Institute

Impact of script sets on the neural representations of reading



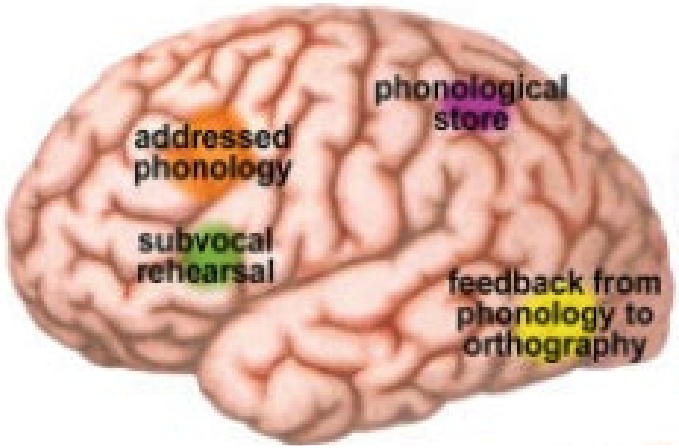
Reading transparent script (Tamil) engaged more sublexical processing in the dorsal stream (IFG) compared to English

Reading opaque script (Chinese) recruits lexico-semantic processing in the ventral stream (frontal areas), while reading English showed divergence for nonwords, engaging more dorsal (parietal) areas

	Chinese	English	Tamil
Word (W)	低	teach	கால்
Pseudoword (PW)	拔	smake	காத்
Nonword (NW)	彪	prtwn	ச்டச
Dummy (DM)	绿	green	நீலம்
Checkerboard (CB)	▩	▩	▩

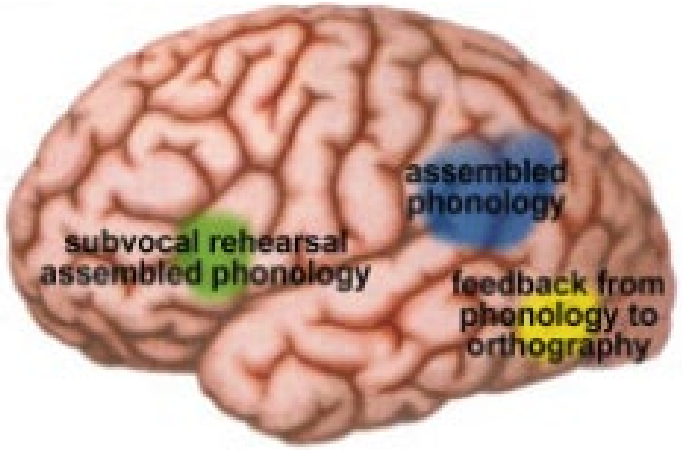
Wu et al. (2020)

Dual Route Model



a. Neural systems for phonological processing of written Chinese characters

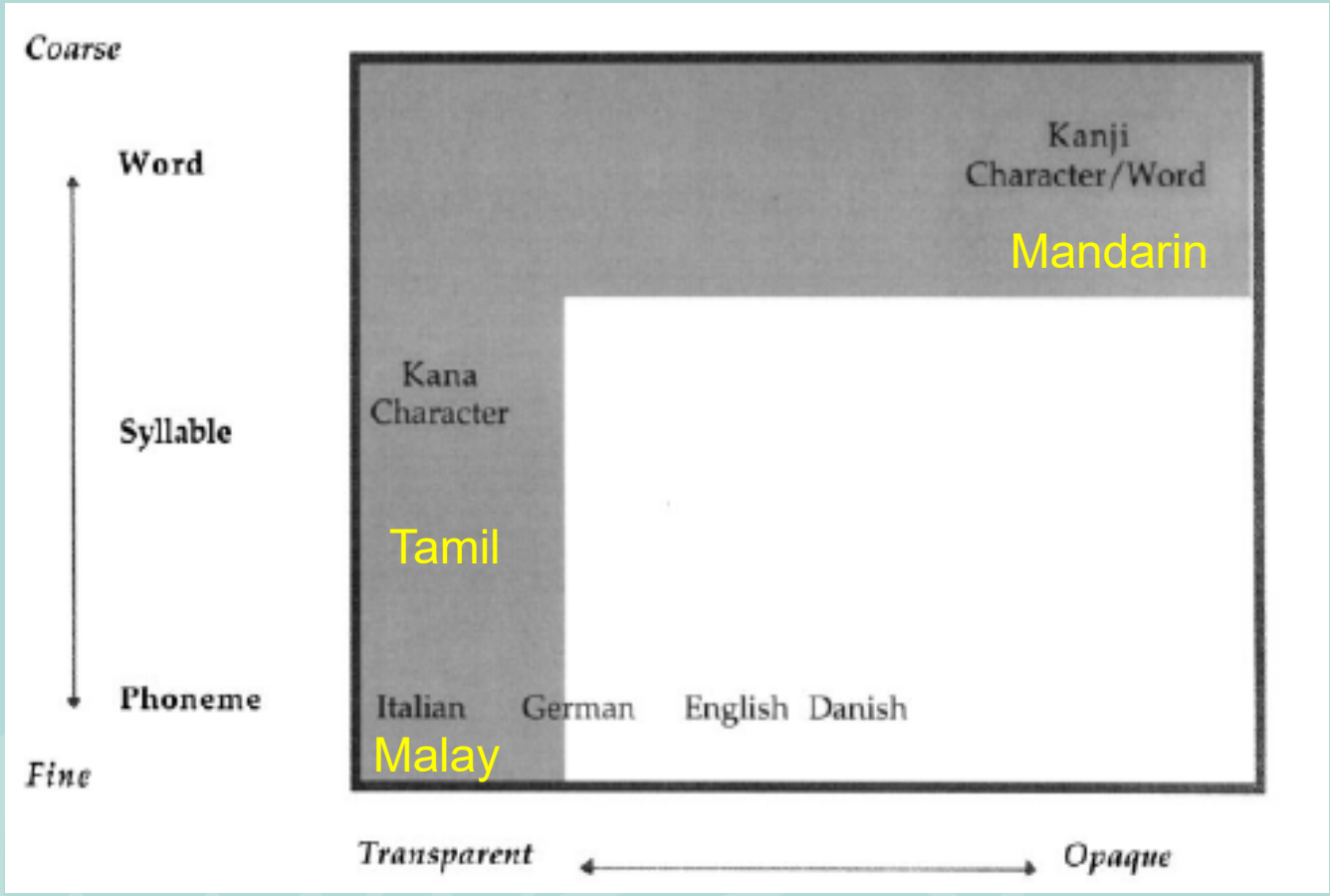
b. Neural systems for phonological processing of written alphabetic words



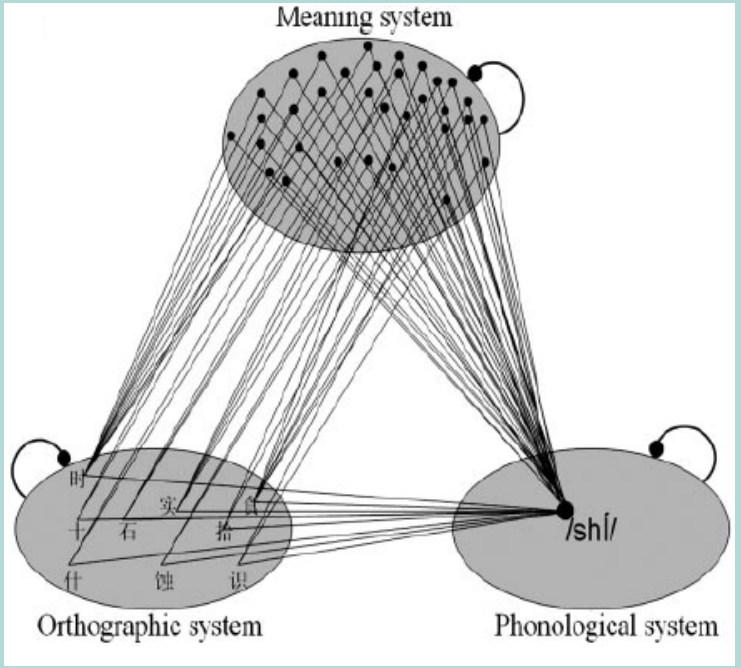
Tan et al. (2005)

Reading models across languages

Cognitive demands of reading



Wydell & Fern-Pollack (2012)



Tan, Spinks, Eden, Perfetti, Siok (2005)

Lexical Constituency Model
What are the relevant units that specify word identification?

Reading models across languages

Predictors of reading ability

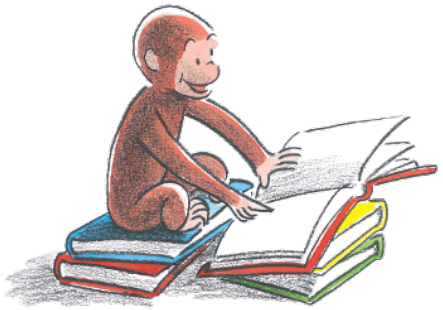
Phonological Awareness – strongest predictor of word reading ability in alphabetic languages (r=0.57); best discrimination of children with dyslexia (*ES* = - 1.37)

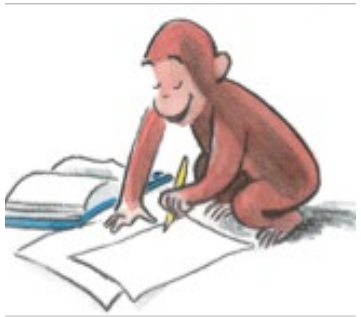
(Melby-Lervåg et al., 2012)

Phonological Awareness stronger correlate of reading in English than in Chinese
Morphological Awareness in Chinese produced significantly larger correlations with reading accuracy

(Ruan et al., 2017)

“Building blocks of reading may, therefore, comprise a variety of adjustments related to early word recognition” (p. 63, McBride, 2016)



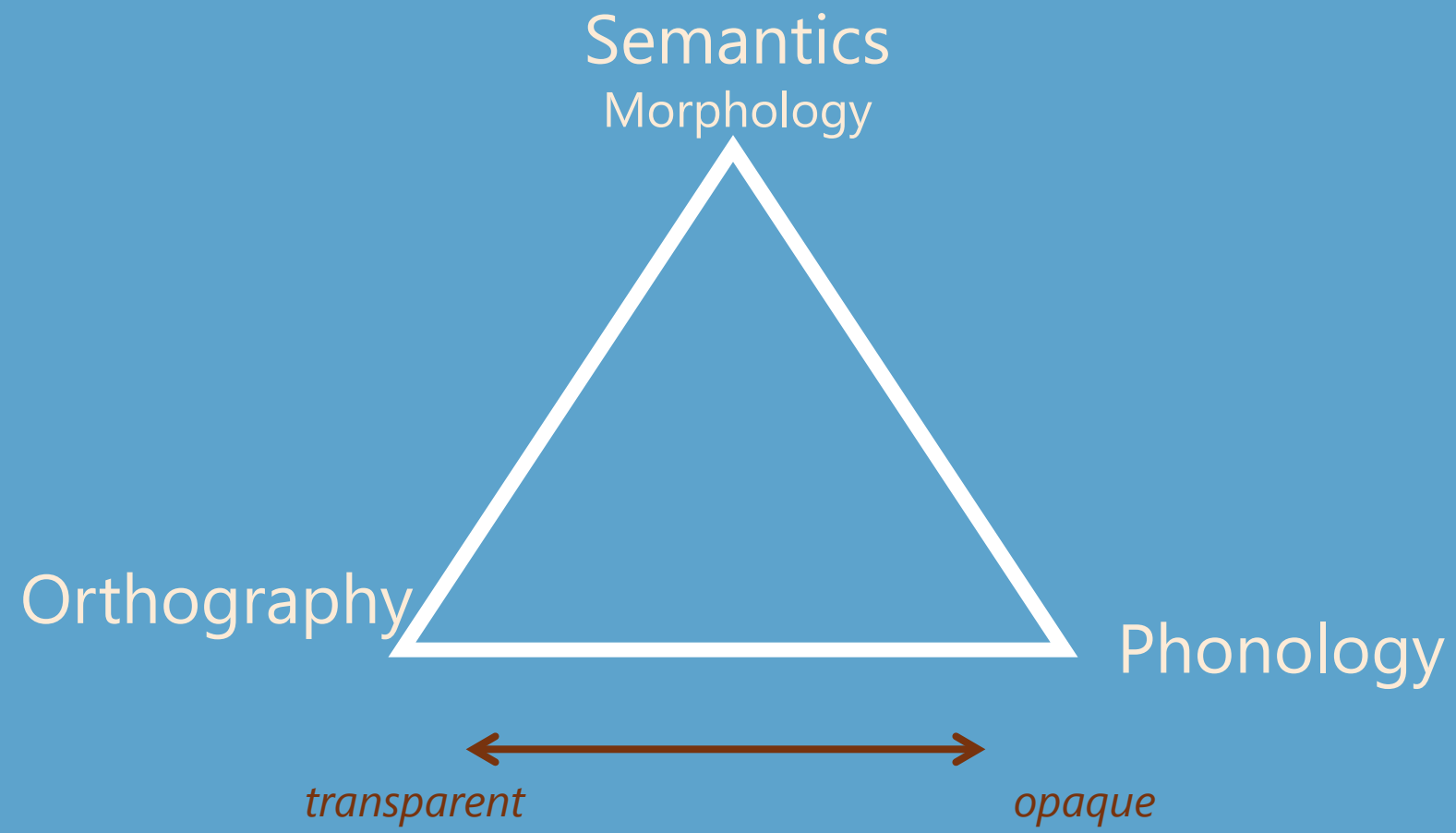


What about spelling?

Spelling

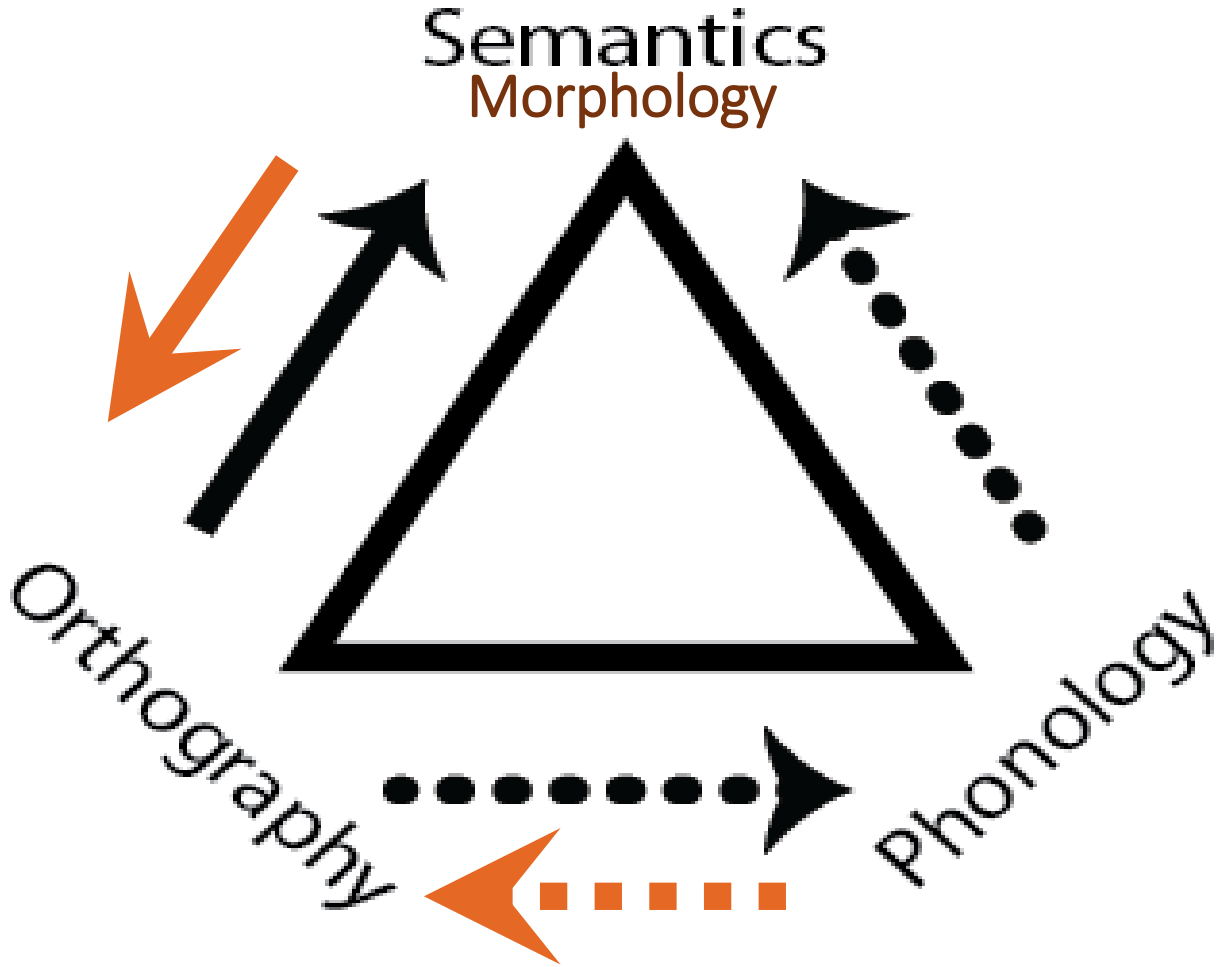
Cognitive demands across scripts

What information do spellers use in different languages?



Triangle Model

Triple Word Form Theory (Triple Code)



Plaut (1997)

Bahr et al. (2015)

Spelling error coding schemes

POMAS *Bahr et al. (2012)* **CoST** *Daffern et al. (2015)*

Phonological

An incorrect representation of the sounds. This type of error includes the use of an allophone, an omission or addition of phonological elements, which can also include tone, stress, and retroflex (supra segmental).

Orthographic

Errors are defined as spelling conveying the same phonology but with incorrect, ambiguous letters (e.g., pseudohomophones)

Morphological

Misspelling the target word/character with one that preserves the correct representation of sound but that has a different semantic, or a substitution with a semantically related word (e.g., homophones). This includes words, or parts of the word, that sound alike but have different meanings

Study 1

What information do spellers use in different languages?

PARTICIPANTS

568 Primary 1 bilingual children in Singapore

Mean age 6 years, 8 months

- 128 English + Malay
- 119 English + Tamil
- 321 English + Chinese

MEASURES

Completed a word dictation task, with 10 items selected from school curriculum list

SINGAPORE KINDERGARTEN IMPACT PROJECT

O'Brien, Habib, Arshad & Lim (2020)

script features

	stone	*pergi*
	English script:	**Malay script**:
Orthographic inventory	• 26 letters	• 25 letters (not X)
Phonological Syllable units	• Complex (6 types)	• Short , agglutinative, reduplication
Phonological representation	• morphophonemic	• phonemes > meaning
Word forms	• multi-letter words	• multi-letter words
Subword forms	• Consonant clusters, vowel digraphs	• Few consonant clusters, vowel pairs split at syllable boundary
Graphemic elements	• letter strokes, upper lower case	• letter strokes, upper lower case

script features

	Tamil script: Akshara	**Chinese script**: Characters
Orthographic inventory	• large (247)	• larger (K's)
Phonological Syllable units	• alphasyllabic	• morphosyllabic, non-alphabetic
Phonological representation	• phonemes > meaning	• meaning > phonology
Word forms	• multi-akshara words	• single to multi-character words
Subword forms	• consonant-vowel glyphs, diacriticals	• character components ~ semantic/phonetic radicals
Graphemic elements	• strokes ~ linear – usually L/R	• strokes ~ square – L/R, top/bottom, surround

Predictions

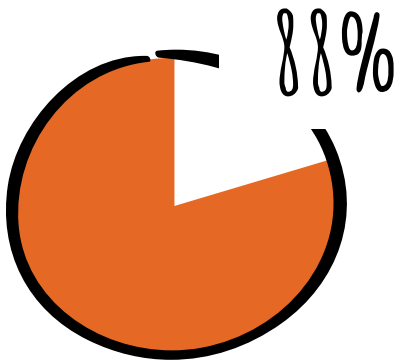
❖Script differences

We expected different types of spelling errors based on the most unambiguous units in the script:

- For Malay language, phonological errors would be most frequent, given that it is a very transparent alphabet; although previous studies suggests morphological awareness contributes to better spelling ability
- For Tamil, most frequent error types would be phonological, as akshara are linked to phonological syllables and this follows previous findings, although orthographic errors might be expected given the orthographic breadth
- For Chinese, morphological and orthographic errors would be most prevalent given the opaque relation of print to phonology and a lexical level of word identification.

Percentages of spelling error types

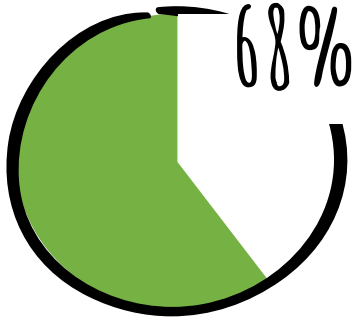
primary 1 students



Chinese

Morphemic errors (inc. Homophone substitutions), or wrong character

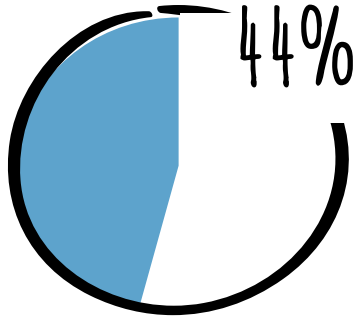
坐 instead of 做
(zuò) 'sit' for 'do'



Malay

Phonological errors
Vowel substitutions or omissions

'dena' instead of 'dan' (and/with)



Tamil

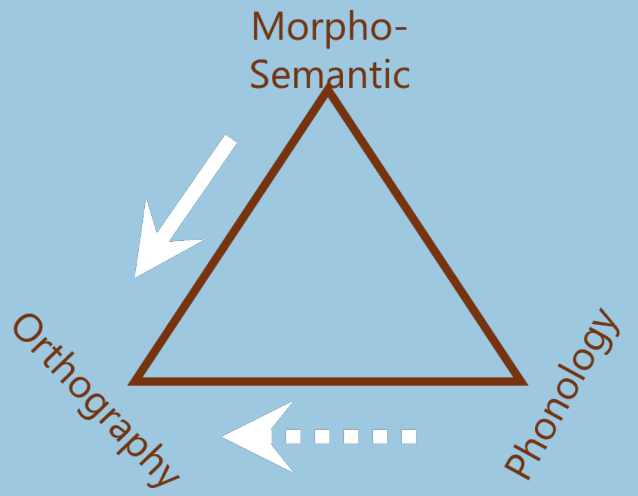
Phonological errors
Consonant, retroflex substitutions

இன்பம் instead of
இம்பம் (joy)

O'Brien, Habib, Arshad & Lim (2020)

SKIP SINGAPORE KINDERGARTEN IMPACT PROJECT

Results



Script differences

- ❖ Phonological Errors

  Malay* > Tamil > Chinese *issues with vowels

- ❖ Morphological Errors

  Chinese > Malay > Tamil* *no issues

- ❖ Orthographic-graphemic Errors

  Malay > Chinese* > Tamil *little issues with character configuration

- ❖ Other errors (blanks or unrelated words) predominated for Chinese responses

limitations

Item lists were ecologically valid, but offered uneven and limited opportunities for some error code types

Dictation task required full word response, yielding blank response or guesses

Cross-sectional study at one age level

The bilingual children’s English spelling was not analysed within the Triple Code framework

Study 2

What information do spellers use in different languages?

PARTICIPANTS

Bilingual cohorts in Singapore:

- Kindergarteners (Mean age = 4.90 years old)
- 1st Graders (Mean age = 6.86 years old)
- 3rd Graders (Mean age = 8.81 years old)

- 390 English + Malay
- 253 English + Tamil
- 761 English + Chinese

MEASURES. Spelling test using a cloze procedure to target specific features identified as exemplars of triple word form constructs (based on Daffern et al., 2015, CoST).

PROCEDURE. Children were asked to look at a target word and to fill in the blank with the correct letter(s)/akshara(s)/character to form the word (e.g., _ook for “book”). Each word has a corresponding picture above to aid in the identification of the word. As additional guidance to the kindergarten students, an audio of each word is played.

items designed for each age level

10.		11.		12.		13.		14.
	s___kor		ke____		mende___ar		_____hari	
15.		16.		17.		18.		19.
	___tas							

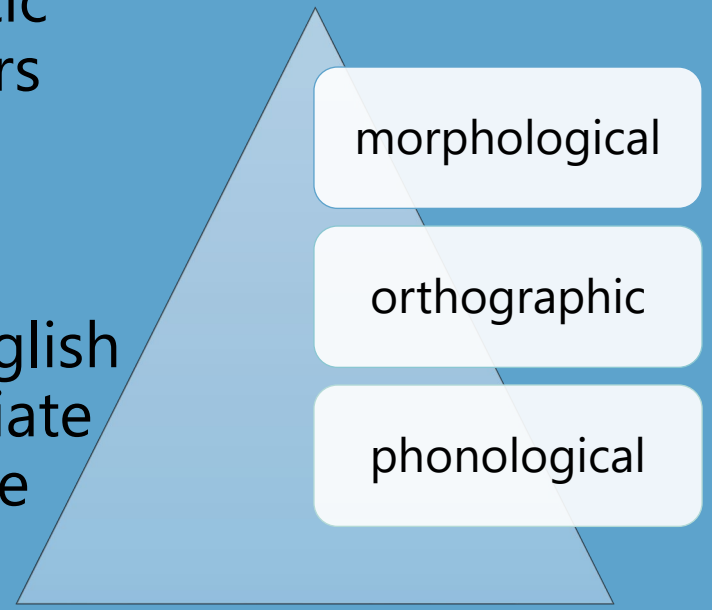
6.		7.		8.		9.	
	b___k		d___ll		cat___		di___
11.		12.		13.		14.	
	c___n		day___		an___mals		s___l

Predictions

- Script differences

  Following previous results, transparent alphabetic scripts would yield more phonological errors, non-alphabetic scripts and opaque scripts more morphological errors

Alternatively, shared strategies across English and each language may be determined by their typological distance; where Malay would most closely reflect English error types (phonological), while Chinese would deviate with more morphological errors, and Tamil with more orthographic errors

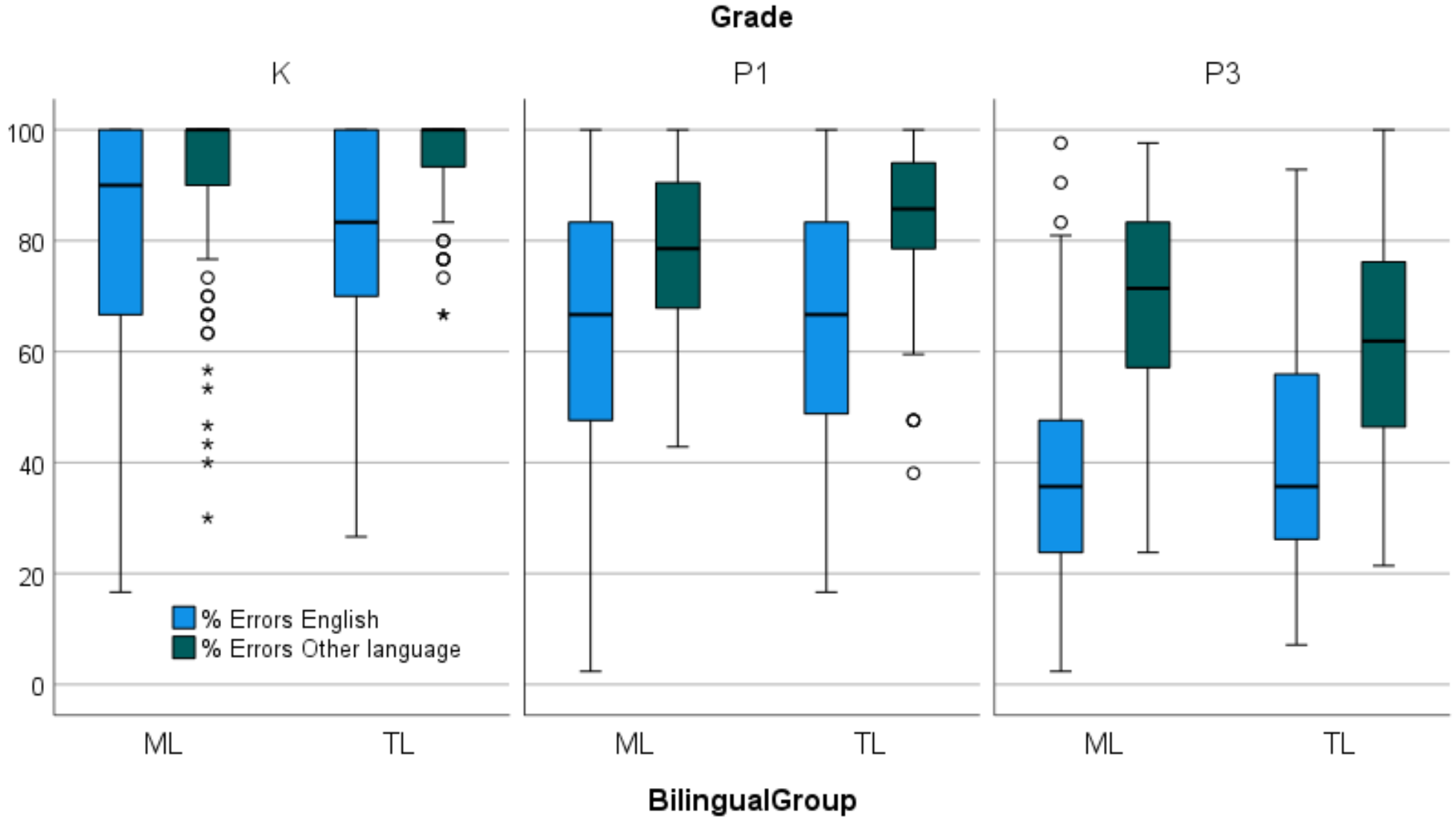


- Age differences

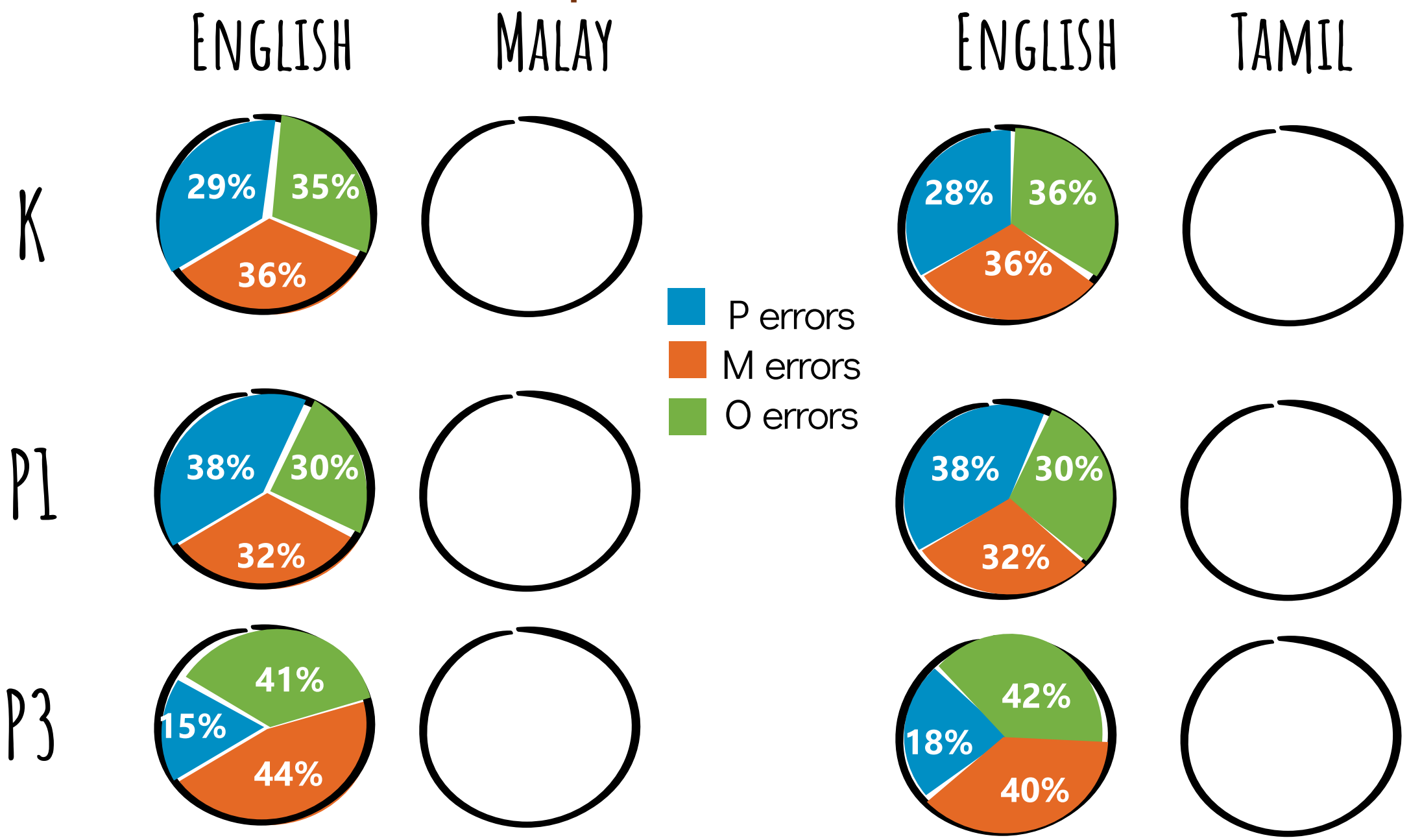
  Similar age differences per script, following spelling development phases (sound (P) → pattern (O) → meaning (M))

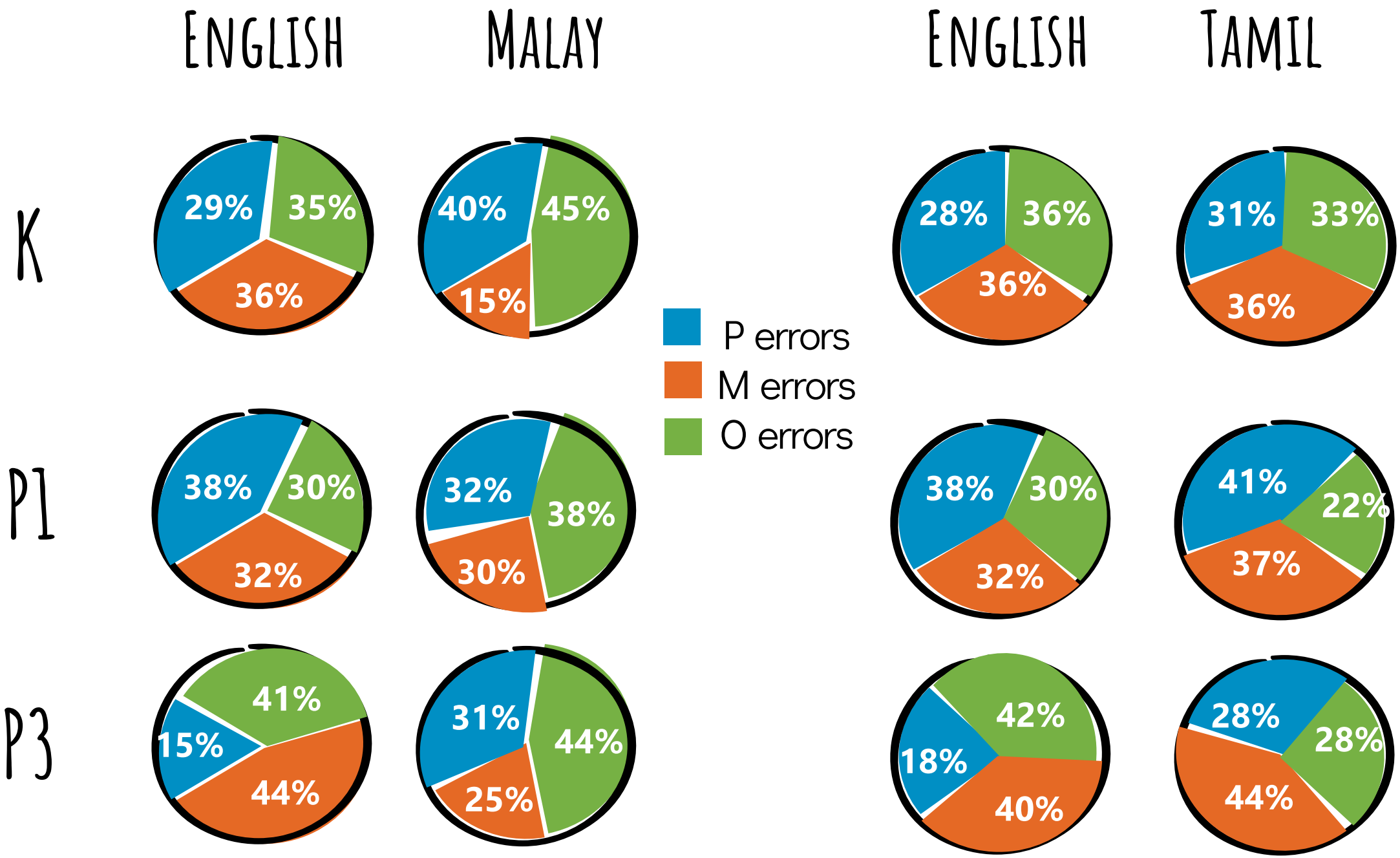
Results (preliminary)

% spelling errors
by Group and Grade levels



Triple Code Errors,




English
Malay
English
Tamil
K
P1
P3
29%
35%
36%
40%
45%
15%
28%
36%
36%
31%
33%
36%
38%
30%
32%
32%
38%
30%
38%
30%
32%
41%
22%
37%
15%
41%
44%
31%
44%
25%
18%
42%
40%
28%
28%
44%
P errors
M errors
O errors

❖Script differences

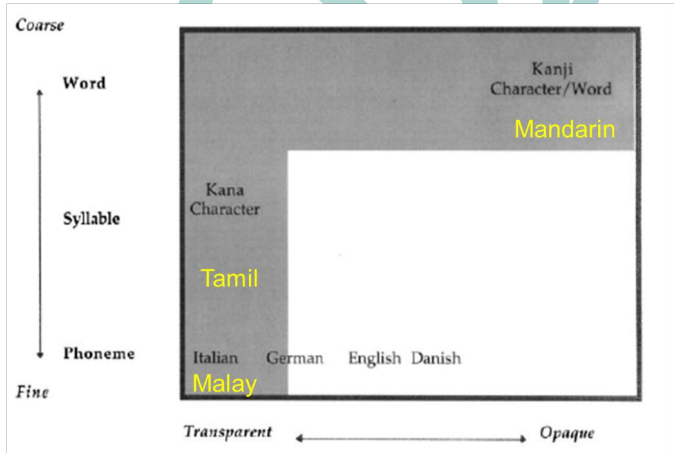
While Malay and Tamil are each relatively transparent scripts, they can code oral language at different grainsizes.

Phonological errors were not the most common for Malay (more orthographic errors), but they were as common or more than orthographic errors for Tamil.

Morphological errors played least role for Malay, but stronger role for Tamil.

As far as cross-linguistic strategies, both bilingual groups show similar spelling error type patterns for English words, without regard to typological distance between their script sets.

Cognitive demands of reading

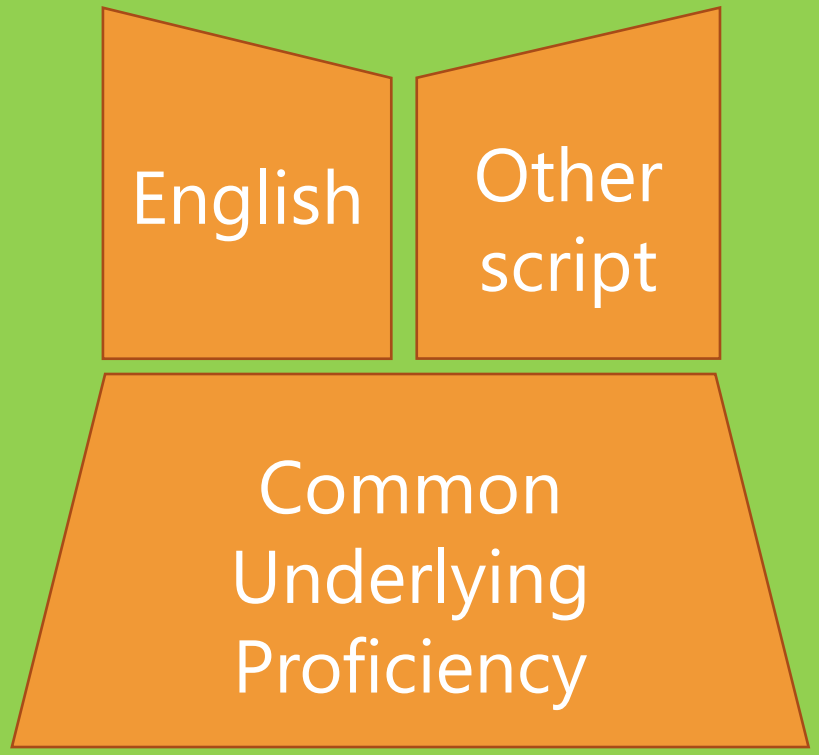


❖Age differences

Across grades, phonology played a more important role at primary entry level than at preschool or middle primary – in particular for English and Tamil. Whereas for Malay, the same pattern of errors persisted across these grade levels.

Thus, changes in spelling development may not follow developmental phases universally, but might be more affected by the nature of the script, the cognitive demands of spelling in that script, and strategies that spellers adopt to meet the demands.

Bilingual Literacy Development



Interdependence Hypothesis

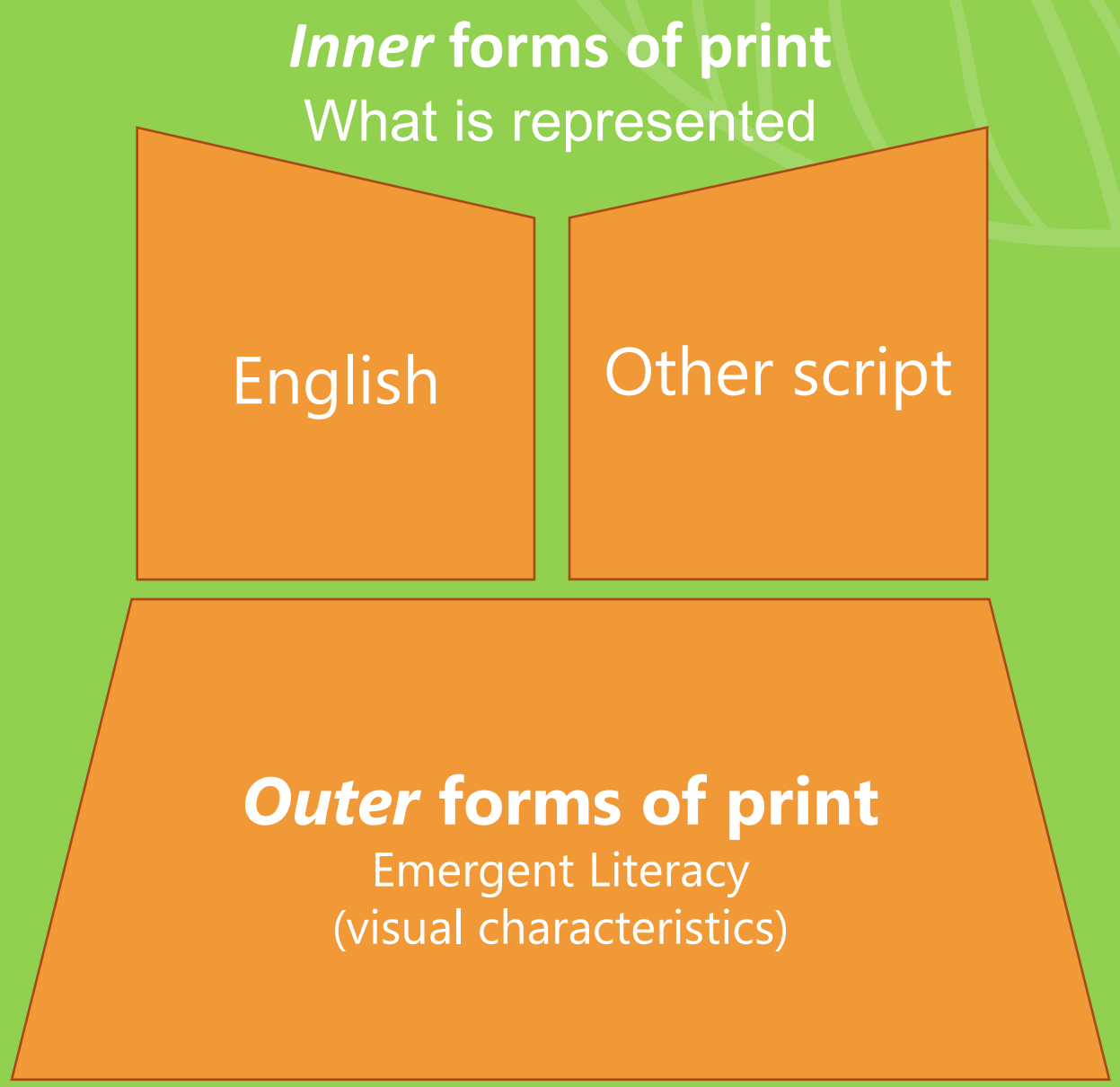
(Cummins, 1991)

a common underlying proficiency that supports both first and second languages

The linguistic and orthographic proximity hypothesis

(Kahn-Horwitz, Schwartz and Share, 2011)

Overlap depends on the degree of proximity between linguistic as well as orthographic structure of L1 and L2



Buckwalter & Lo (2002); Puranik & Lonigan, (2011); Treiman & Kessler (2014)

Cross language effects

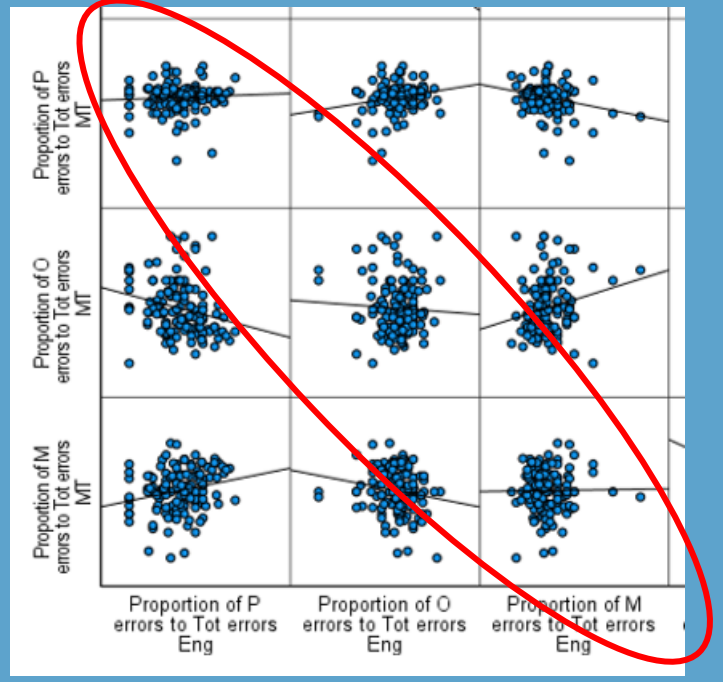
	P errors	O errors	M errors
English-Malay	**-0.145**	**0.228**	**-0.357**
English-Tamil	**0.545**	**0.247**	0.129

grade partialed out, bold $p < .01$

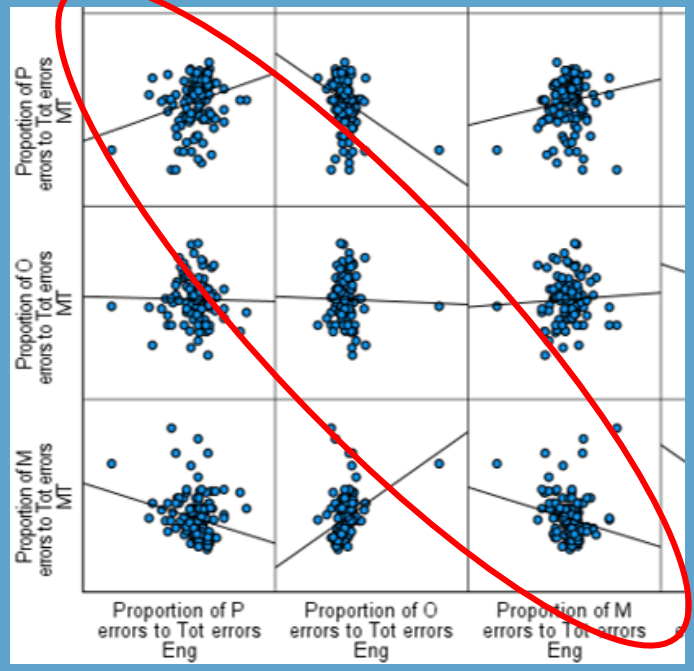
English-Malay

Cross-Language Relations by grade

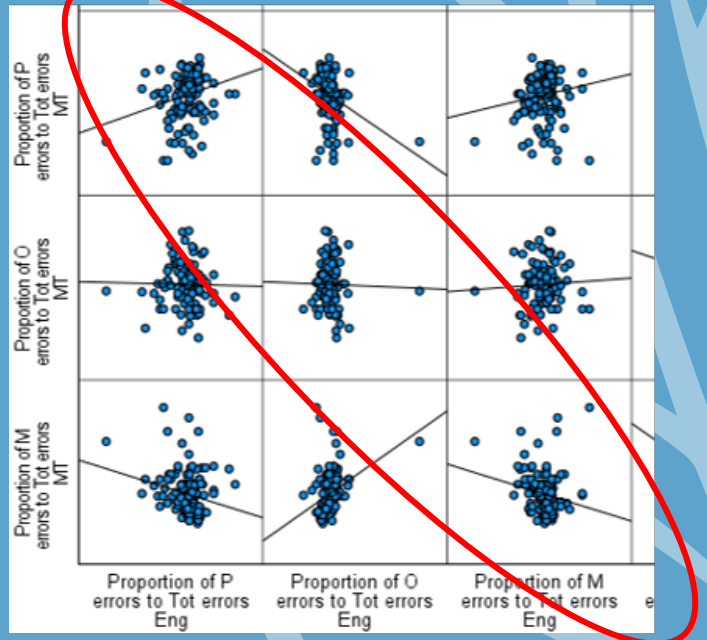
P3



P1



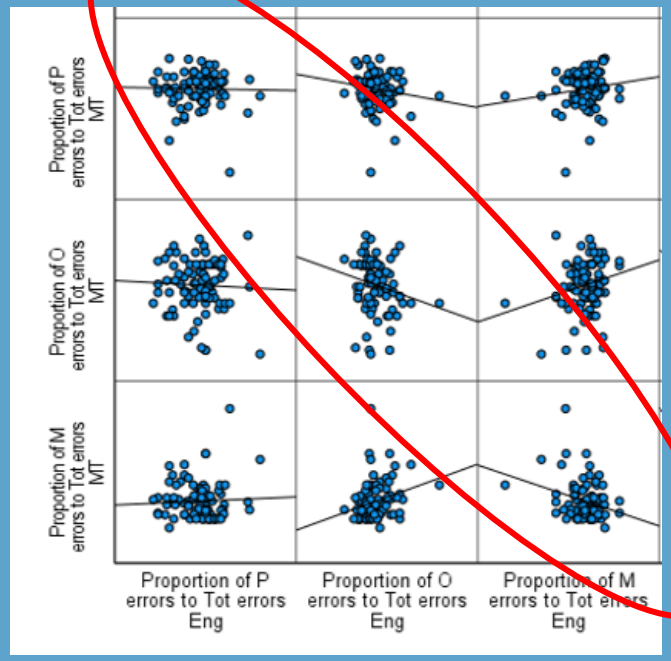
K



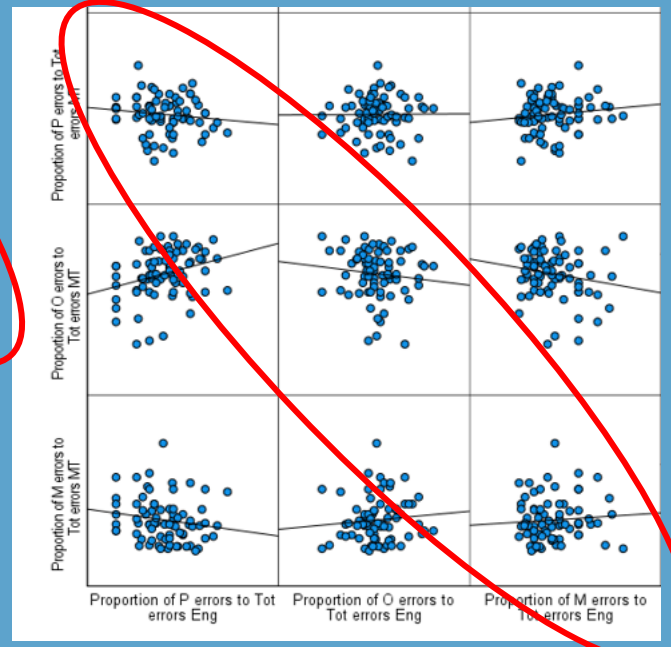
English-Tamil

Cross-Language Relations by grade

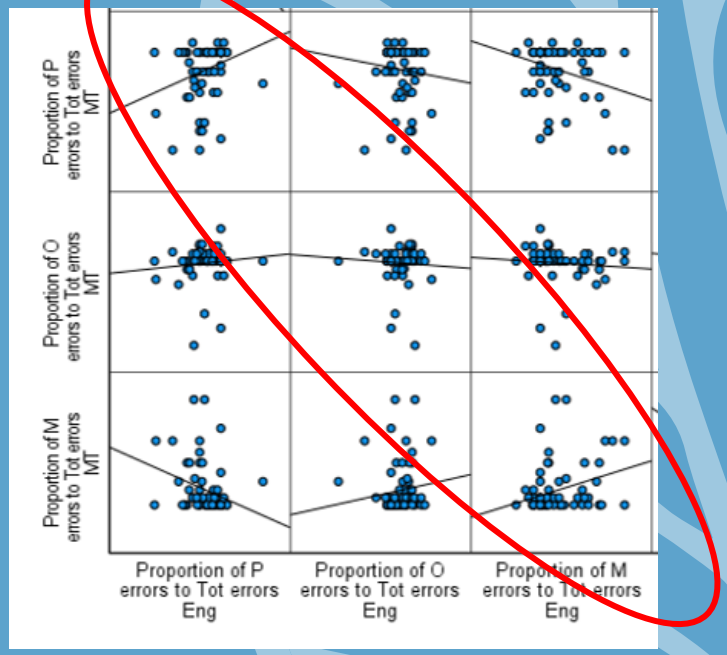
P3



P1



K



Conclusion

Children learning to read and write in more than one language need to meet the challenge of how each language specifies oral language.

While biliteracy studies suggest children will transfer their knowledge across languages, less is known about children acquiring literacy in multiple languages simultaneously.

Identifying universal and language-specific components of this learning process will help children, and their teachers, leverage on transferrable knowledge.

Implications

Teachers of beginning biliterates may use opportunities to draw children's attention to cross-language similarities as they learn to decode and encode print, or to process shared narratives

Pre-readers	Beginning readers	Developing readers	Older readers	Assessment
Importance of print and alphabetic knowledge	Decoding and encoding words	Managing more difficult words and sentence structures	Use of strategies to understand text	To identify at-risk individuals, test in multiple languages
Outer forms of print Can point to similarities where possible	To highlight strongest relations between codes (phonology / morpho-semantic – orthography) – Inner forms of print	Add strategies for decoding/encoding besides the strongest link	Metacognitive strategies can be applied across languages	Strategically plan skill assessment for more universal aspects of languages, along with script-specific aspects


Thank you for your attention

Thank you to our sponsors

SINGAPORE KINDERGARTEN IMPACT PROJECT

NIE NATIONAL INSTITUTE OF EDUCATION SINGAPORE

Questions & Discusson