

---

# First language effects on incidental vocabulary learning through bimodal input

A multisite, preregistered, and close replication of Malone (2018)

## ordered alphabetically by last name

Yingzhao Chen, Michigan State University (USA)

Jianwu Gao, Capital Normal University (China)

Bronson Hui, University of Maryland, College Park (USA)

Benjamin Kremmel, University of Innsbruck (Austria)

Shuang Ma, Capital Normal University (China)

Ryo Maie, The University of Tokyo (Japan)

Eva Puimège, KU Leuven (Belgium)

John Rogers, The Hong Kong Polytechnic University (Hong Kong)

Micheline Wilson, University of Maryland, College Park (USA)



@BronsonHui@mastodon.online

@BronsonHui

...

Just wanted to put it out there: If anyone would like to form a team to replicate an important vocab study, I would love to be involved.



**Kevin McManus** @KMcManus84 · Dec 15, 2021

Call for papers: #Replication in Second Language Research. Special Issue of Studies in Second Language Acquisition #OpenScience #SSLA  
#appliedlinguistics [twitter.com/linguistlist/s...](https://twitter.com/linguistlist/s/)

6:15 AM · Dec 15, 2021

---

1 Retweet 14 Likes

---



**Kevin McManus** @KMcManus84 · Dec 15, 2021

...

Replying to [@BronsonHui](#)

This would be a great topic!! Don't hesitate to get in touch if you have any questions about a proposal



3



# Incidental vocabulary learning (IVL)

- Acquiring vocabulary through a meaning-focused activity
- Learning after one exposure possible (e.g. Chen & Truscott 2010; Webb 2007)
- More exposures = more learning (e.g. Uchihara et al. 2019)
- Effectiveness of aural enhancement: mixed results

Beneficial (Webb & Chang 2022)

No difference (Brown et al. 2008)

# Malone (2018)

## Design

- 2x2 design:
  - Reading only & aural enhancement/ Reading while listening
  - 2 & 4 exposures
- Participants:
  - 80 intermediate ESL learners
  - Various L1s

## Results

- Effect of aural enhancement:
  - Form recognition:
    - AE effects for 2 exposures
  - Form-meaning:
    - AE effects for 2 & 4 exposures
- number of exposures
  - All groups gained knowledge
  - 4 exposures led to more knowledge
- working memory capacity
  - Lower WM did not cancel out AE benefits

# Malone (2018)

## Design

- 2x2 design:
  - Reading only & aural enhancement/ Reading while listening
  - 2 & 4 exposures
- Participants:
  - 80 intermediate ESL learners
  - **Various L1s**

## Results

- Effect of aural enhancement:
  - Form recognition:
    - AE effects for 2 exposures
  - Form-meaning:
    - AE effects for 2 & 4 exposures
- number of exposures
  - All groups gained knowledge
  - 4 exposures led to more knowledge
- working memory capacity
  - Lower WM did not cancel out AE benefits

# L1 Effects

## L1 effects found for...

- L2 lexical processing (e.g. Wang, Koda & Perfetti 2003)
- L2 word learning (e.g. Hamada & Koda 2008)
- L2 word decoding & reading (Ben-Yehudah et al. 2019)

## No L1 effects found for...

- L2 word recognition (Lemhöfer et al. 2008)
- L2 incidental vocabulary learning (Tang & Chan 2021)

# Malone (2018)

## Our design (close replication)

- 2x2 design:
  - Reading only & aural enhancement/ Reading while listening
  - 2 & 4 exposures
- Participants:
  - 131 EFL learners
  - 4 L1 groups:
    - 2 Germanic
    - 2 Chinese

## Results

- Effect of aural enhancement:
  - Form recognition:
    - AE effects for 2 exposures
  - Form-meaning:
    - AE effects for 2 & 4 exposures
- number of exposures
  - All groups gained knowledge
  - 4 exposures led to more knowledge
- working memory capacity
  - Lower WM did not cancel out AE benefits

# Research Questions

RQ1: To what extent can we replicate the findings of Malone?

(a) form recognition, (b) meaning connection, and (c) role of working memory capacity

RQ2: What are the effects of L1s of different levels of grapheme-correspondence with English?

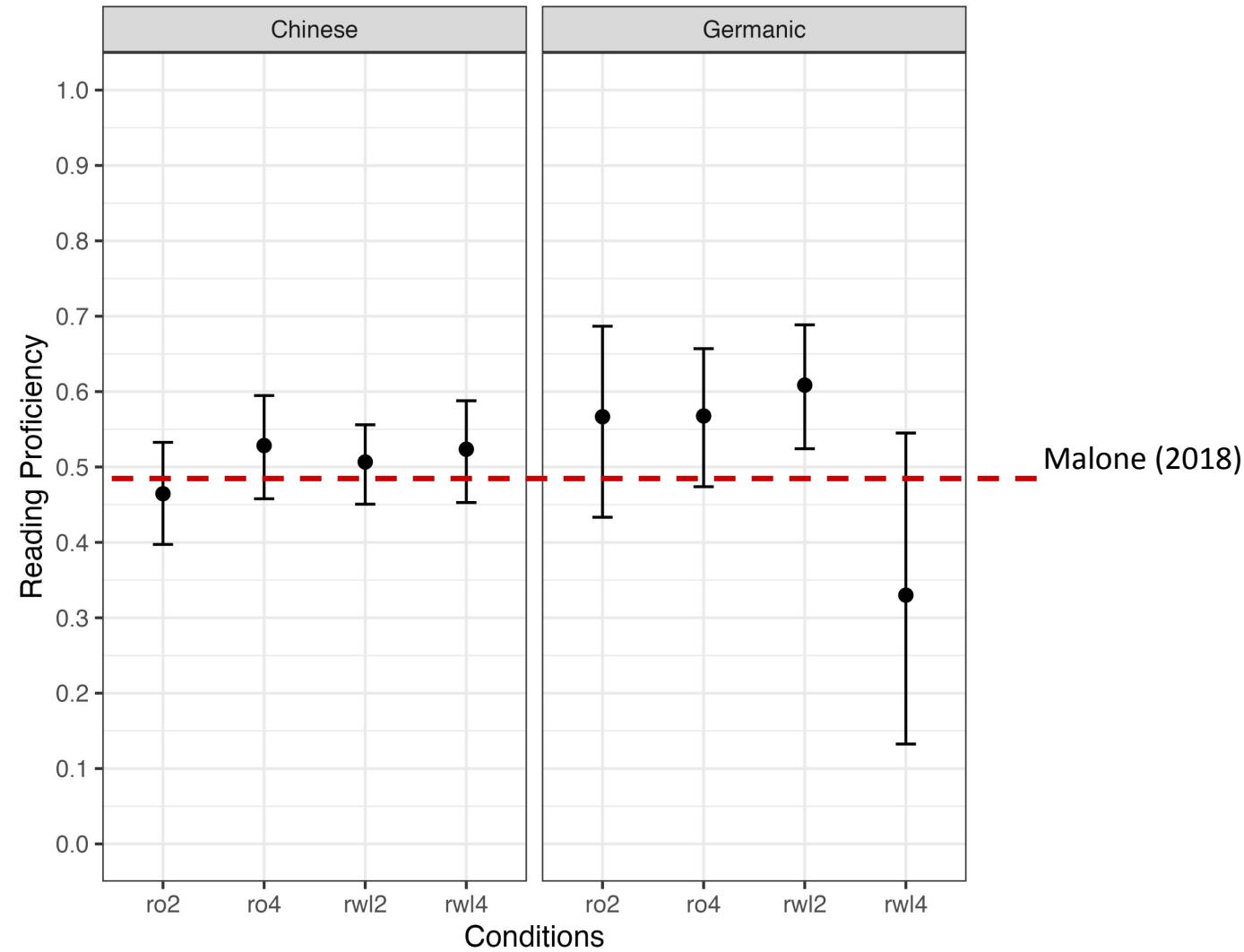
(a) form recognition, (b) meaning connection, and (c) role of working memory capacity

# Methodology

# Participants

- 131 L2 learners of English
  - 84 L1 speakers of Chinese (Mandarin & Cantonese)
  - 47 L1 speakers of Germanic languages (German & Dutch)
- Proficiency:
  - Comparable with Malone (2018)
  - Measured by a cloze test (Brown, 1980)

# Proficiency Comparison



# Materials

- 4 short stories
  - 96% words from the 1K-4K frequency bands in COCA
- Reading comprehension questions during reading (4 per story)
- 32 target words
  - Low-frequency concrete nouns (e.g., *ibis*, *petrel*, *civit*, *fescue*)

# Treatment Conditions

- 2 x 2 between-subject design: 4 treatment groups
- Variable 1: Frequency of exposure to target words: 2 or 4 times
- Variable 2: Modality: reading-while-listening or reading only
  - **Group 1:** 2 exposures - reading-only
  - **Group 2:** 4 exposures - reading-only
  - **Group 3:** 2 exposures - reading-while-listening
  - **Group 4:** 4 exposures - reading-while-listening

# More on modality

- Reading-only: participants only read the text
- Reading-while-listening:
  - Participants read and listen to the text simultaneously
  - Audio speed: 120 - 140 words / minute
- In both modes, a line of text will disappear after a certain time
  - Reading-while-listening: the line disappears following its recording
  - Reading-only: the display duration of a line matches that in reading-while-listening

# Vocabulary Posttests (Form Recognition)

- |                                   |                                    |                                    |
|-----------------------------------|------------------------------------|------------------------------------|
| <input type="checkbox"/> Fossa    | <input type="checkbox"/> Kestrel   | <input type="checkbox"/> Morel     |
| <input type="checkbox"/> Bootleg  | <input type="checkbox"/> Sorrel    | <input type="checkbox"/> Lemming   |
| <input type="checkbox"/> Folktale | <input type="checkbox"/> Sinew     | <input type="checkbox"/> Sumac     |
| <input type="checkbox"/> Riptide  | <input type="checkbox"/> Spoonbill | <input type="checkbox"/> Heathen   |
| <input type="checkbox"/> Mastic   | <input type="checkbox"/> Heifer    | <input type="checkbox"/> Freesia   |
| <input type="checkbox"/> Melange  | <input type="checkbox"/> Nadir     | <input type="checkbox"/> Atoll     |
| <input type="checkbox"/> Gentian  | <input type="checkbox"/> Prattle   | <input type="checkbox"/> Fracas    |
| <input type="checkbox"/> Holler   | <input type="checkbox"/> Subplot   | <input type="checkbox"/> Yuppy     |
| <input type="checkbox"/> Hovel    | <input type="checkbox"/> Paging    | <input type="checkbox"/> Huckster  |
| <input type="checkbox"/> Burble   | <input type="checkbox"/> Civet     | <input type="checkbox"/> Liftoff   |
| <input type="checkbox"/> Briar    | <input type="checkbox"/> Ibis      | <input type="checkbox"/> Chervil   |
| <input type="checkbox"/> Widget   | <input type="checkbox"/> Baggie    | <input type="checkbox"/> Pontoon   |
| <input type="checkbox"/> Fescue   | <input type="checkbox"/> Sculpin   | <input type="checkbox"/> Bullock   |
| <input type="checkbox"/> Snafu    | <input type="checkbox"/> Pinkie    | <input type="checkbox"/> Shortwave |

# Vocabulary Posttests (Form-Meaning Connections)

Q2

(1) A **cabal** is...

- a. a group of people
- b. a type of bird
- c. a type of food

# Working Memory Tasks

Nonword span task: phonological short-term memory

- Nonwords presented in lists of 7
- Participants then saw another list of 14 nonwords and decided whether they had seen the nonwords before in the task

# Working Memory Tasks

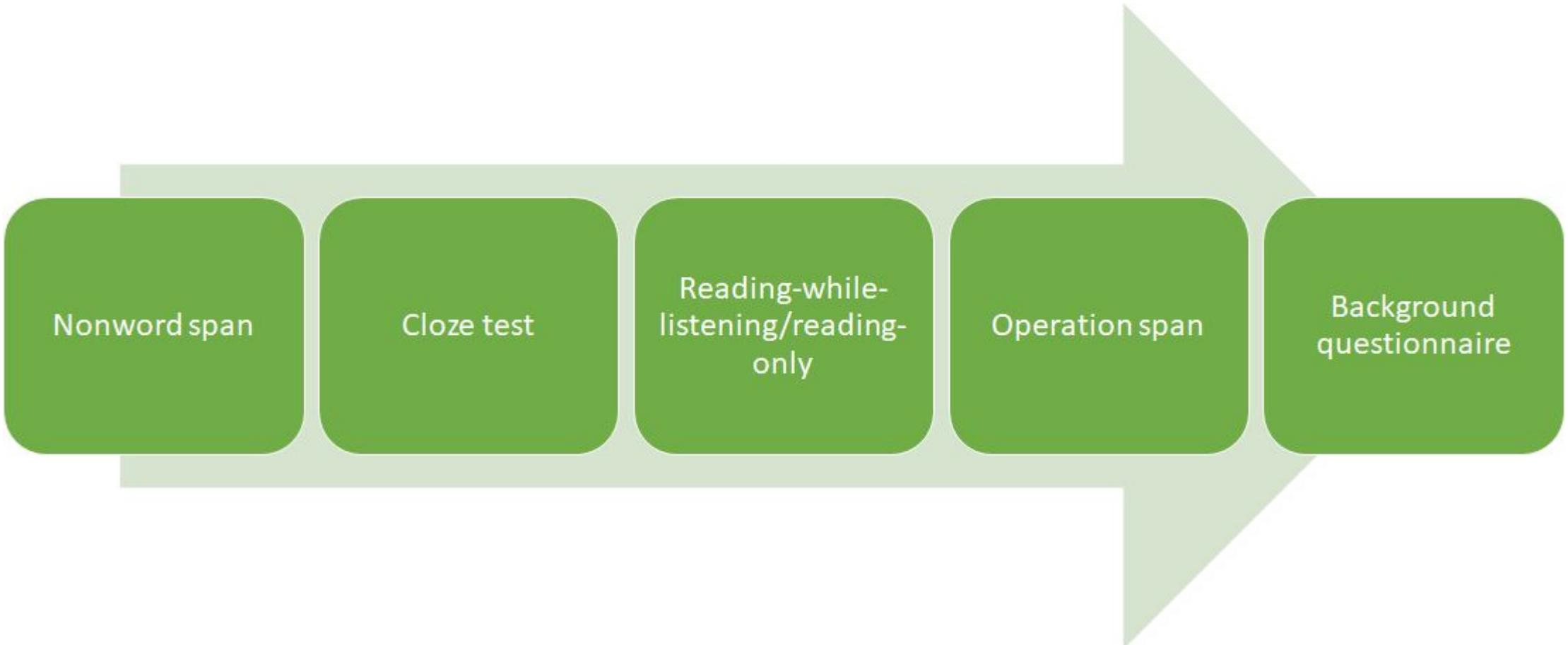
Operation span task: complex measure of working memory

- Participants saw a mathematical equation and a letter
- Example:  $3 + 7 = 9$  U
- Participants read aloud the equation
- Participants decided whether the equation was correct
- Participants recalled the letters in groups of 2 to 5

# Working Memory Tasks

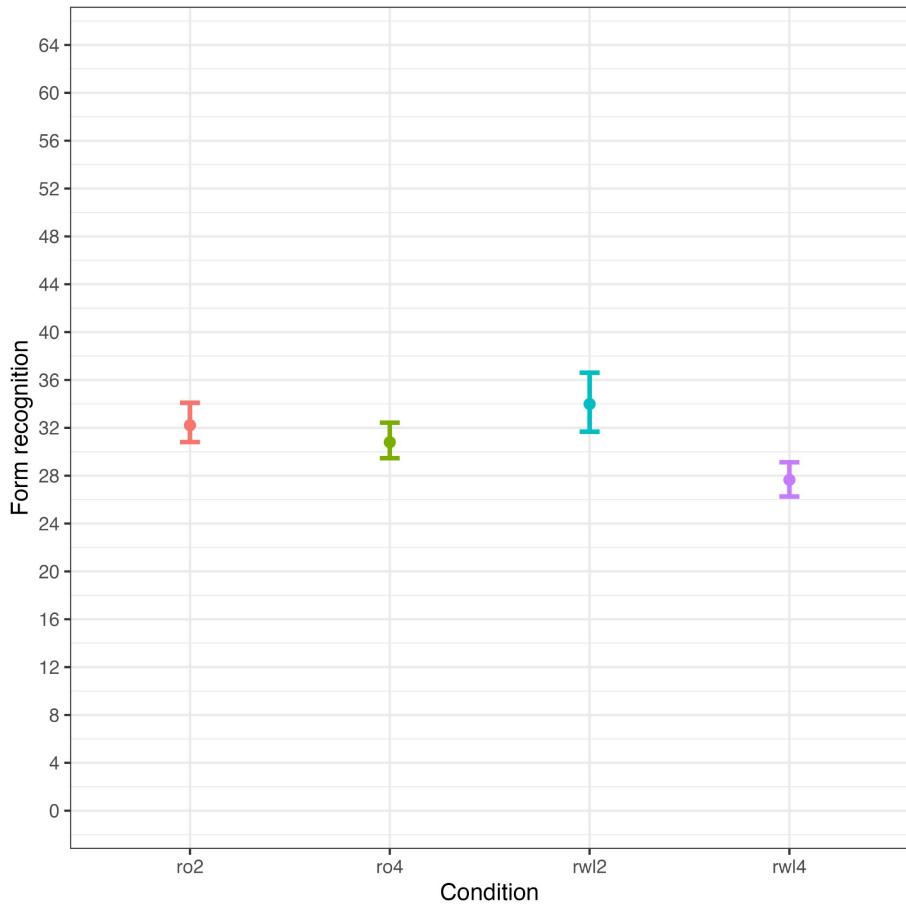
- The Shapebuilder Task (Atkins et al., 2013): nonlinguistic working memory
- Dropped in the current study due to lack of access to it

# Procedure: ~ 2 hours; online data collection



# Results

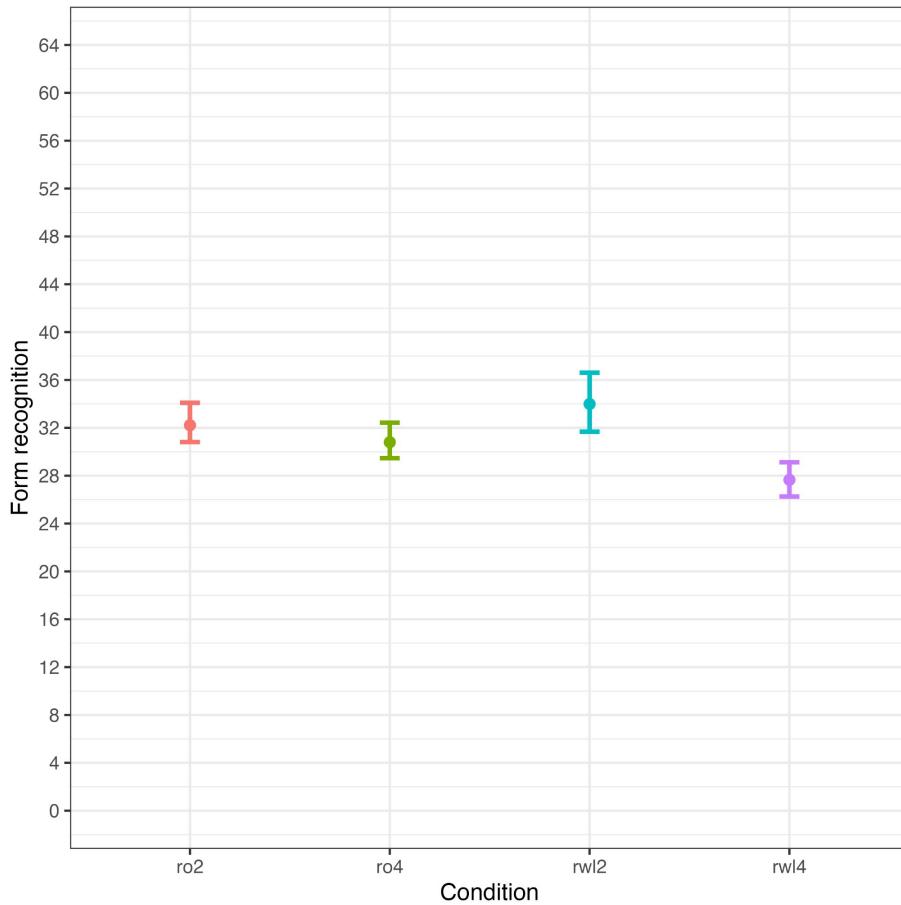
# Replicating Malone (2018) - Form recognition (RQ1a)



	Hypothesis	p(H data)	Malone	Us
H0:	$\text{rwI4} = \text{rwI2} = \text{ro4} = \text{ro2}$	.011		
H1:	$\text{rwI4} > \text{rwI2} = \text{ro4} > \text{ro2}$	.000		
H2:	$\text{rwI4} = \text{rwI2} = \text{ro4} > \text{ro2}$	.001		
H3:	$\text{rwI4} = \text{rwI2} > \text{ro4} > \text{ro2}$	.000	✓	
Ha:	<b>no specific restriction</b>	.988		✓

There is at least an effect but it is none of the original hypotheses

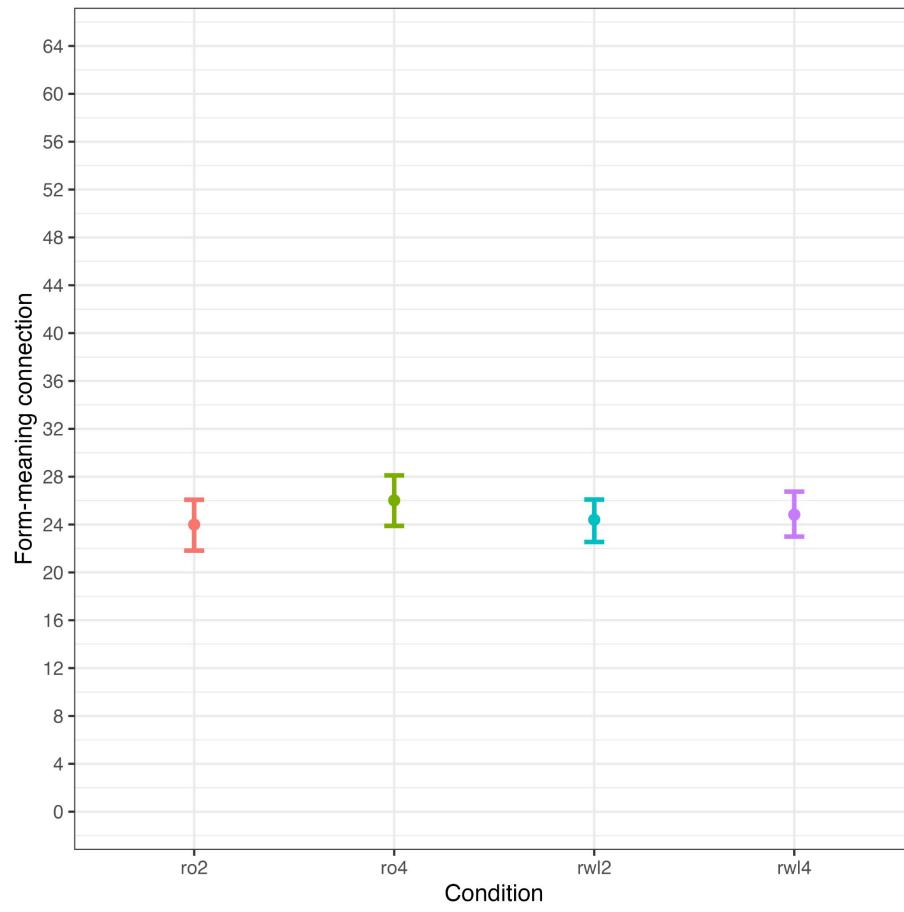
# Replicating Malone (2018) - Form recognition (RQ1a)



	Hypothesis	p(H data)	Malone	Us
H0:	$\text{rwI4} = \text{rwI2} = \text{ro4} = \text{ro2}$	.000		
H1:	$\text{rwI4} > \text{rwI2} = \text{ro4} > \text{ro2}$	.000		
H2:	$\text{rwI4} = \text{rwI2} = \text{ro4} > \text{ro2}$	.000		
H3:	$\text{rwI4} = \text{rwI2} > \text{ro4} > \text{ro2}$	.000	✓	
H4:	$\text{rwI2} > \text{ro4} = \text{ro2} > \text{rwI4}$	.957		✓
Ha:	no specific restriction	.042		

When adding a data-driven hypothesis, it becomes the winner outright!

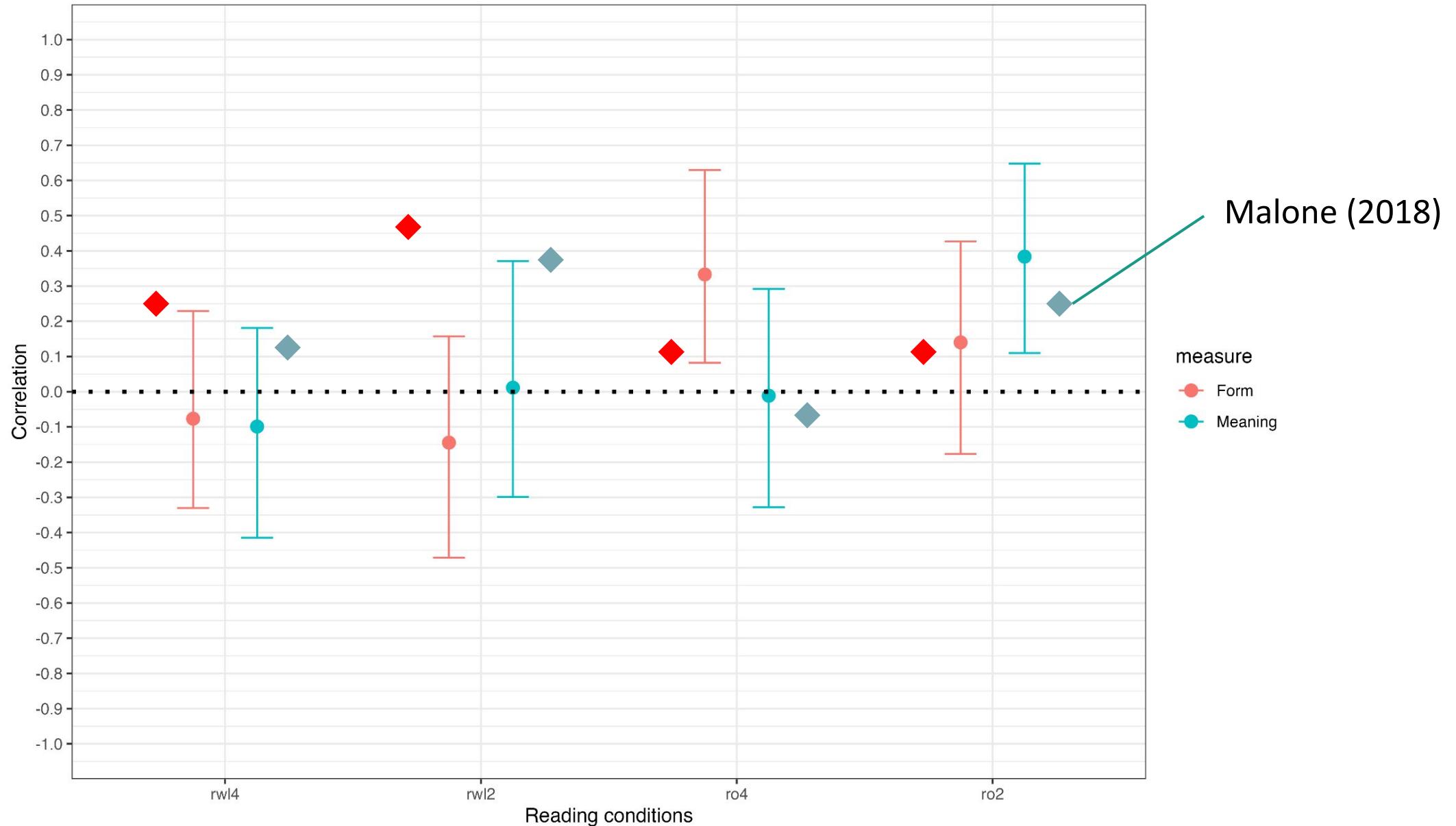
# Replicating Malone (2018) - Form-meaning connection (RQ1b)



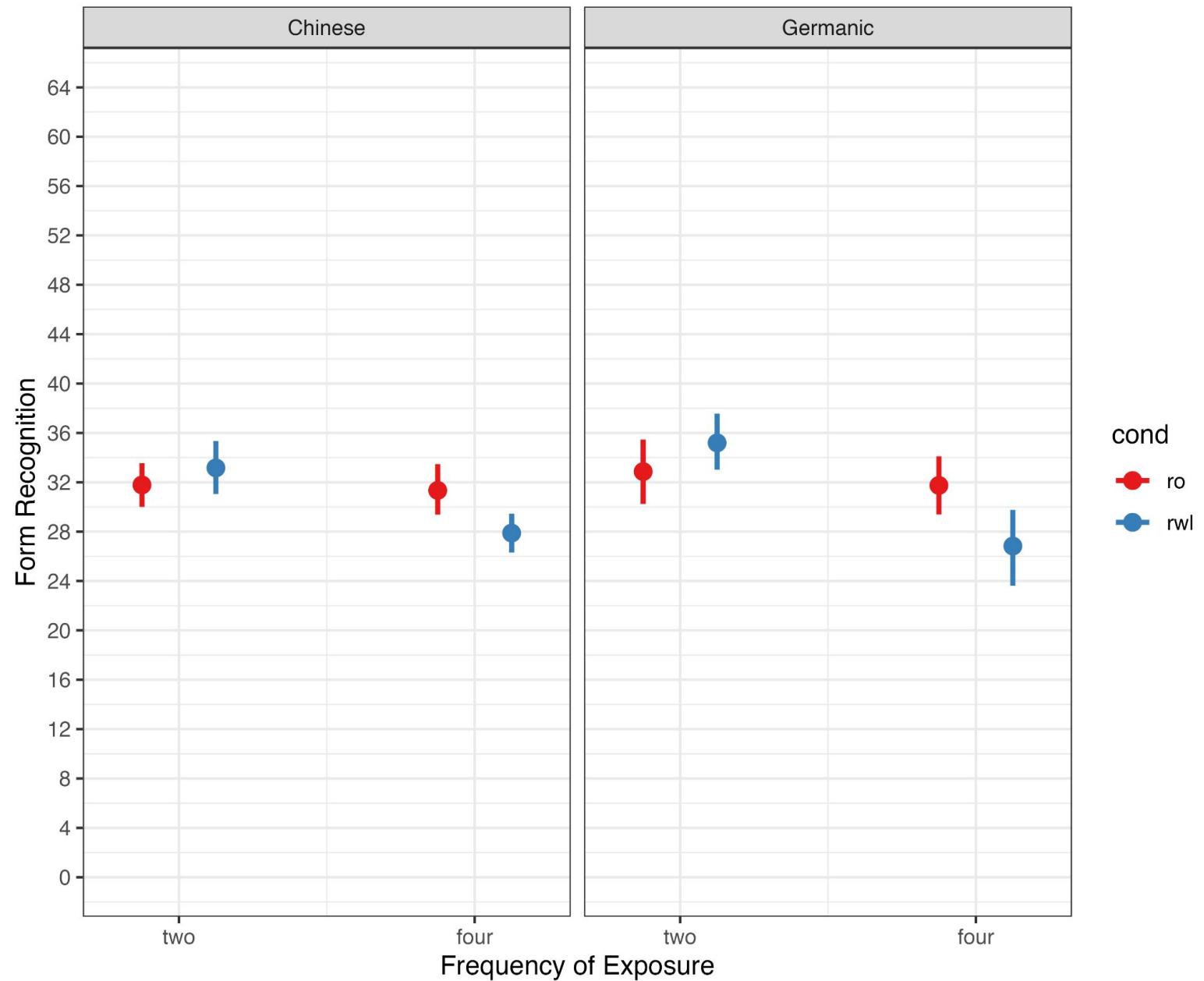
	Hypothesis	p(H data)	Malone	Us
H0:	$\text{rwI4} = \text{rwI2} = \text{ro4} = \text{ro2}$	.850		✓
H1:	$\text{rwI4} > \text{rwI2} > \text{ro4} > \text{ro2}$	.004		
H2:	$\text{rwI4} > \text{rwI2} = \text{ro4} > \text{ro2}$	.032	✓	
H3:	$\text{rwI4} = \text{rwI2} \& \text{ro4} > \text{ro2}$	.105		
Ha:	no specific restriction	.009		

There is **no effect** of learning condition or frequency → There may be an L1 effect!

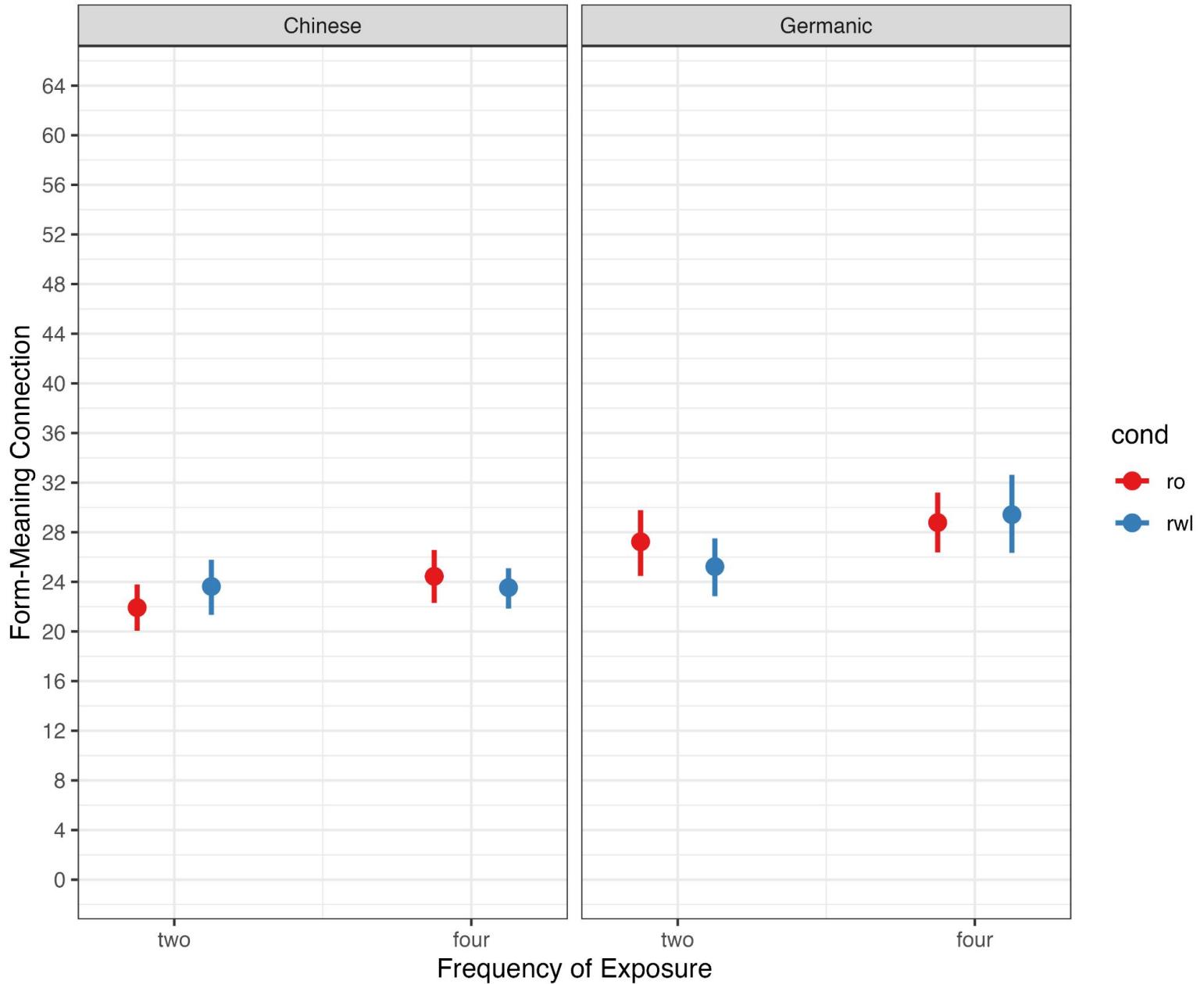
# Replicating Malone (2018) - Working memory (RQ1c)



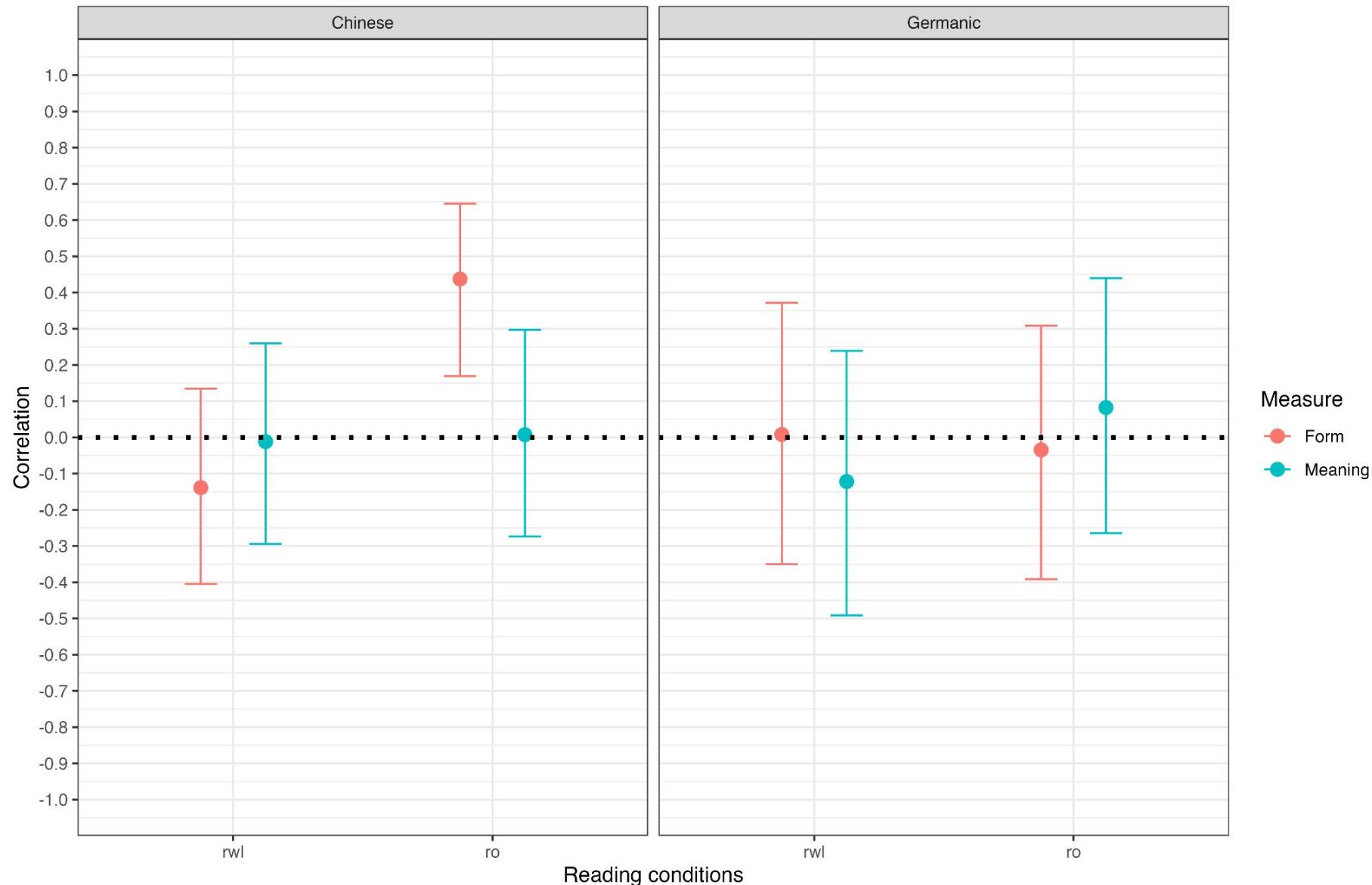
## RQ2a: L1 effects (form recognition)



# RQ2b: L1 effects (form-meaning connection)



# RQ2c: Working memory



# **Discussion**

# Did the findings replicate?

## Malone (2018)

- Effect of aural enhancement:
  - Form recognition:
    - AE effects for 2 exposures
  - Form-meaning:
    - AE effects for 2 & 4 exposures
- number of exposures
  - All groups gained knowledge
  - 4 exposures led to more knowledge
- working memory capacity
  - Lower WM did not cancel out AE benefits

## Our Results

- Effect of aural enhancement:
  - Form recognition and Form-meaning:
    - no clear advantage
- number of exposures
  - effects not observed
- working memory capacity
  - generally consistent with Malone

# Is there an L1 effect?

- Form recognition and form-meaning connections:  
Not really; both groups showed similar patterns
- Working memory:  
Some potential effects - Chinese groups showed different patterns in the RO condition

# It is not clear *why*

In terms of the effects of aural enhancement and no. of exposures

- Data collection is still ongoing
  - large variability in reading proficiency in rwI4
- Differences in L1 background, coupled with learning contexts (ESL vs. EFL)
- Online data collection (?) (but our sample also had a high reading comprehension score)
- Audio pace (too slow for proficient learners?, e.g., Conklin et al. 2020)
- No. of exposures not different enough (2 vs. 4)?

# **It is not clear *why***

In terms of the psychological mechanism behind vocab learning through RWL

- strengthen spoken and written form connections
- speakers of both language families performed similarly

# Thank you

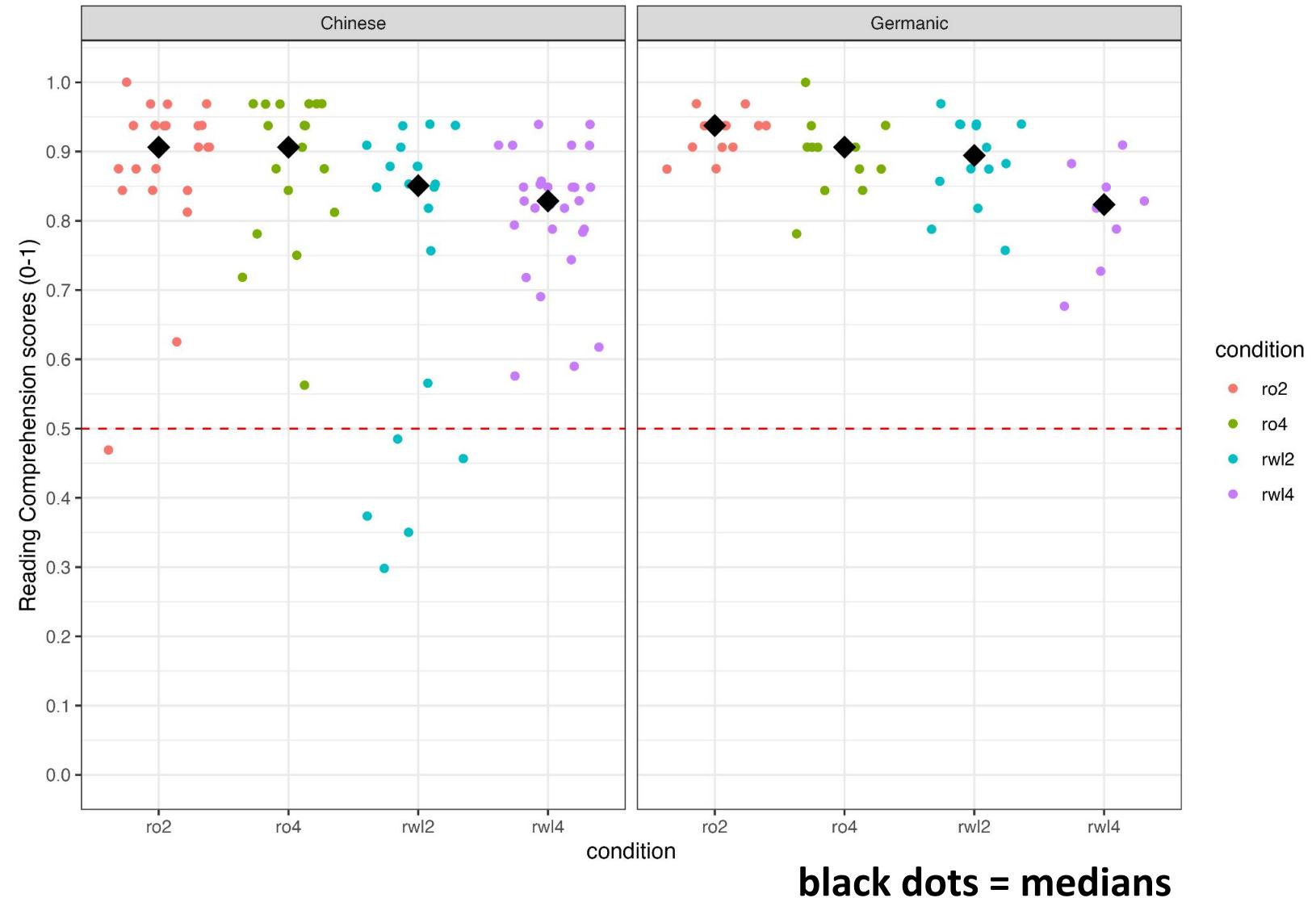
More details:

- OSF project page:  
<https://osf.io/vntra/>
- SSLA Replication Special Issue:  
<https://www.cambridge.org/core/journals/studies-in-second-language-acquisition/call-for-proposalssla>

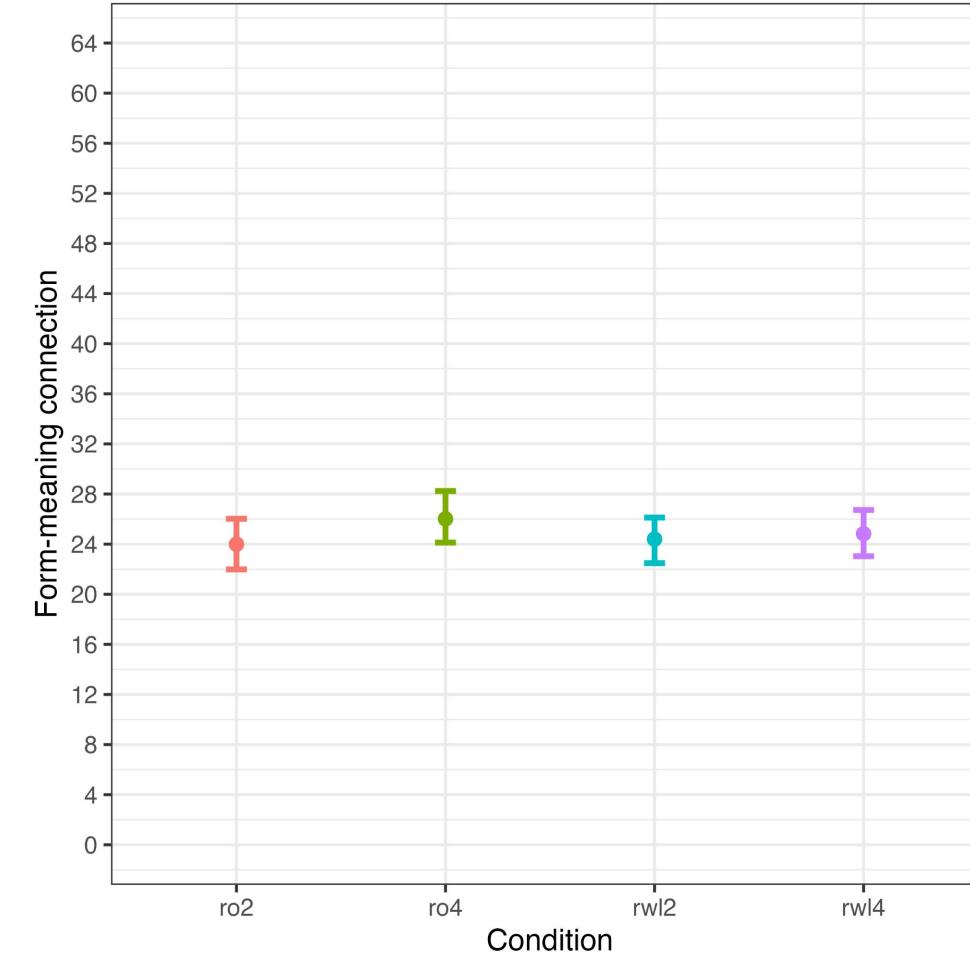
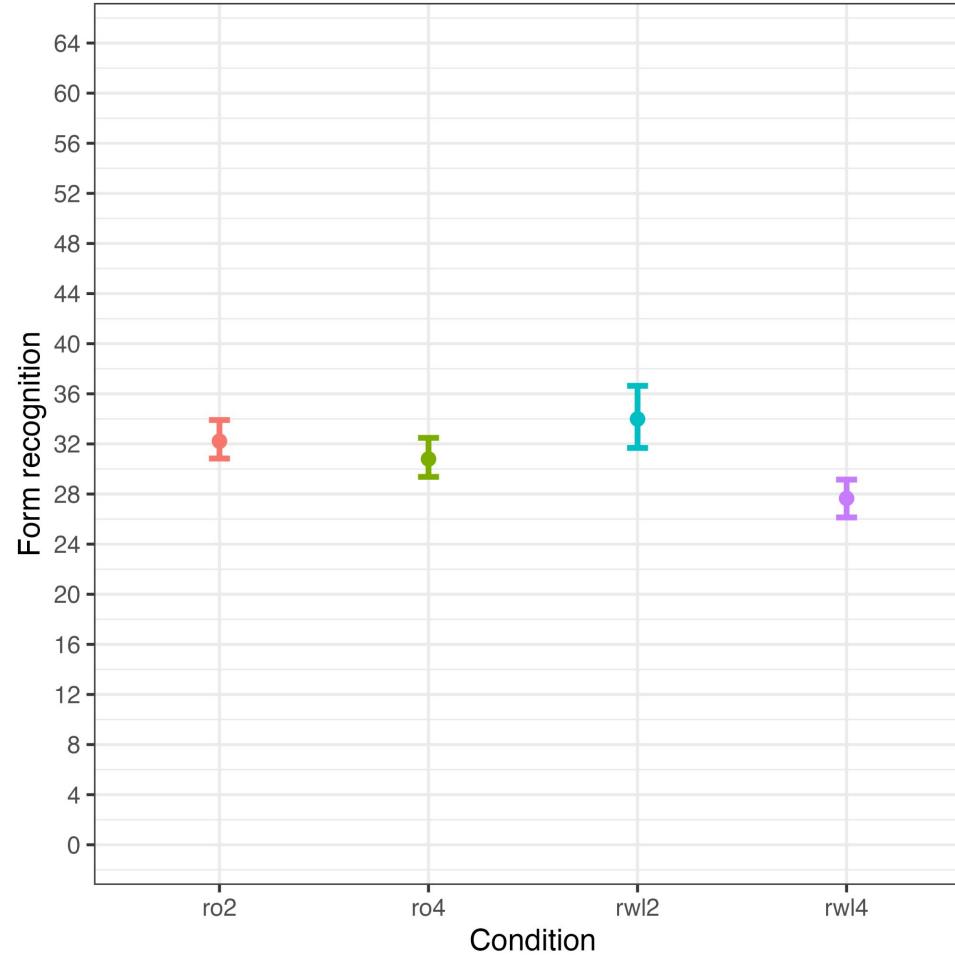


# Supplementary slides

# Reading Comprehension



# Descriptive Statistics (learning scores)



# Analysis: RQ1

## Correlation

- Between form recognition or form-meaning connection scores and working memory scores
- Bayesian approach to compute **the probability of correlation being different from zero.**

## Malone (2018)

### Form recognition

- reading while listening  
(four)
- reading while listening  
(two)
- reading only (four)



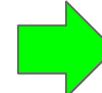
### Form-meaning connection

- reading while listening  
(four)
- reading while listening  
(two)
- reading only (four)

# Analysis: RQ 2

## Regression Modeling

- L1 (Chinese vs. Germanic)
- Reading mode (reading-while listening vs. reading only)
- Frequency (Four vs. Two)
- Working memory (composite scores)
- Reading comprehension (cloze test)



- brms package (Bürkner, 2017)
- Bayesian models
  - Interpret the results based on
    - estimated means of each group
    - estimated effects of WM

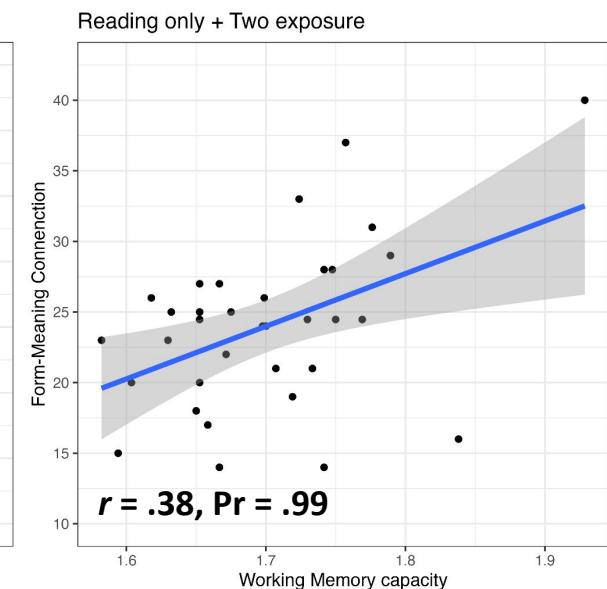
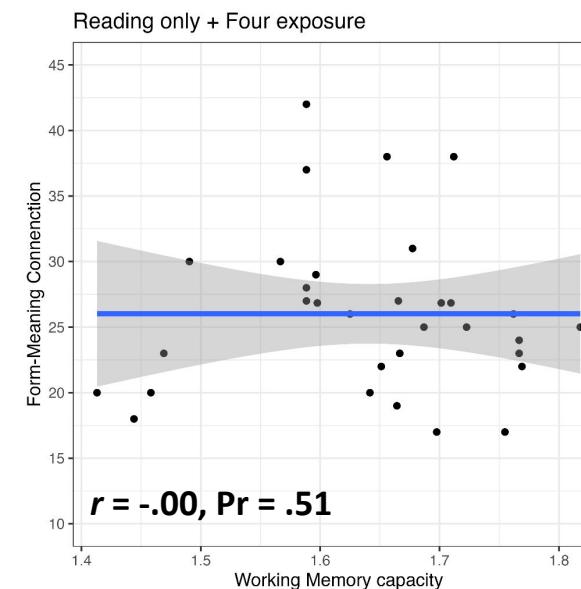
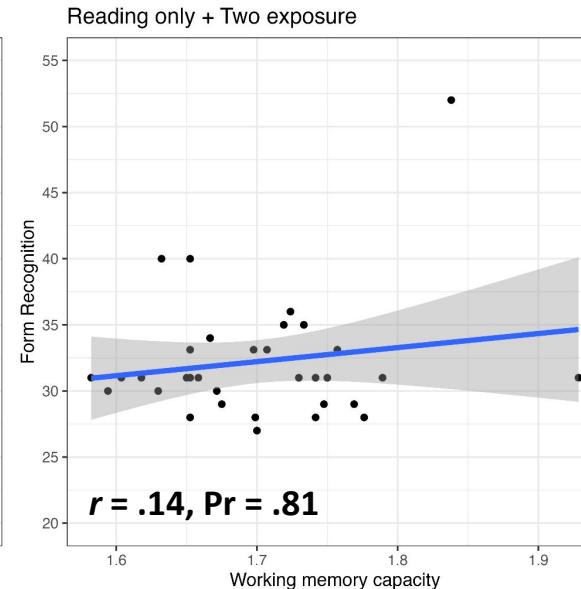
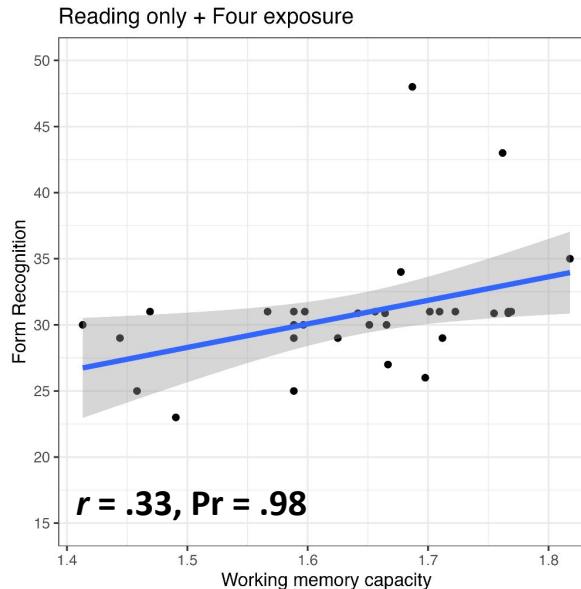
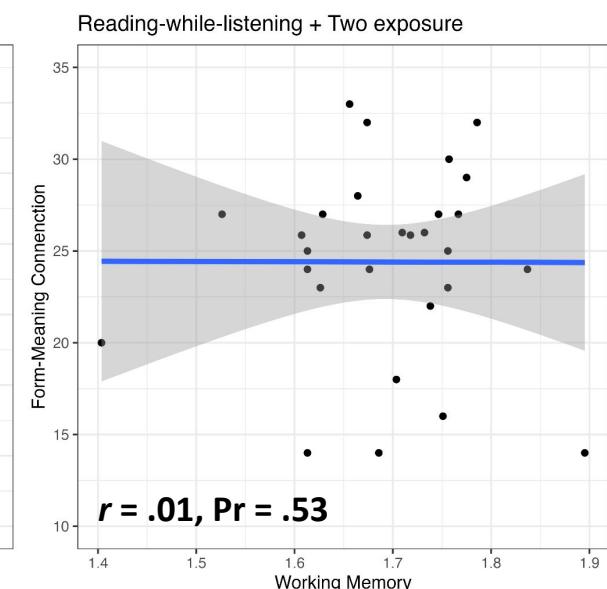
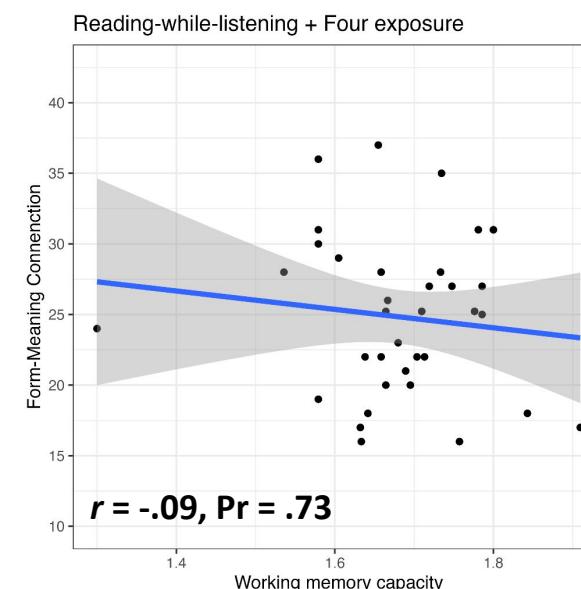
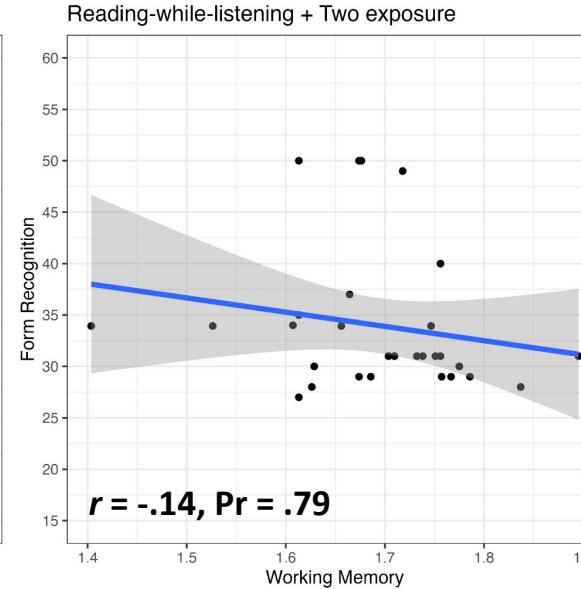
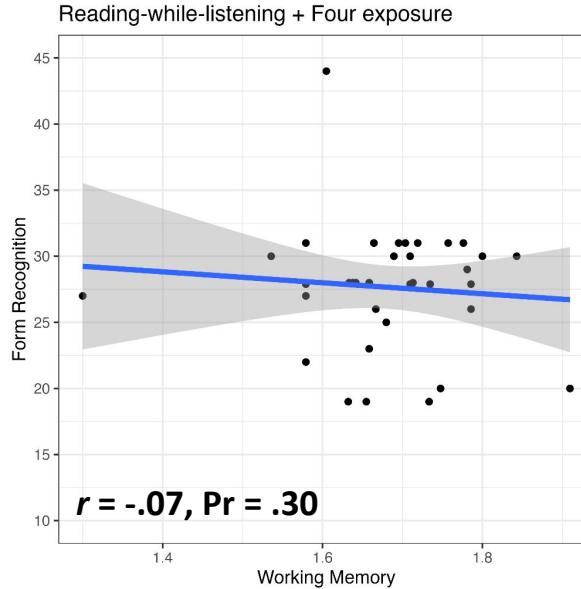
learning ~ L1 + cond + freq + wm + cloze +  
L1:freq + L1:cond + freq:cond + L1:wm + cond:wm  
L1:freq:cond + L1:cond:wm



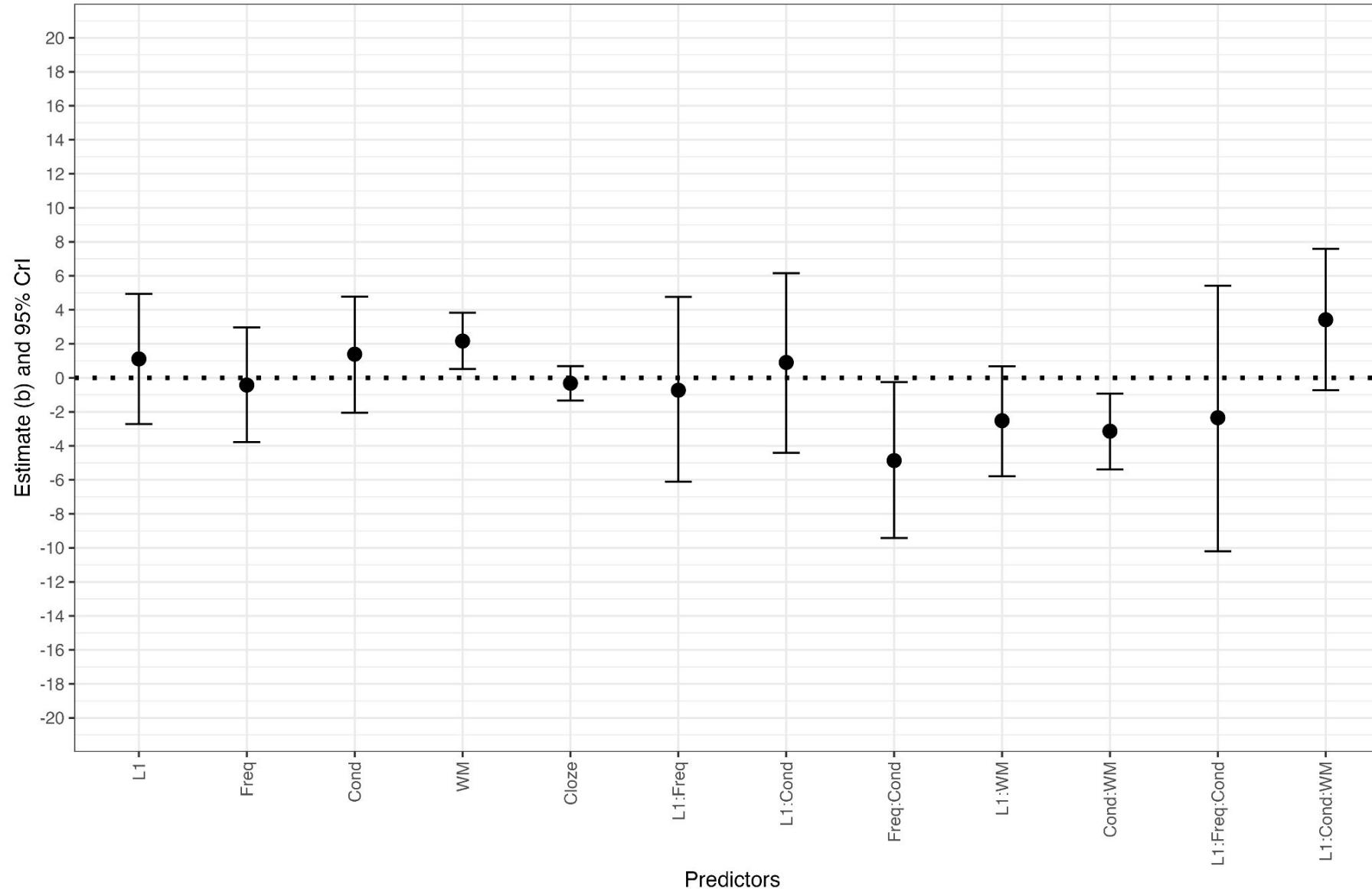
# Replicating Malone (2018) - Working memory

Supplementary

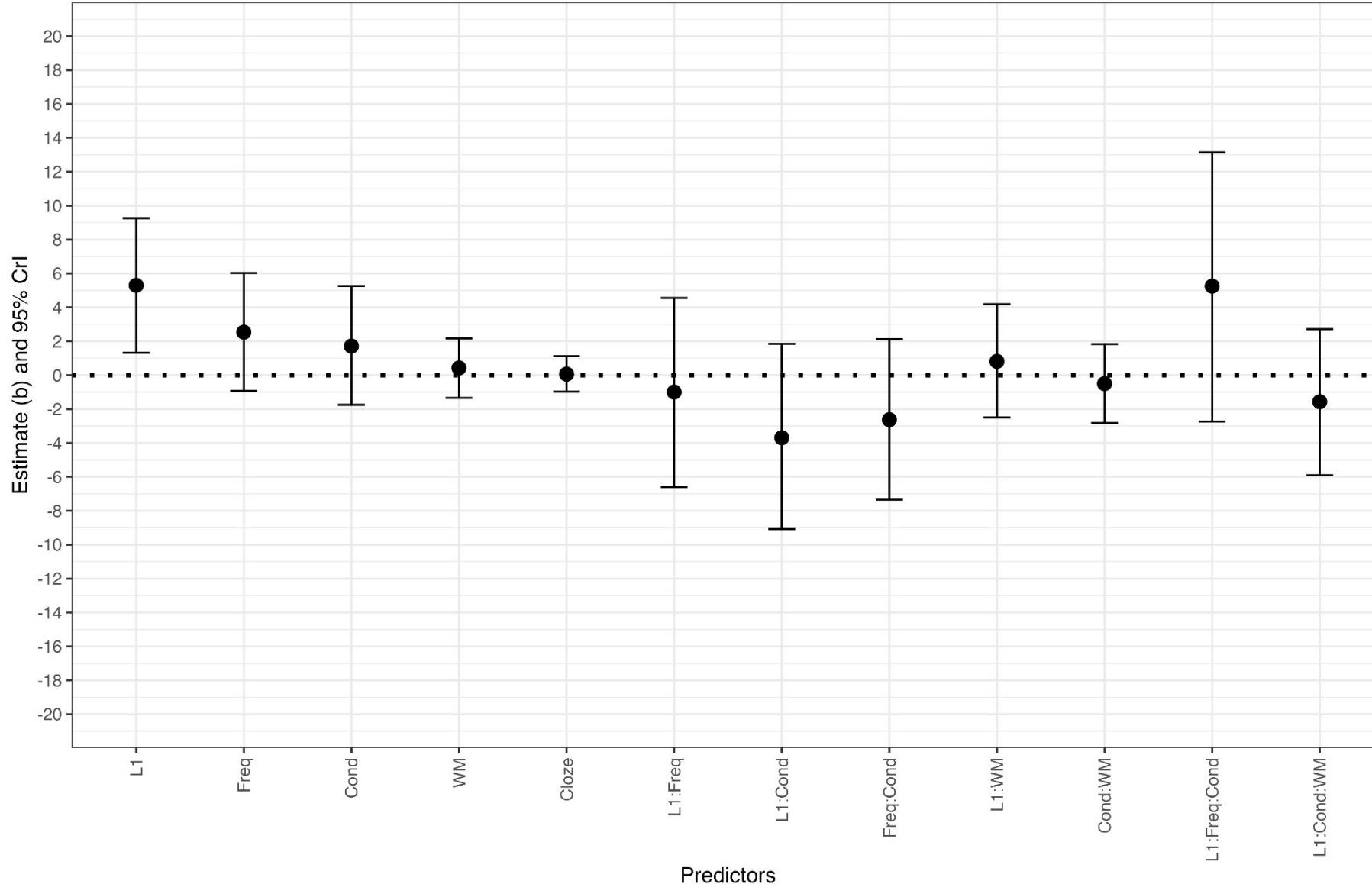
## Form recognition



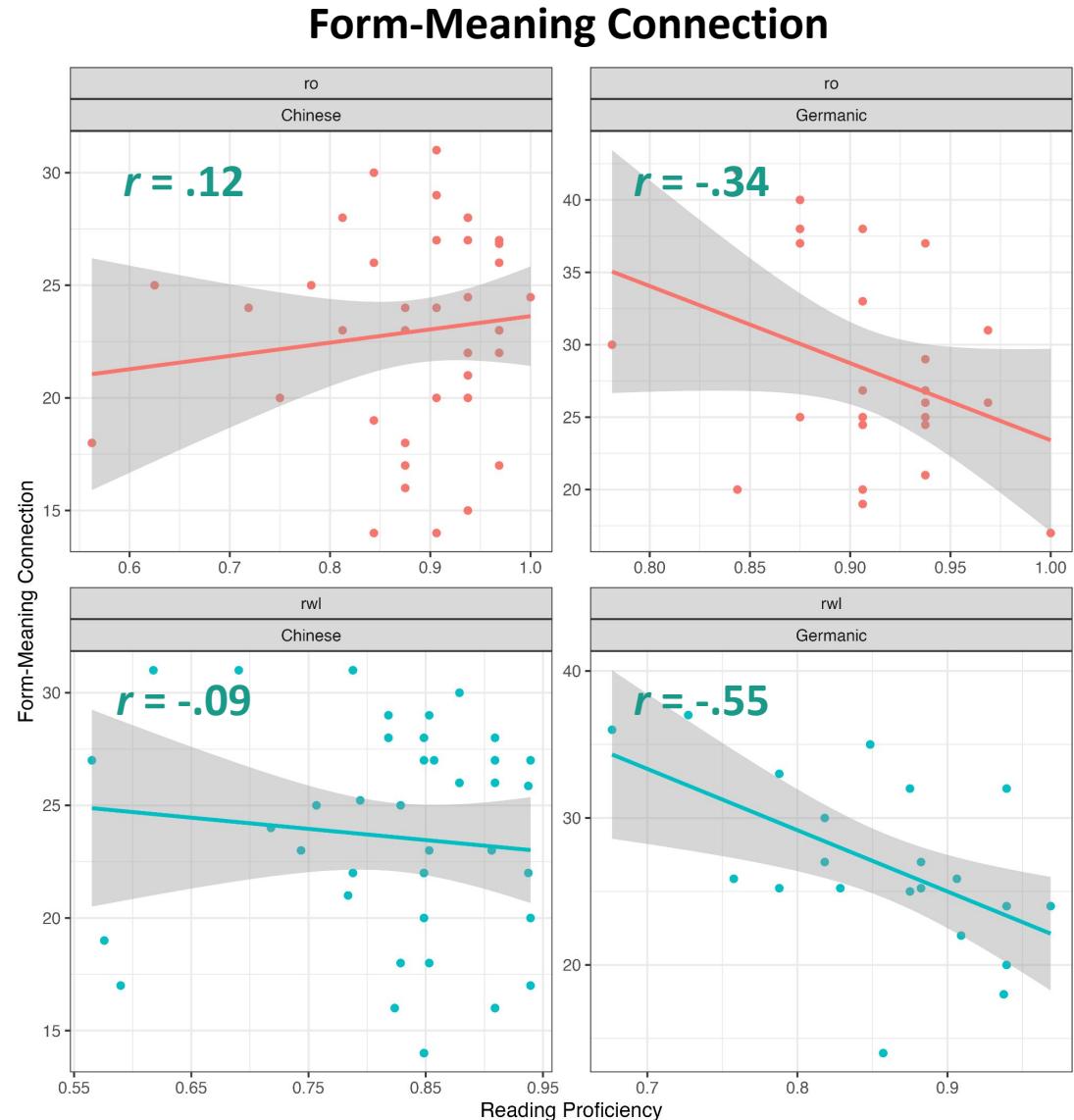
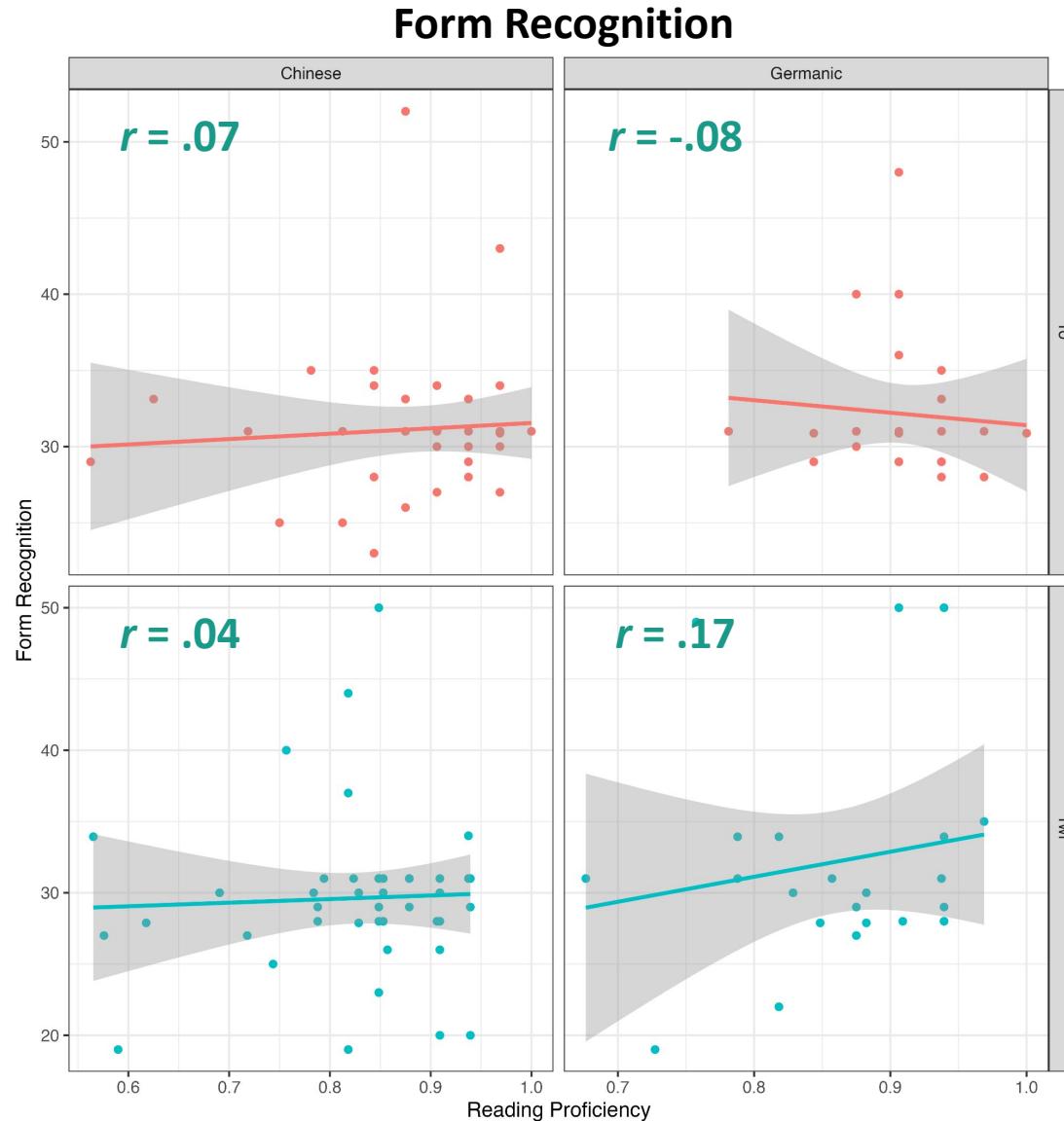
# RQ2: Model estimates (form recognition)



# RQ2: Model estimates (form-meaning connection)



# L1, Condition, Proficiency



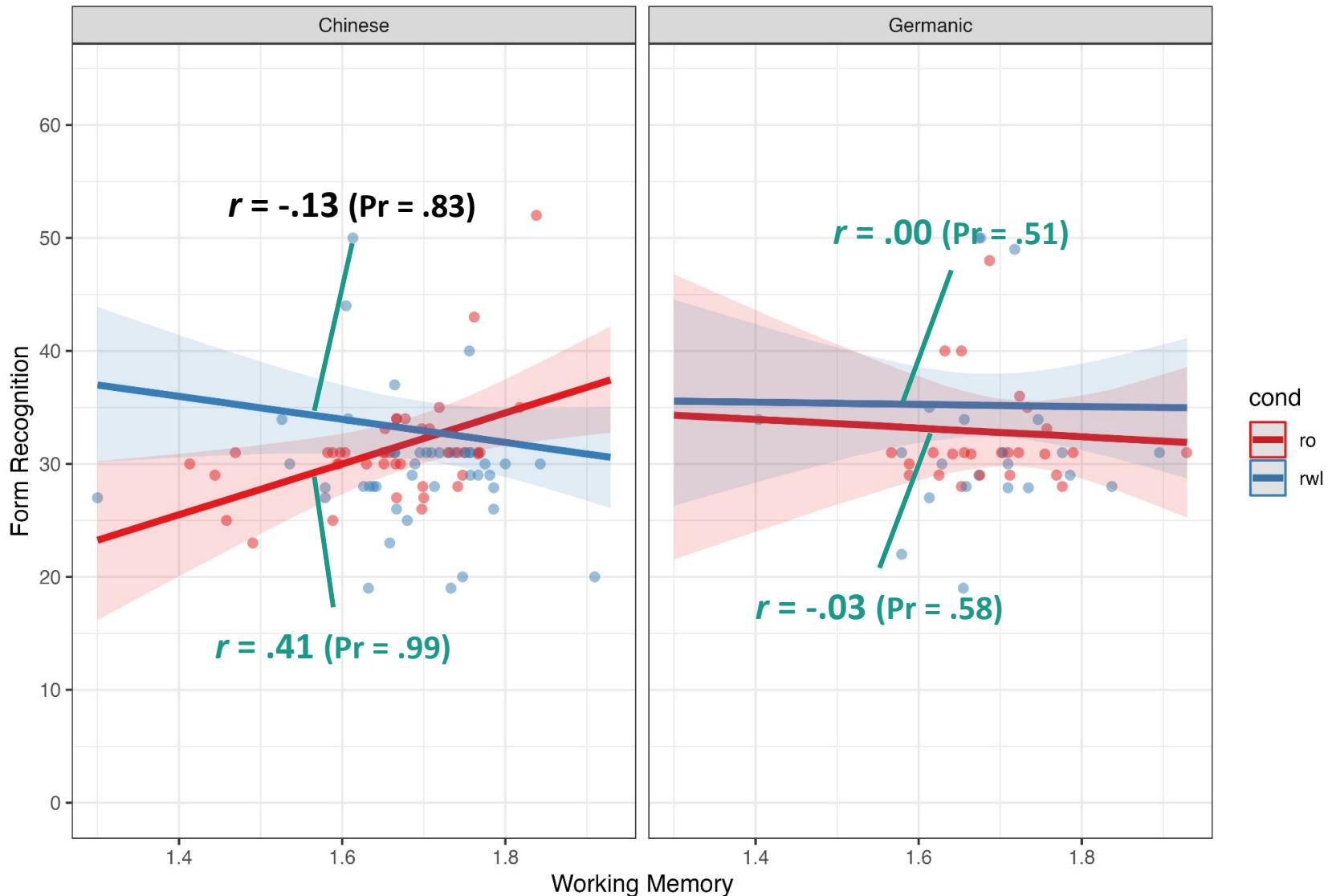
# Correlation Matrix (Chinese)



# Correlation Matrix (Germanic)



# RQ2c: Working memory (form recognition)



# RQ2c: Working memory (form-meaning connection)

