Translation equivalents are not special in bilingual infant vocabulary development: Evidence from **a quantitative model**

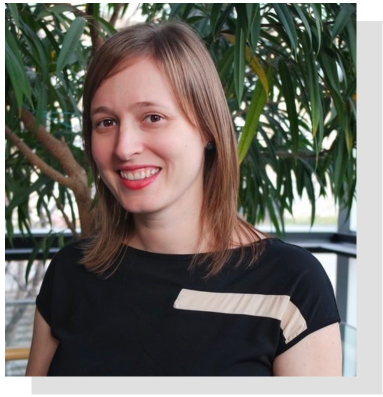
Université Concordia University

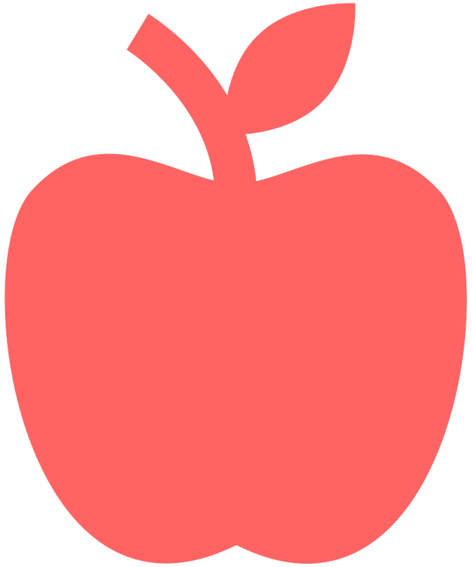
Concordia Infant Research Laboratory
Laboratoire de recherche sur l'enfance de Concordia


Rachel Ka-Ying Tsui


Ana Maria Gonzalez-Barrero


Esther Schott


Krista Byers-Heinlein

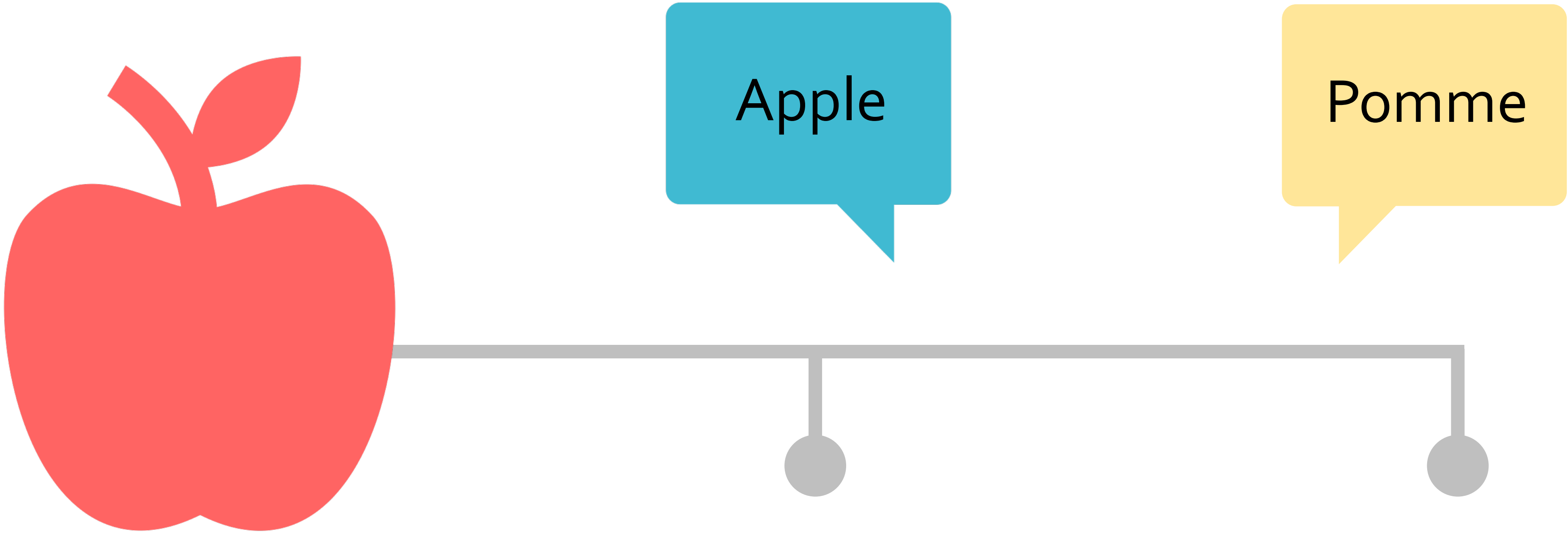


Apple

Pomme

Translation equivalents:

2 labels for the same concept

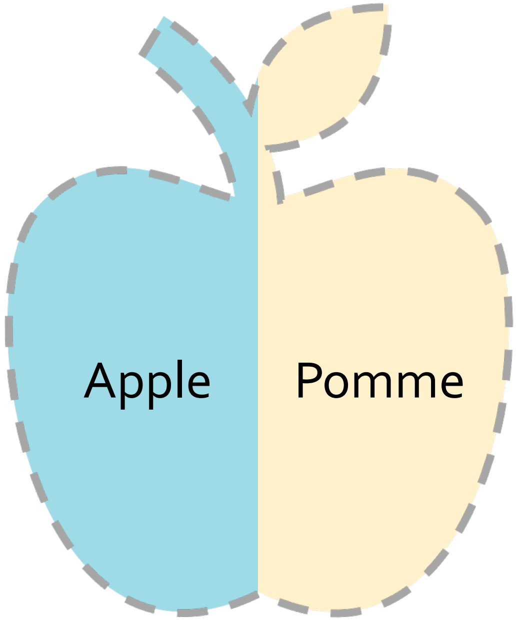


Singlet:
A first label for a concept

Translation equivalents

Translation equivalents are special…

- Learned differently from singlets
- Strong semantic overlap



3 competing theories:
How are translation equivalents learned?

Account #1

Bilingual children **reject** translation equivalents in favour of learning one label for each referent
(Volterra & Taeschner, 1978)

Avoidance Account

Account #2

Bilingual children **favour** learning translation equivalents
(Bilson et al., 2015; Floccia et al., 2020)

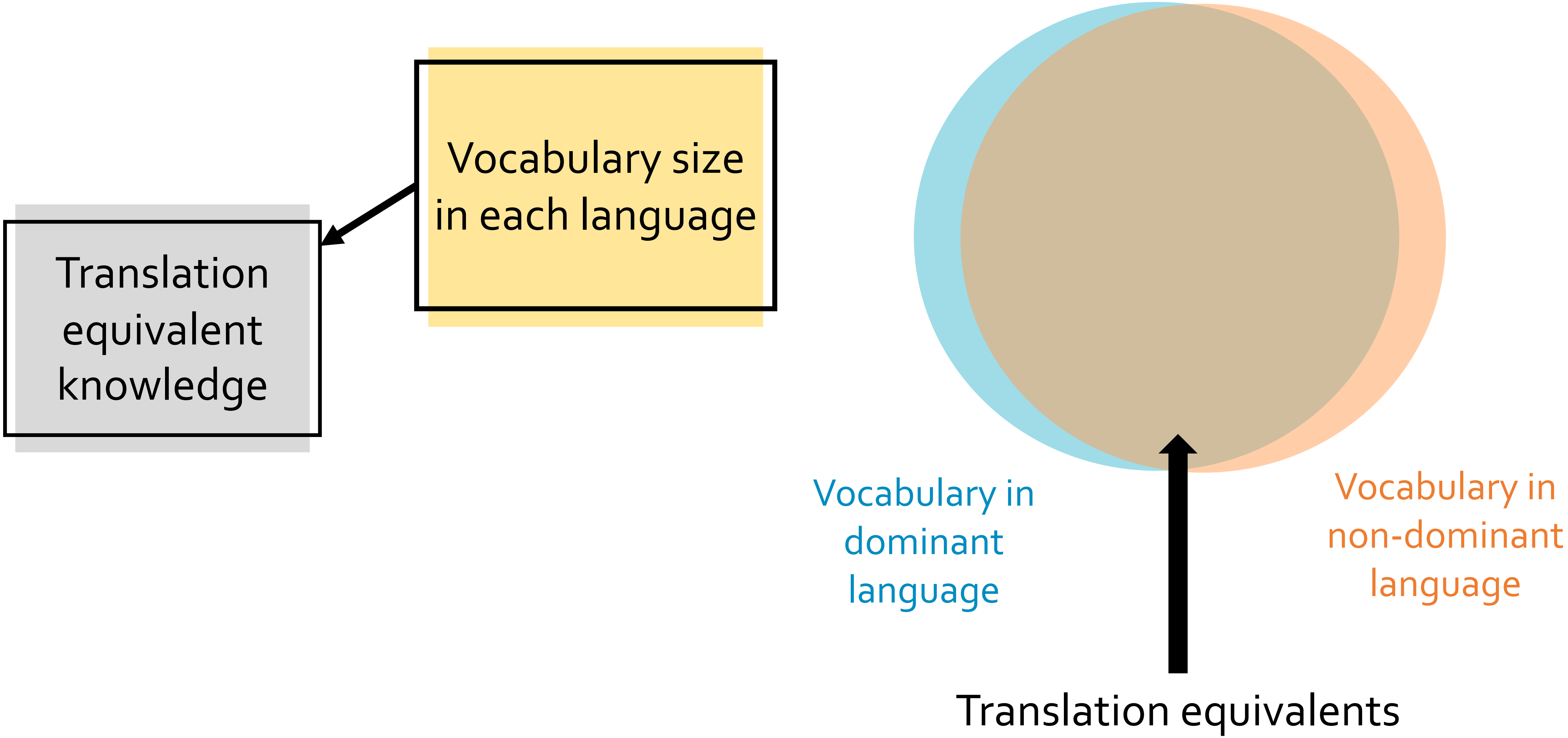
Preference Account

Account #3

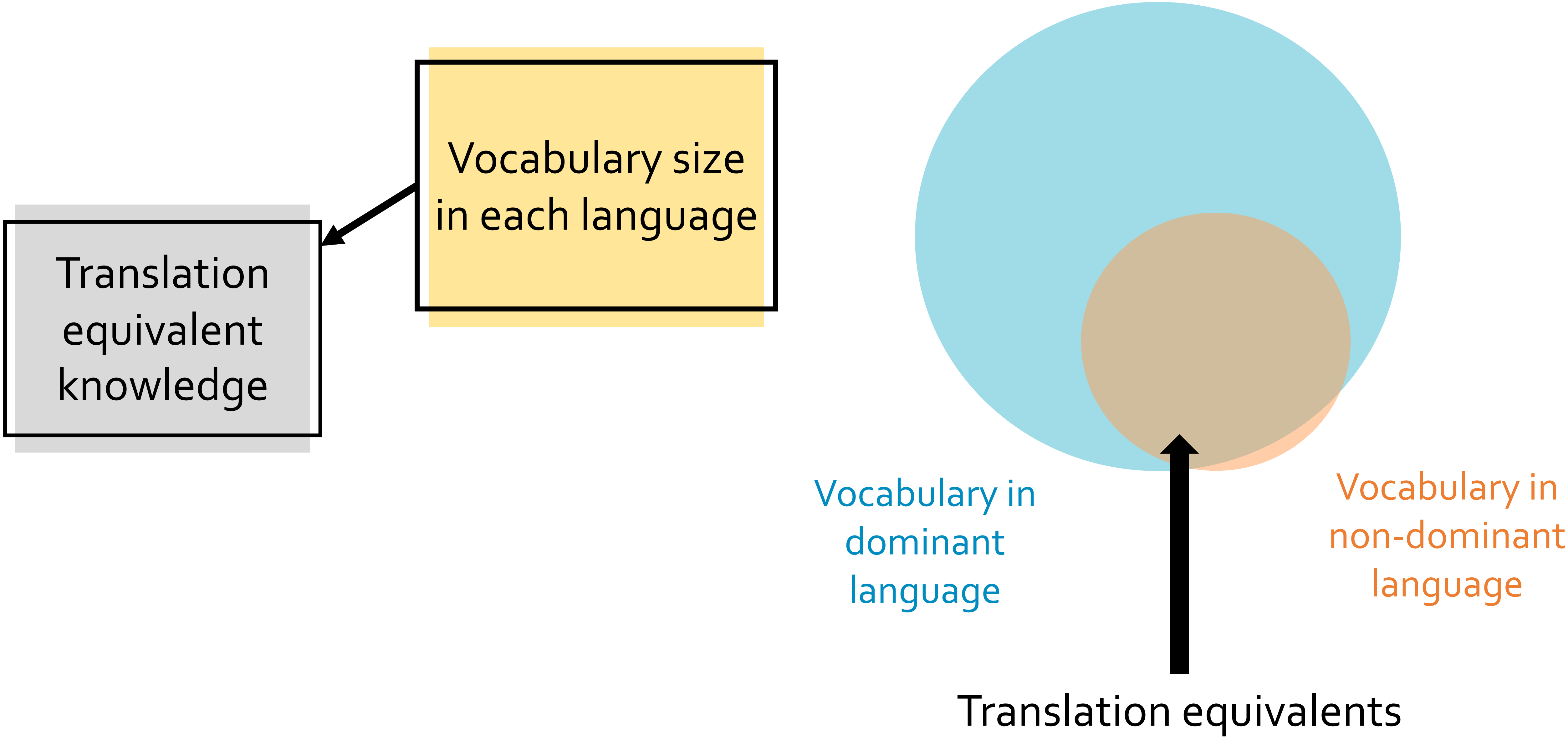
Bilingual children learn translation equivalents and singlets **in a similar way**
(Pearson et al., 1995)

Neutral Account

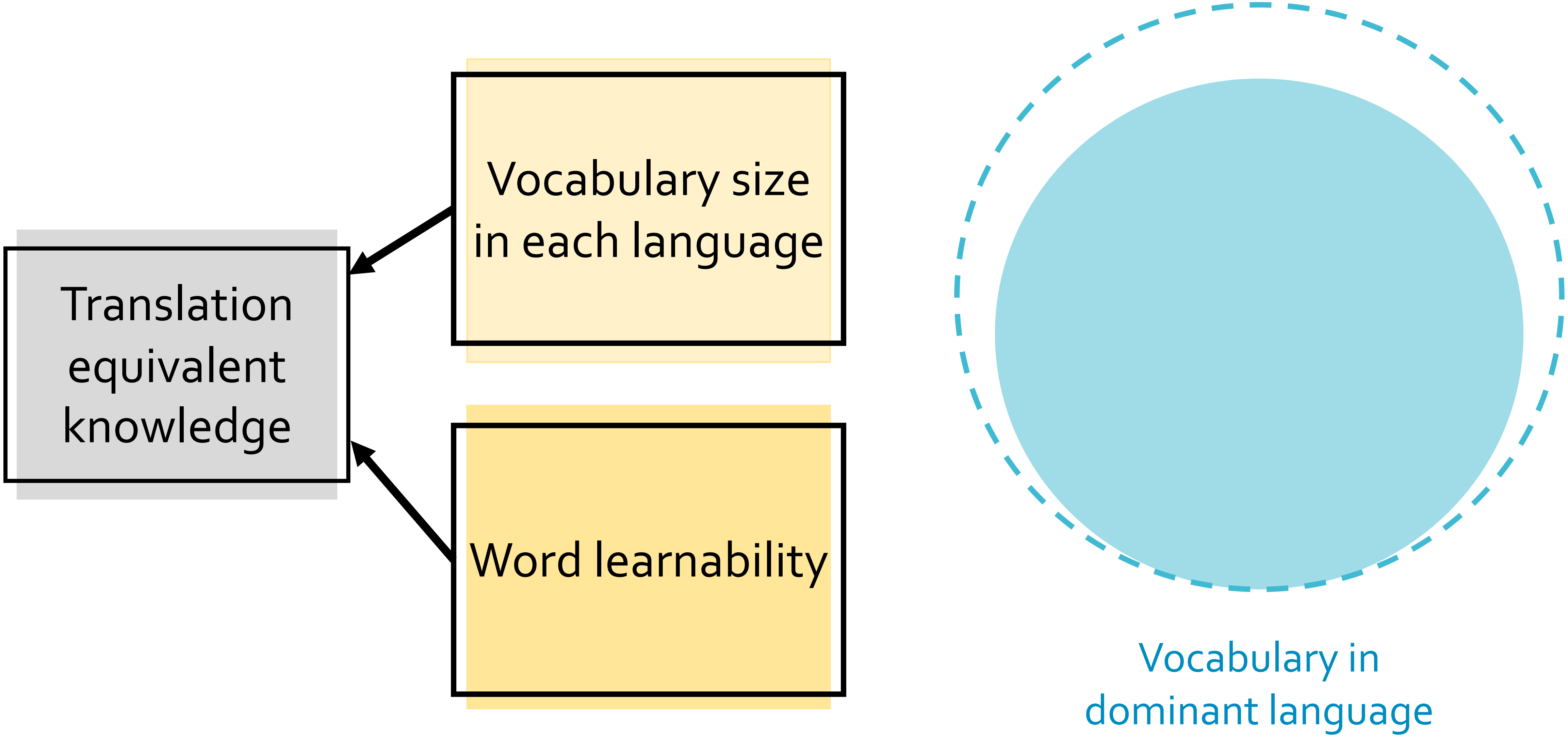
Contributors to translation equivalent knowledge



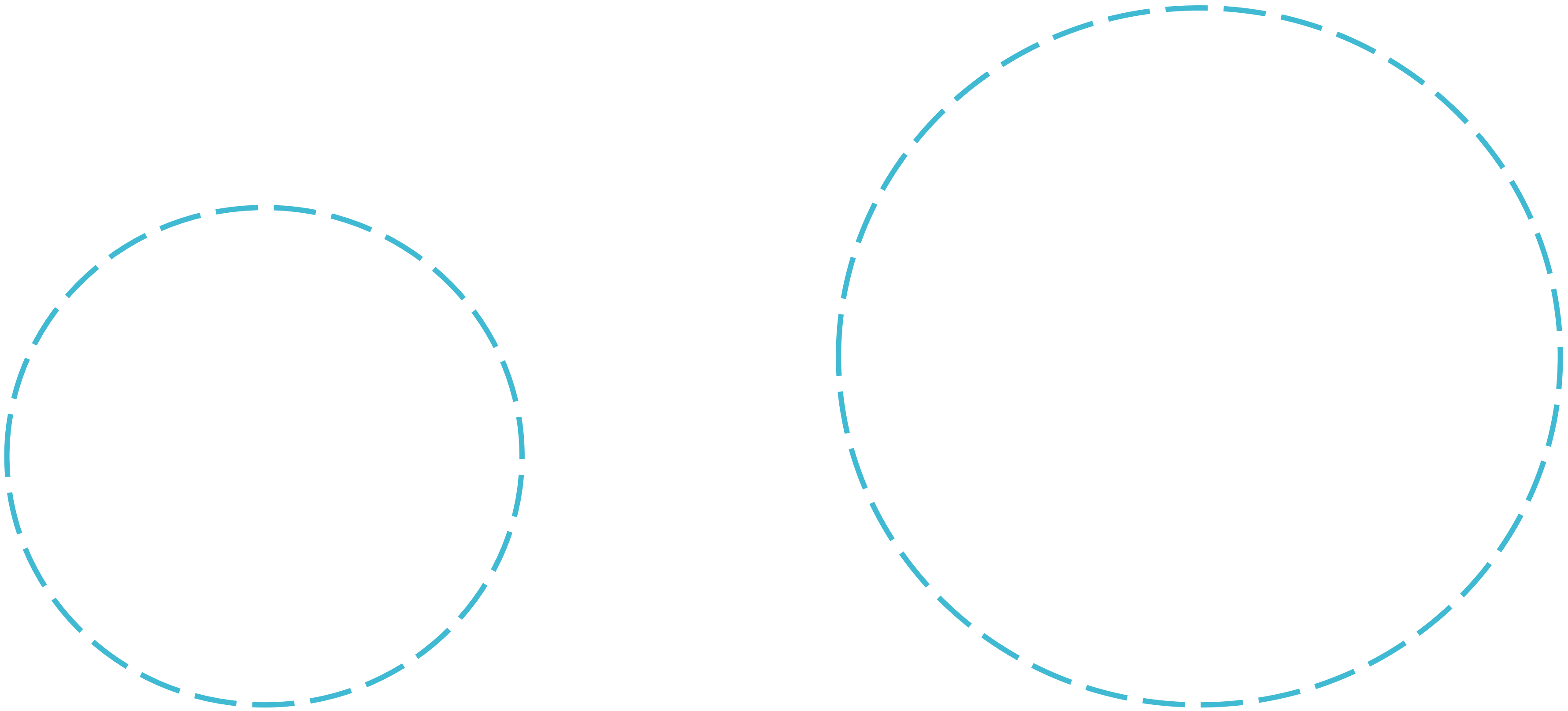
Contributors to translation equivalent knowledge



Contributors to translation equivalent knowledge



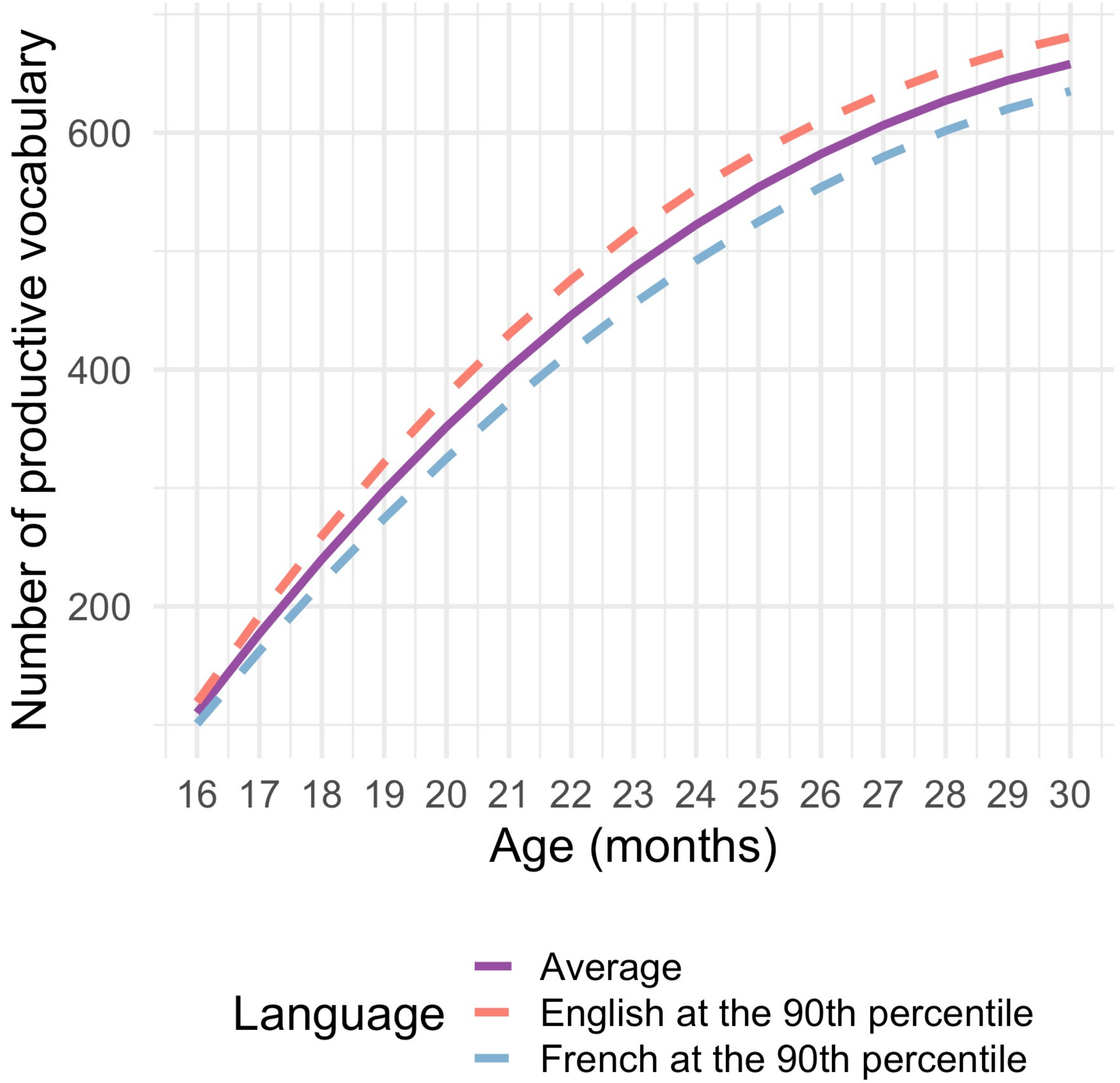
Contributors to translation equivalent knowledge



Vocabulary size of
a 18-month-old

Vocabulary size of
a 30-month-old

Number of CDI words produced at the 90th percentile



Our study

- Translation equivalent knowledge as a function of bilinguals' own vocabulary size in each language
- What is the nature of translation equivalent learning in bilingual children?

Avoidance Account

Translation equivalents are harder to learn than singlets

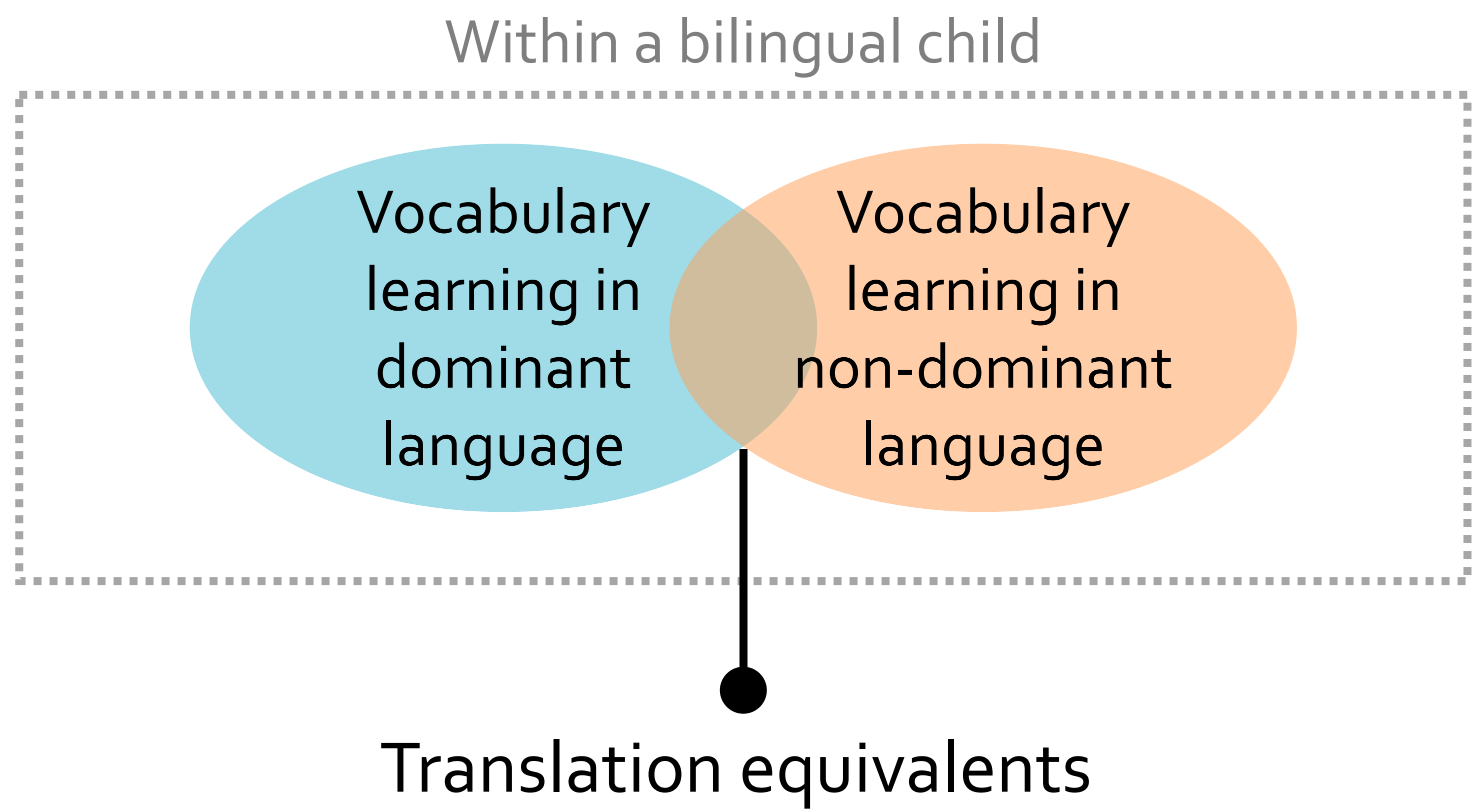
Preference Account

Translation equivalents are easier to learn than singlets

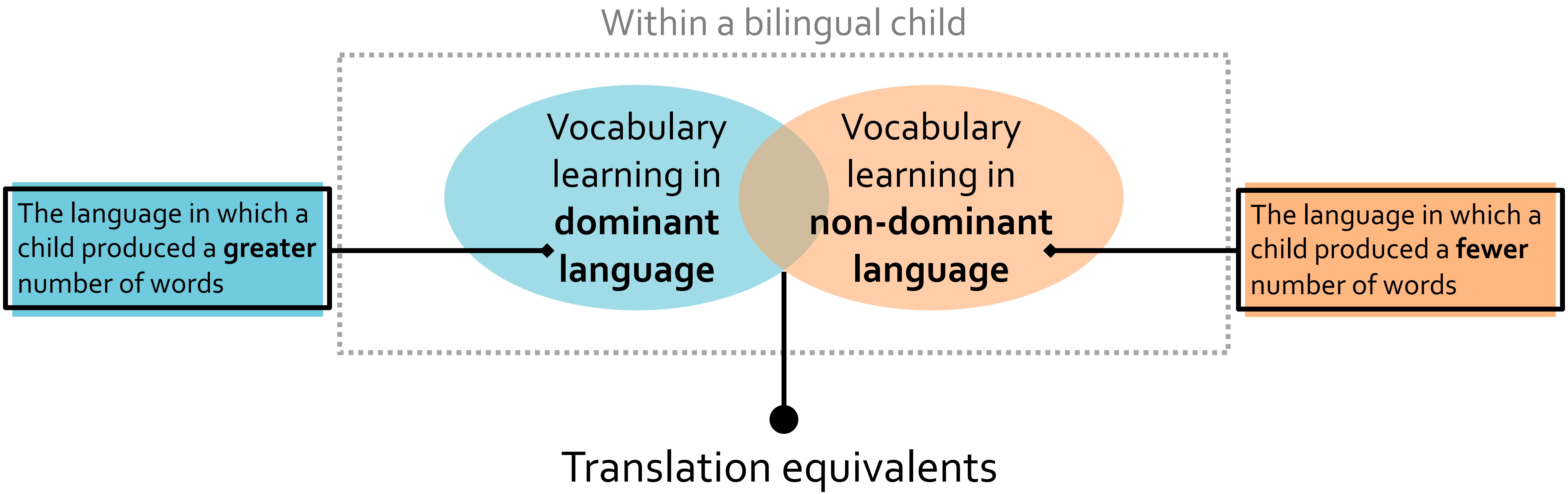
Neutral Account

Translation equivalents are similar to learn as singlets

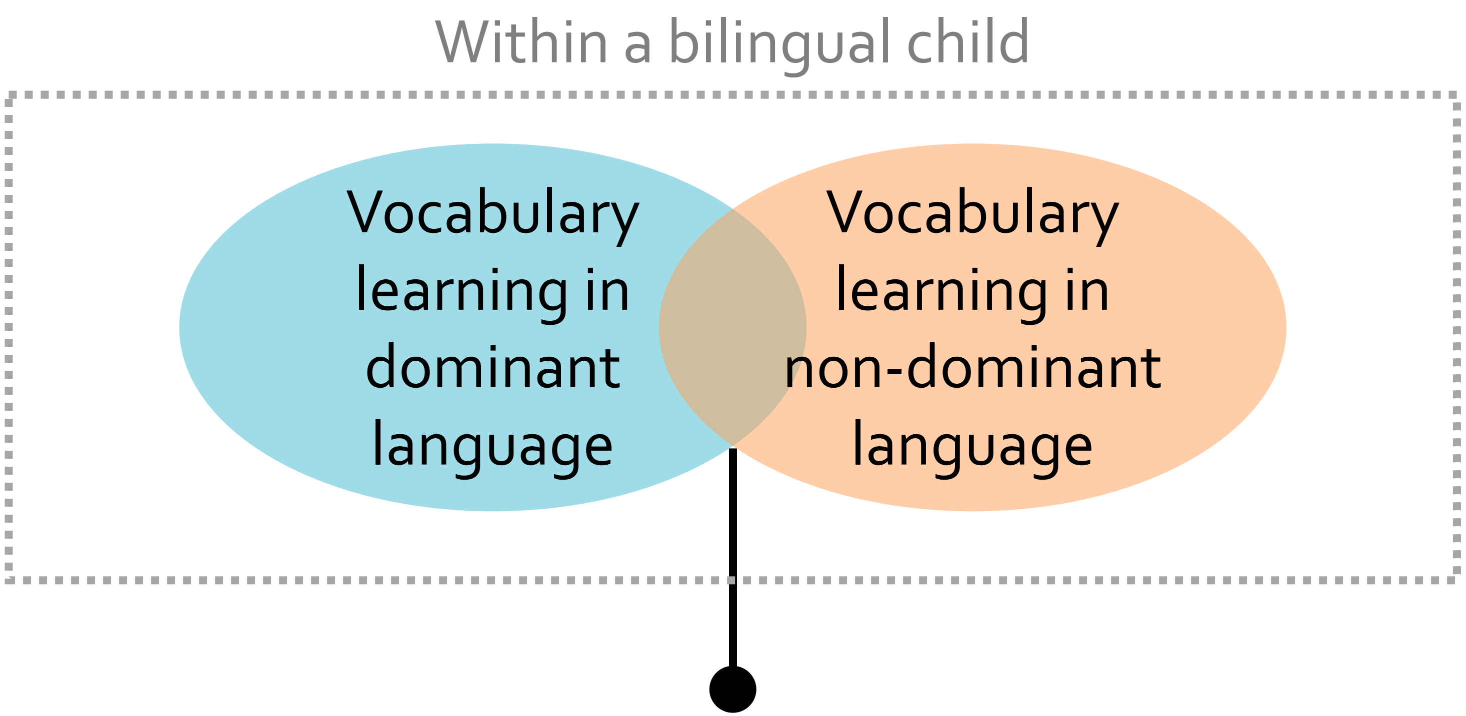
Bilingual Vocabulary Model



Bilingual Vocabulary Model



Bilingual Vocabulary Model

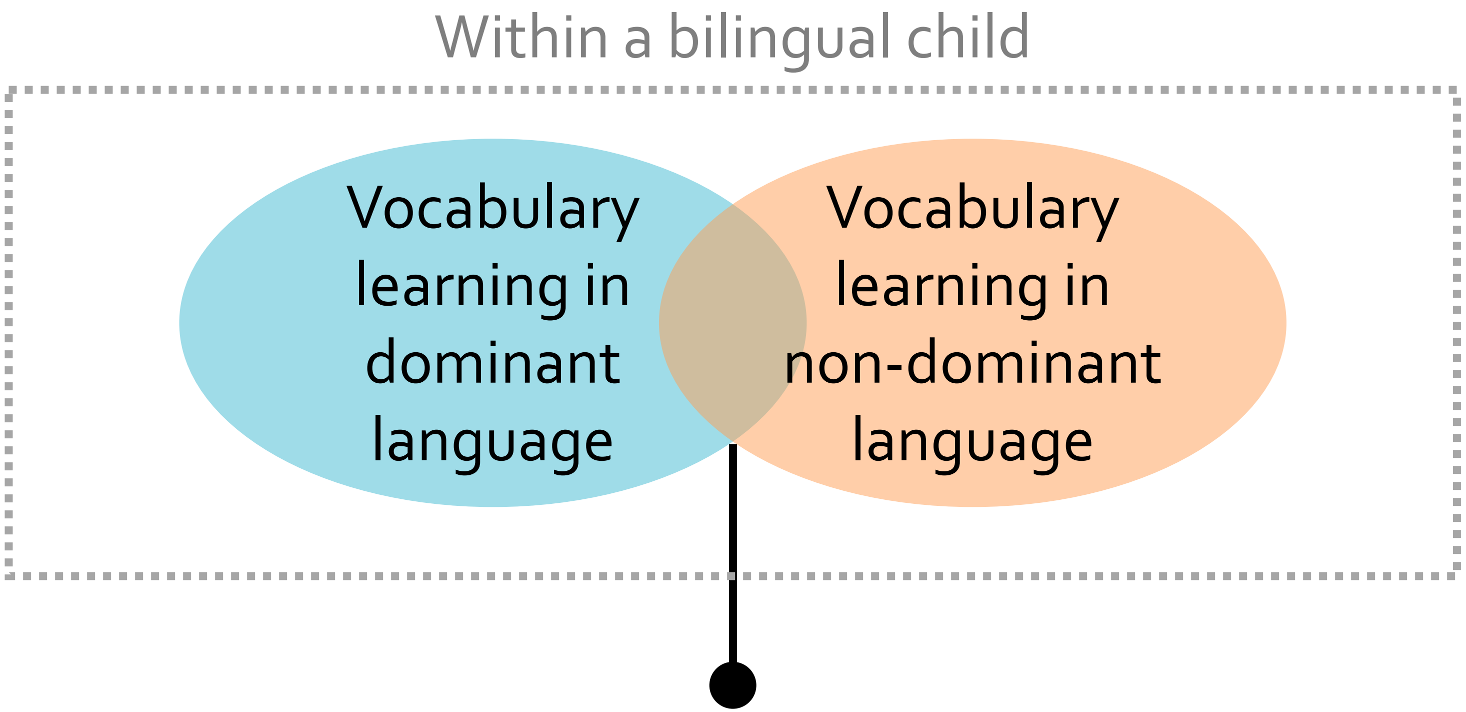


Translation equivalents

=

Joint probability of learning the words in each language

Bilingual Vocabulary Model

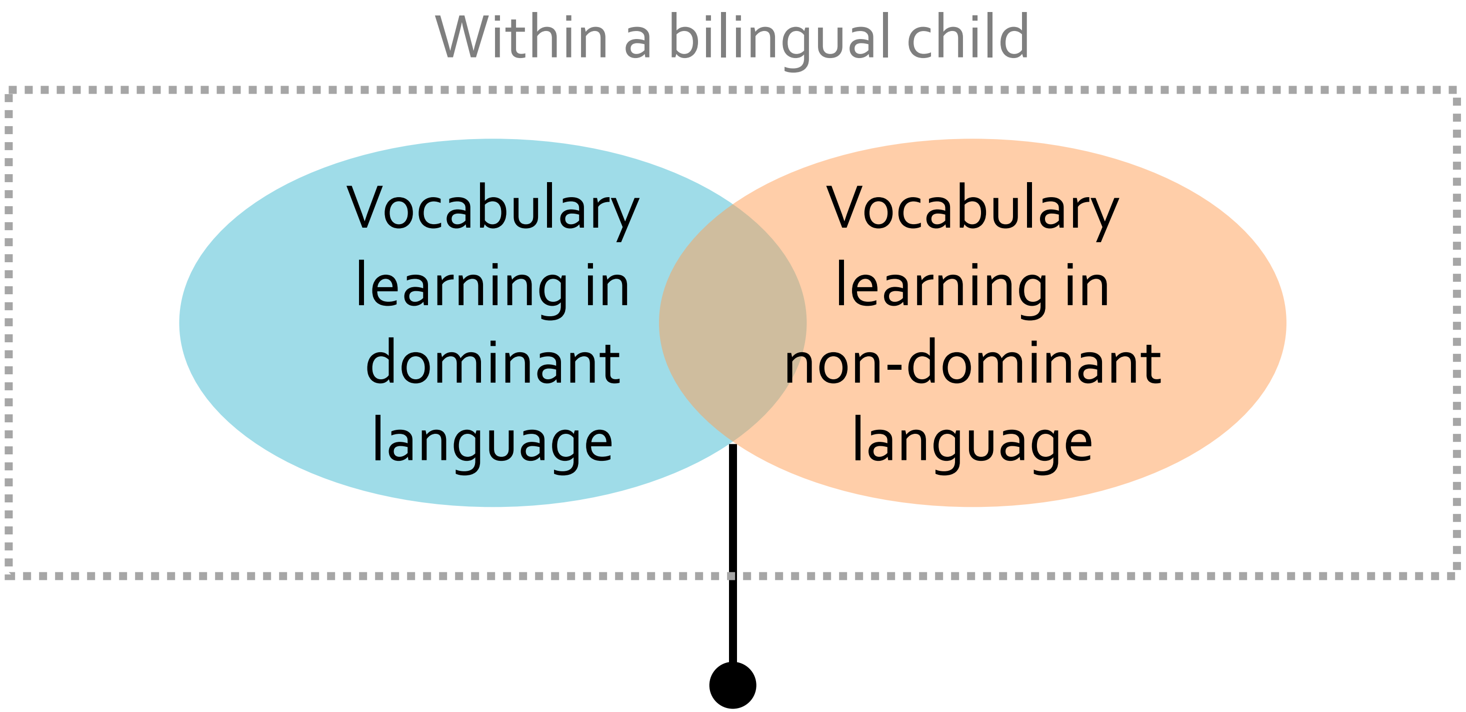


Translation equivalents

=

Independent event: **P(A and B) = P(A) × P(B)**

Bilingual Vocabulary Model



Translation equivalents

=

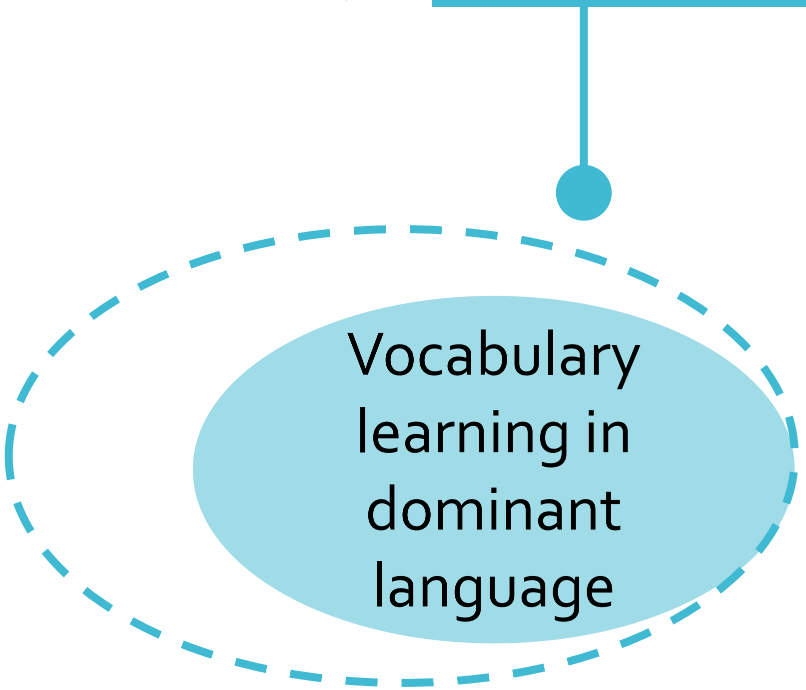
P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)



$$\frac{\text{Number of dominant vocabulary }\textbf{known}}{\text{Number of }\textbf{learnable}\text{ vocabulary}}$$

Bilingual Vocabulary Model

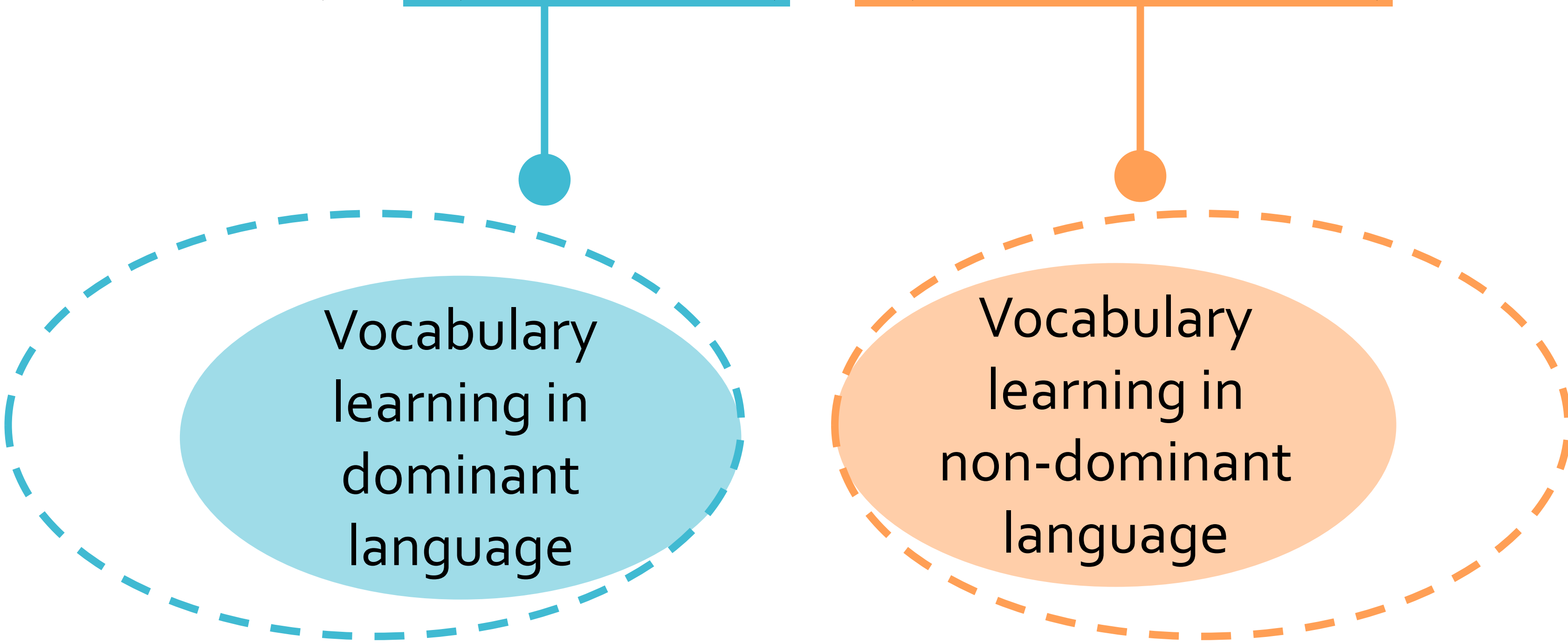
P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

Vocabulary learning in non-dominant language

$$\frac{\text{Number of non–dominant vocabulary } \textbf{known}}{\text{Number of } \textbf{learnable} \text{ vocabulary}}$$

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

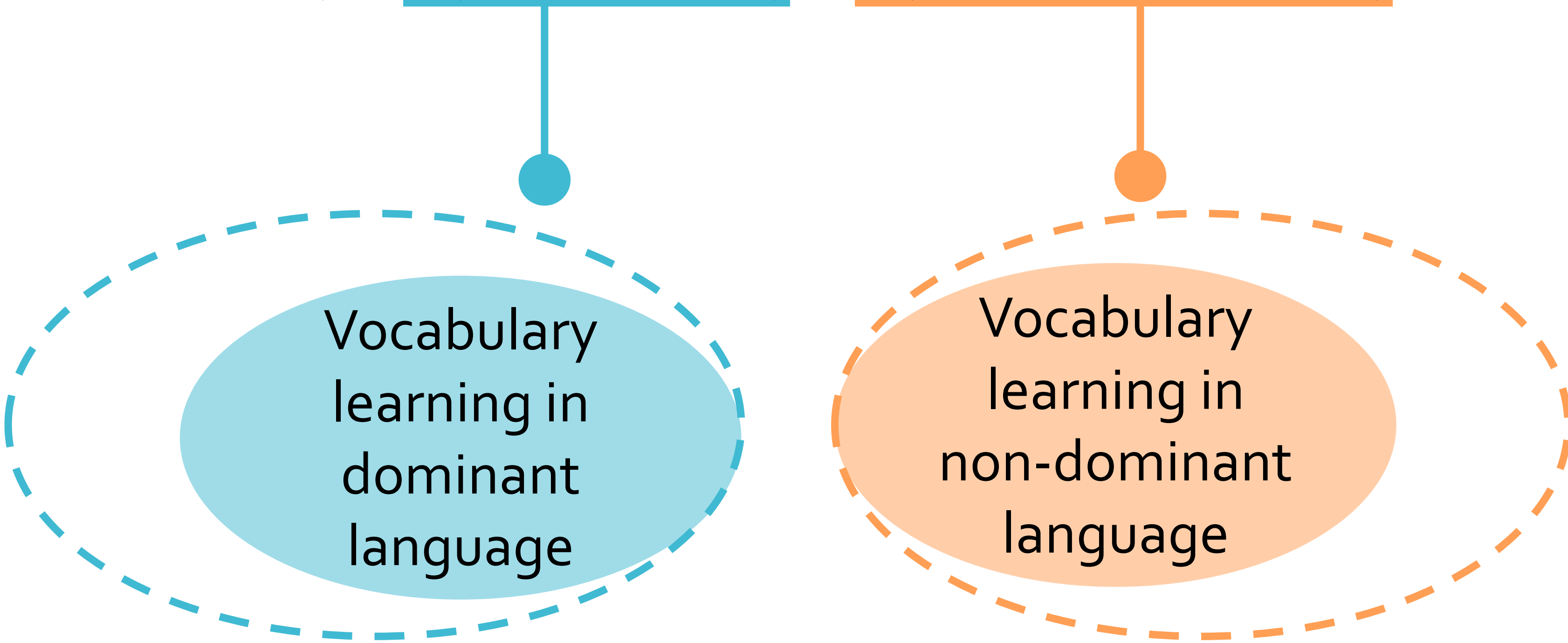


90th percentile for a 21-month-old child

$$\frac{\text{300 dominant vocabulary known}}{\text{400 learnable vocabulary}}$$

Bilingual Vocabulary Model

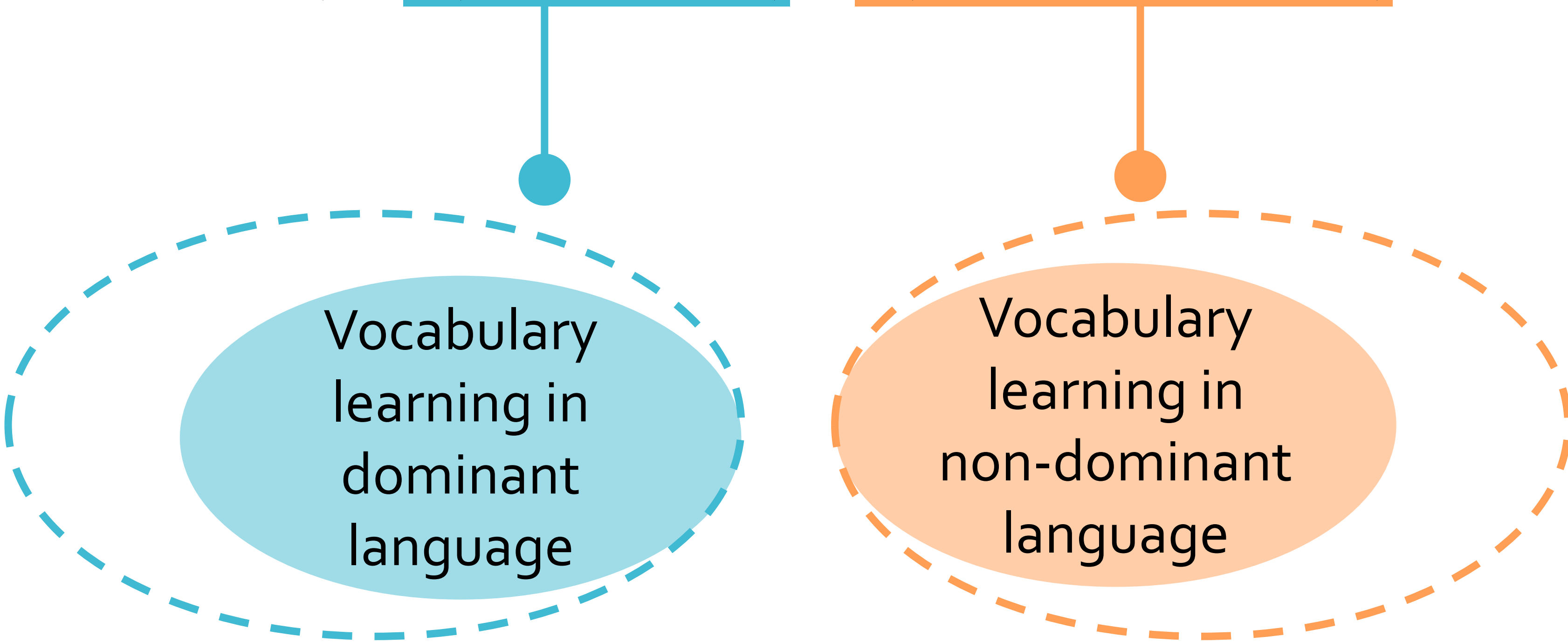
P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)



$$P(\text{Dominant}) = \frac{300}{400}$$

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)



$$P(\text{Dominant}) = \frac{300}{400} \qquad P(\text{Non-Dominant}) = \frac{100}{400}$$

Bilingual Vocabulary Model

$$P(\text{Dominant and Non-Dominant}) = P(\text{Dominant}) \times P(\text{Non-Dominant})$$

$$\frac{300}{400} \qquad \frac{100}{400}$$

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

$$P(\text{Dominant and Non-Dominant}) = \frac{300}{400} \times \frac{100}{400}$$

Expected(Dominant and Non-Dominant) =

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

$$\frac{300}{400} \qquad \frac{100}{400}$$

Expected number of translation equivalents =

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

$$\frac{300}{400} \qquad \frac{100}{400}$$

Expected number of translation equivalents =

P(Dominant and Non-Dominant) × Number of learnable vocabulary

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

$$\frac{300}{400} \qquad \frac{100}{400}$$

Expected number of translation equivalents =

P(Dominant and Non-Dominant) × Number of learnable vocabulary

$$\frac{300}{400} \times \frac{100}{400}$$

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

$$\frac{300}{400} \qquad \frac{100}{400}$$

Expected number of translation equivalents =

P(Dominant and Non-Dominant) × Number of learnable vocabulary

$$\frac{300}{400} \times \frac{100}{400} \times 400$$

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

$$\frac{300}{400} \qquad \frac{100}{400}$$

Expected number of translation equivalents =

$$\frac{300 \times 100}{400}$$

Bilingual Vocabulary Model

P(Dominant and Non-Dominant) = P(Dominant) × P(Non-Dominant)

$$\frac{300}{400} \qquad \frac{100}{400}$$

Expected number of translation equivalents =

No. of dominant vocabulary

No. of non-dominant vocabulary

$$\frac{300 \times 100}{400} = 75$$

No. of learnable vocabulary

To evaluate if translation equivalents are learned independently,

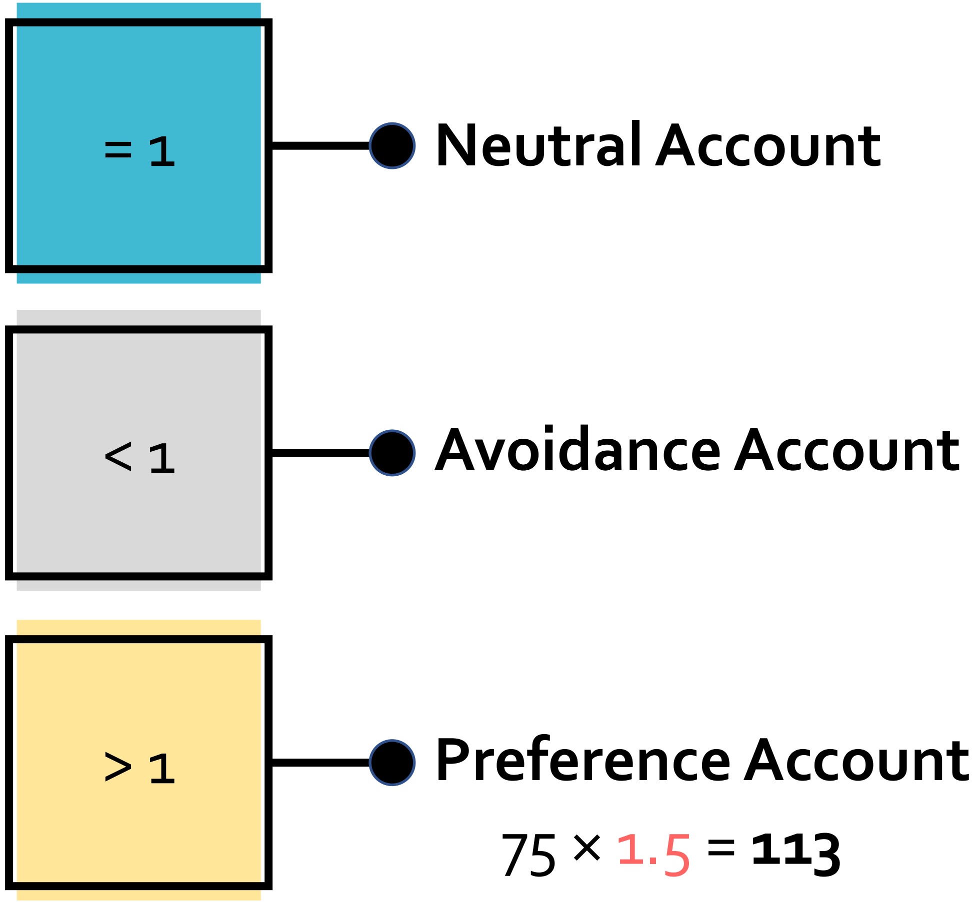
$$\text{Expected no. of Translation equivalents} = \frac{\text{No. of dominant vocabulary} \times \text{No. of non-dominant vocabulary}}{\text{No. of learnable vocabulary}}$$

To evaluate if translation equivalents are learned independently,

$$\text{Expected no. of Translation equivalents} = \frac{\text{No. of dominant vocabulary} \times \text{No. of non-dominant vocabulary}}{\text{No. of learnable vocabulary}} \times \text{Bias parameter}$$

To evaluate if translation equivalents are learned independently,

$$\text{Expected no. of Translation equivalents} = \frac{\text{No. of dominant vocabulary} \times \text{No. of non-dominant vocabulary}}{\text{No. of learnable vocabulary}} \times \text{Bias parameter}$$



Validating the Bilingual Vocabulary Model

1. Running simulations under the Neutral Account

2. Testing the bias parameter with real-life observed data

1 Simulation

Simulated data

- 216 simulated children
- Generated from a range of possible dominant vocabulary from 100 to 600, and a range of non-dominant vocabulary from 0 to 600

Observed data

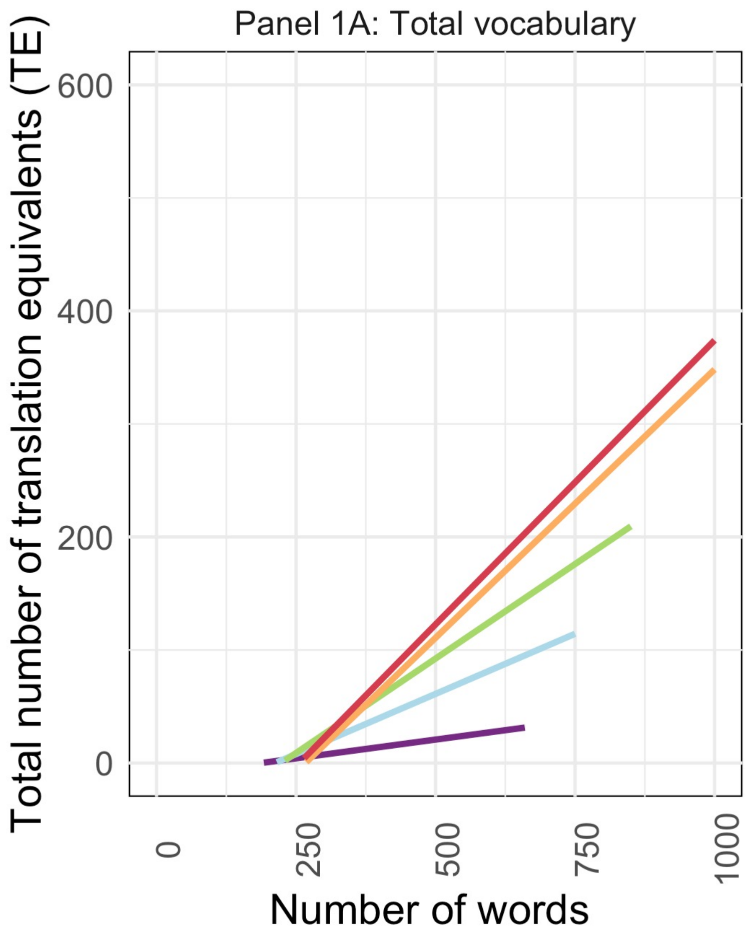
- Archival data collected in Montréal (2010 to 2018)
- 200 English-French bilingual children (18 – 33 months)
- MacArthur-Bates Communicative Development Inventories: Words and Sentences:
  - English (Fenson et al., 2007) and
  - Canadian French (Trudeau et al., 1997)

611 translation equivalents

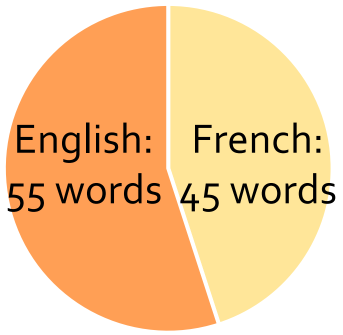
- Identified by 3 proficient bilingual French–English adults

1 Simulation

Simulated data



Proportion of words produced in the non-dominant language relative to the total vocabulary produced

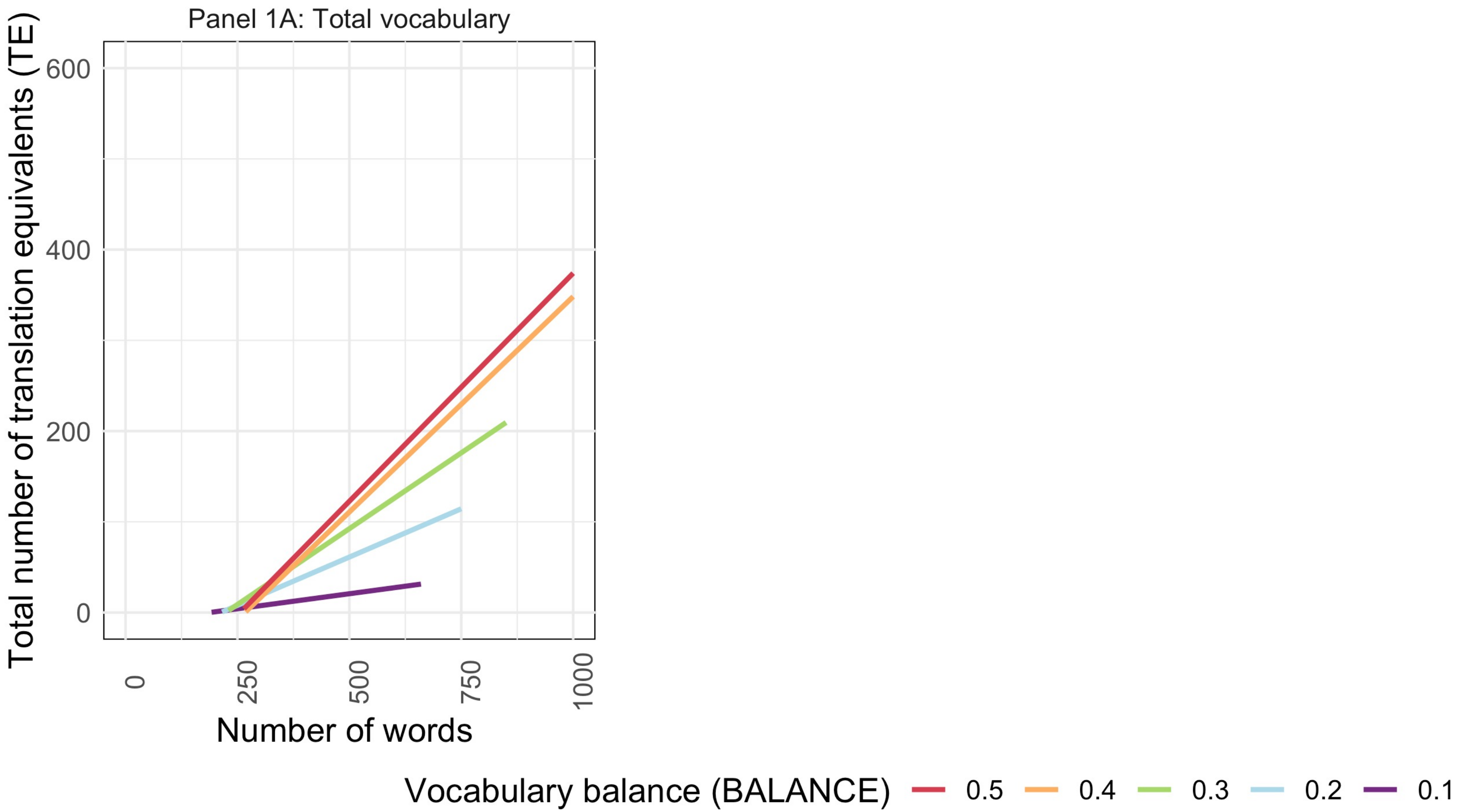


$$= \frac{45}{55 + 45} = \mathbf{0.45}$$

Most balanced ~0.5 ⟷ Least balanced ~0.1

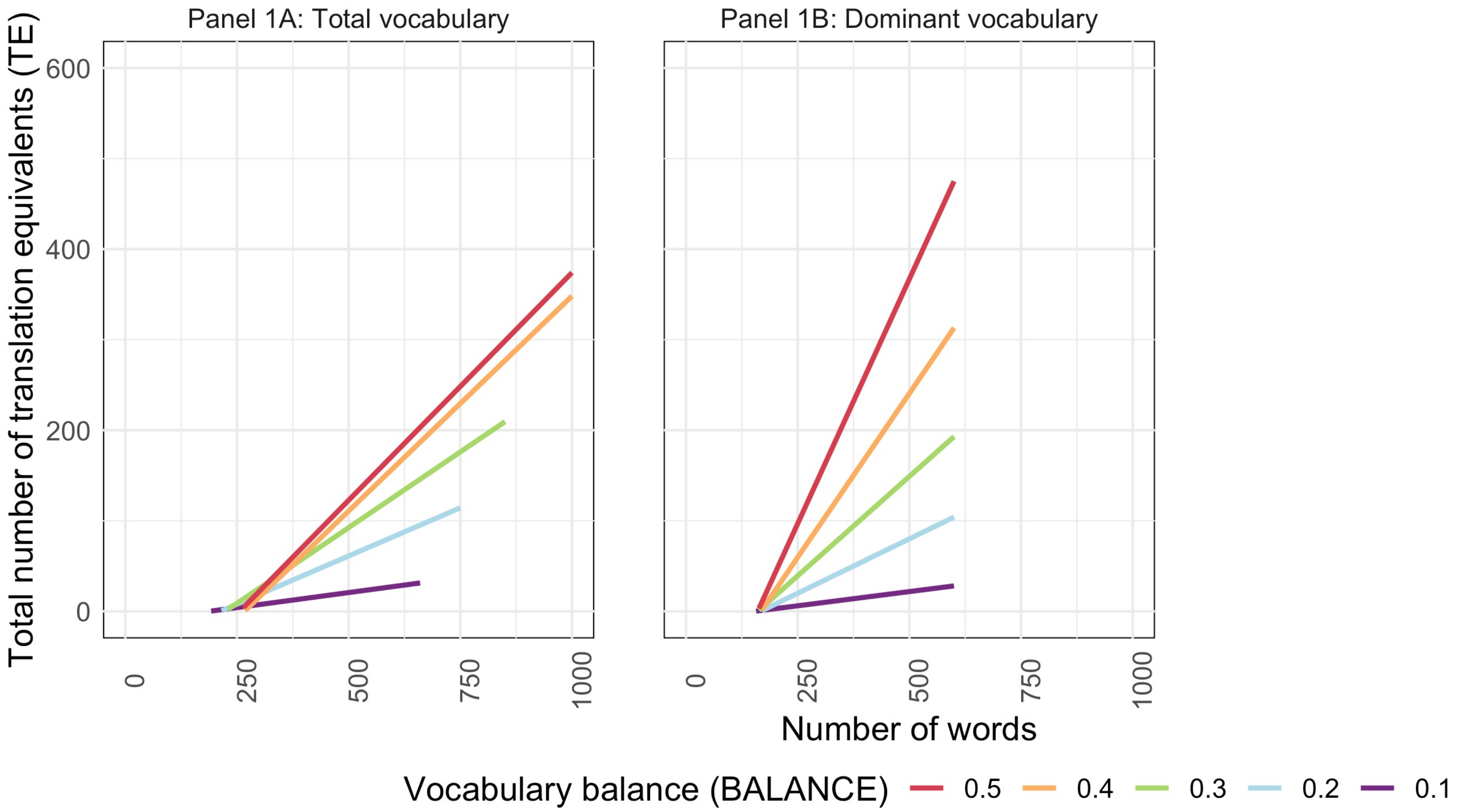
1 Simulation

Simulated data



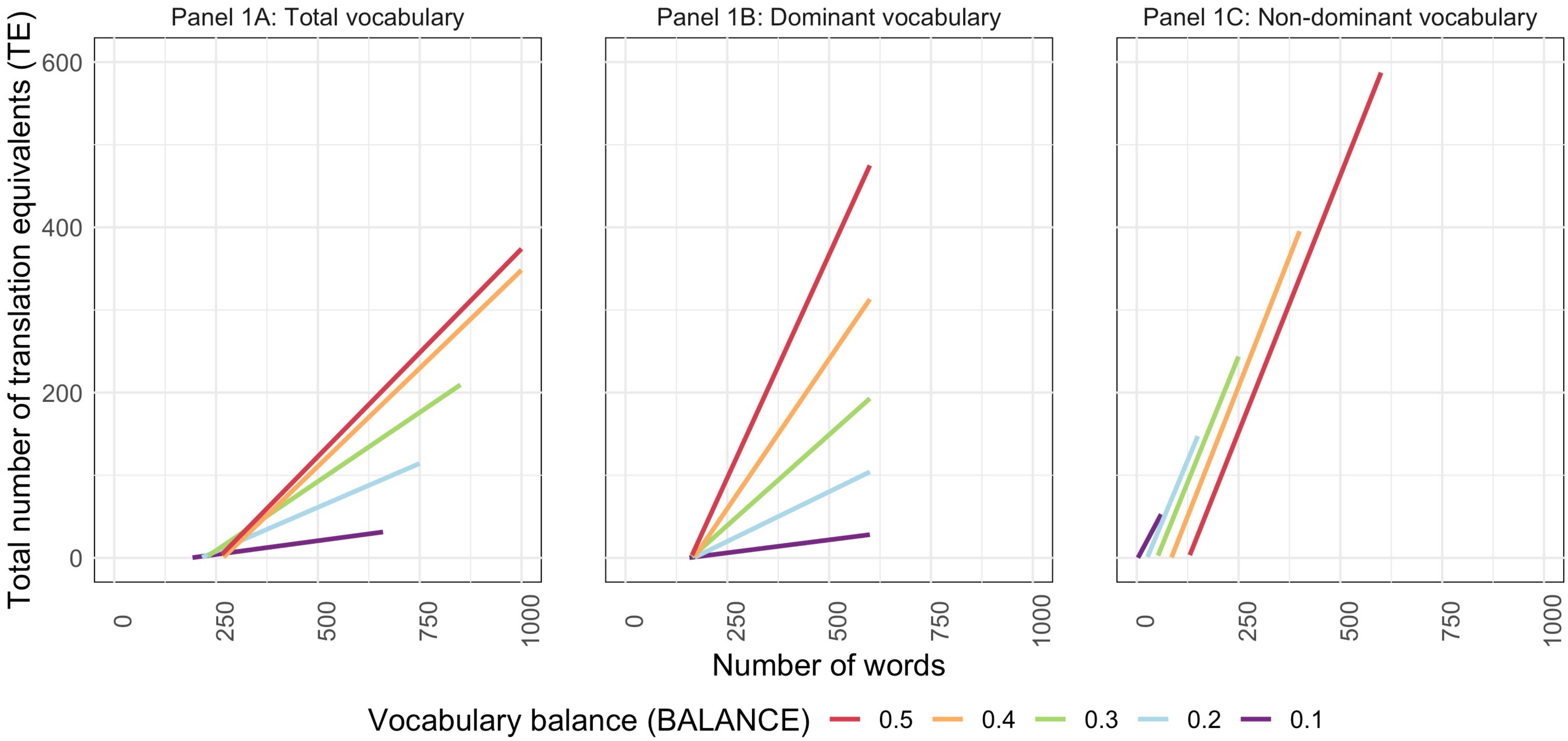
1 Simulation

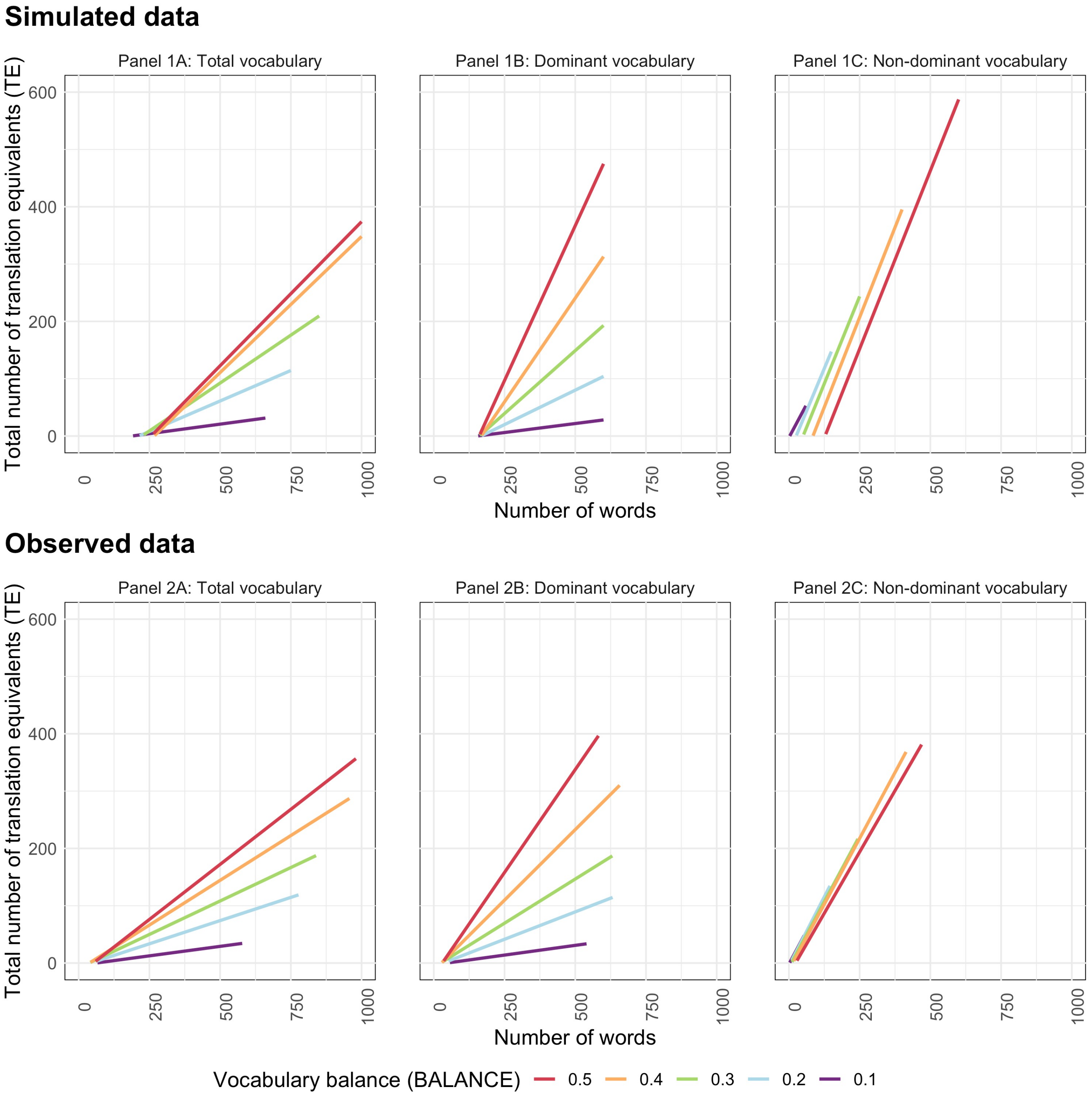
Simulated data



1 Simulation

Simulated data




Simulated data
Panel 1A: Total vocabulary
Panel 1B: Dominant vocabulary
Panel 1C: Non-dominant vocabulary
Total number of translation equivalents (TE)
0
200
400
600
0
250
500
750
1000
Number of words
Observed data
Panel 2A: Total vocabulary
Panel 2B: Dominant vocabulary
Panel 2C: Non-dominant vocabulary
Vocabulary balance (BALANCE)
0.5
0.4
0.3
0.2
0.1

2 Testing the bias parameter

Observed no. of translation equivalents vs.

Expected no. of translation equivalents

$$\frac{\text{No. of dominant vocabulary} \times \text{No. of non-dominant vocabulary}}{\text{No. of learnable vocabulary}} \times \text{Bias parameter}$$

2 Testing the bias parameter

$$\text{Observed no. of translation equivalents} \sim \frac{\text{No. of dominant vocabulary} \times \text{No. of non-dominant vocabulary}}{\text{No. of learnable vocabulary}} \times \text{Bias parameter}$$

2 Testing the bias parameter

$$\text{Observed no. of translation equivalents} \sim 0 + \frac{\text{No. of dominant vocabulary} \times \text{No. of non-dominant vocabulary}}{\text{No. of learnable vocabulary}} \times \text{Bias parameter}$$

No translation equivalents will be produced if a child doesn't produce any vocabulary

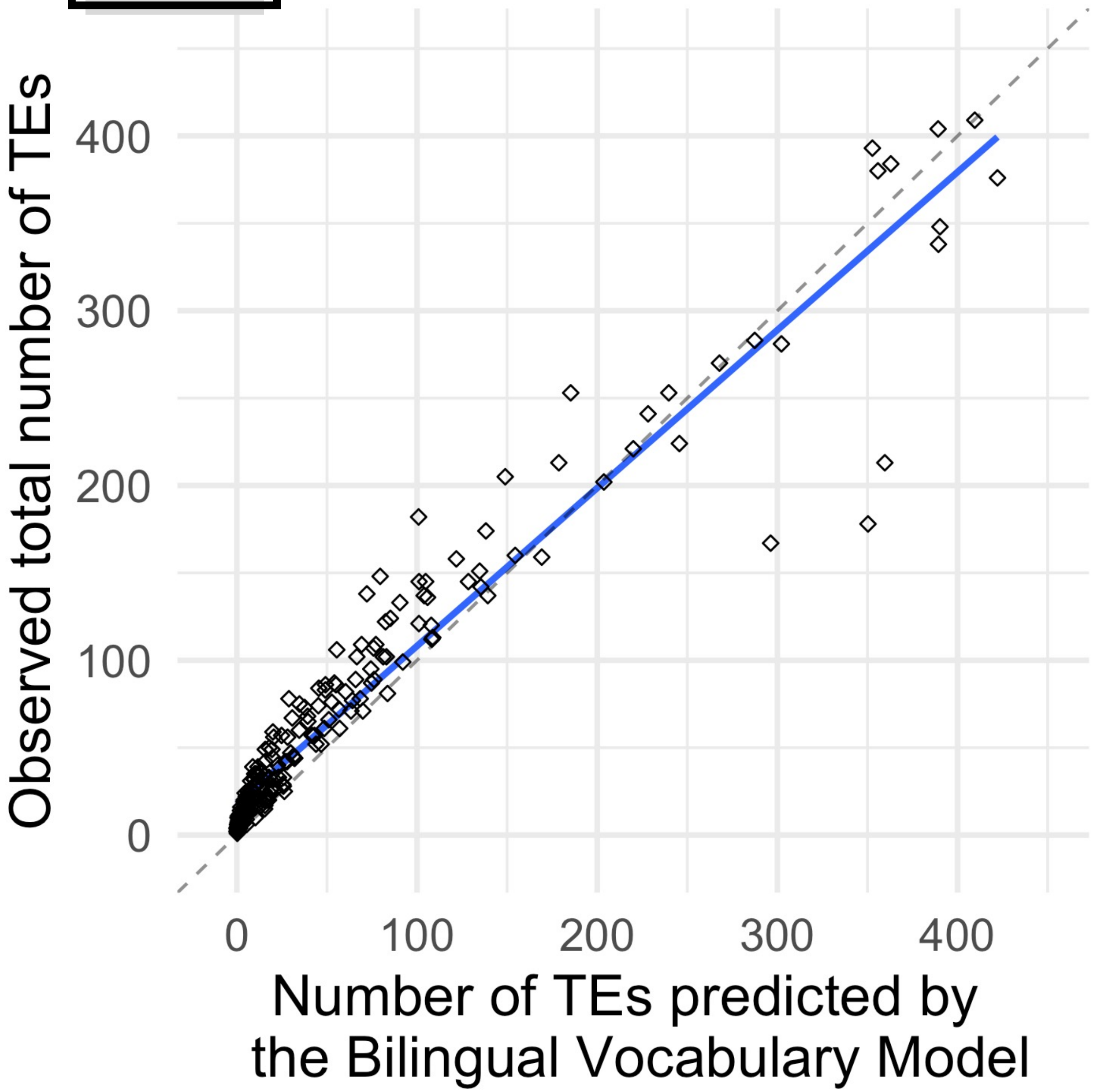
2 Testing the bias parameter

$$\text{Observed no. of translation equivalents} \sim 0 + \frac{\text{No. of dominant vocabulary} \times \text{No. of non-dominant vocabulary}}{\text{No. of learnable vocabulary}} \times \text{Bias parameter}$$

0: No translation equivalents will be produced if a child doesn't produce any vocabulary

Bias parameter: The regression coefficient estimated by the model

2 Testing the bias parameter



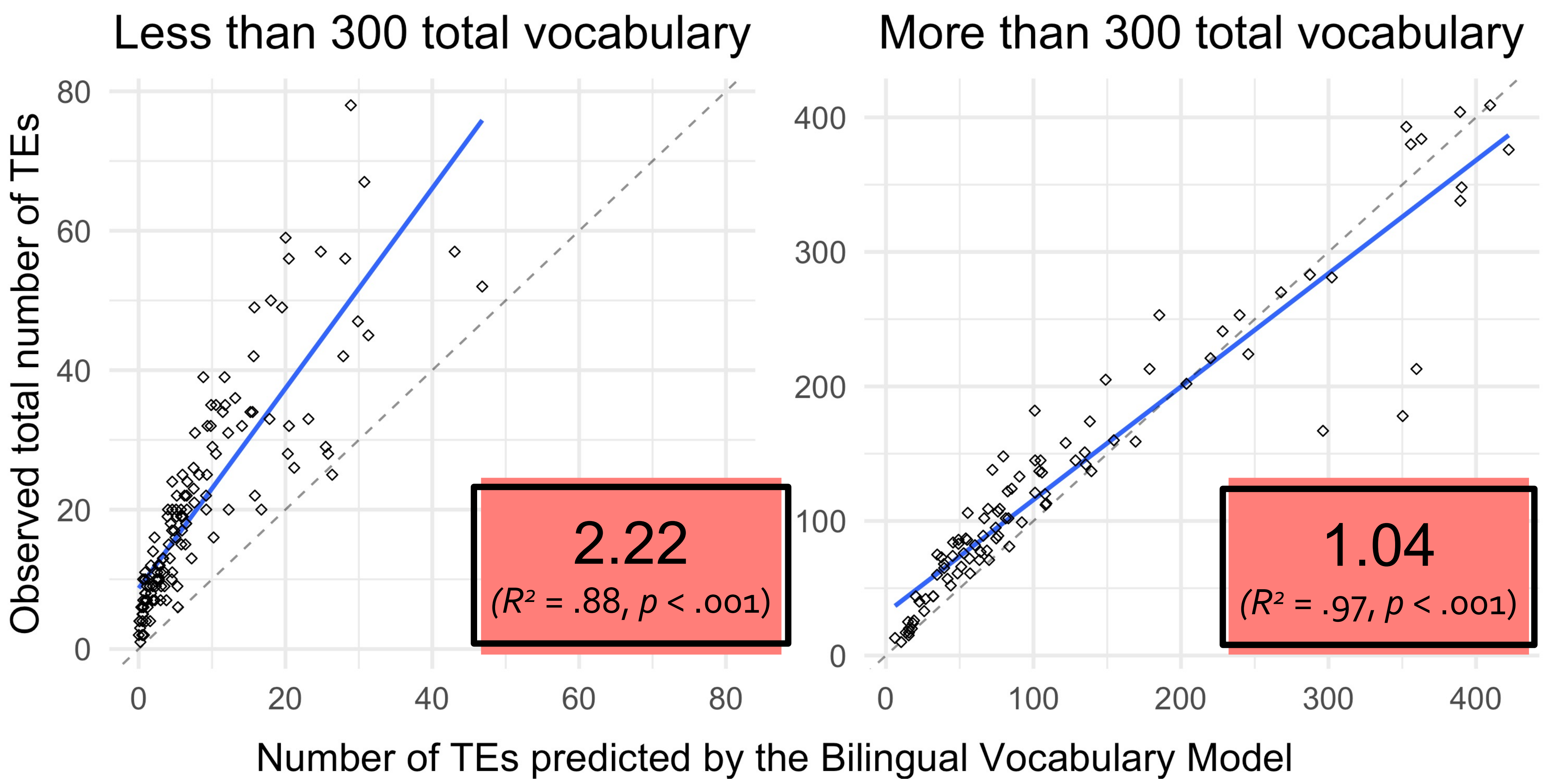
BIAS coefficient

1.02

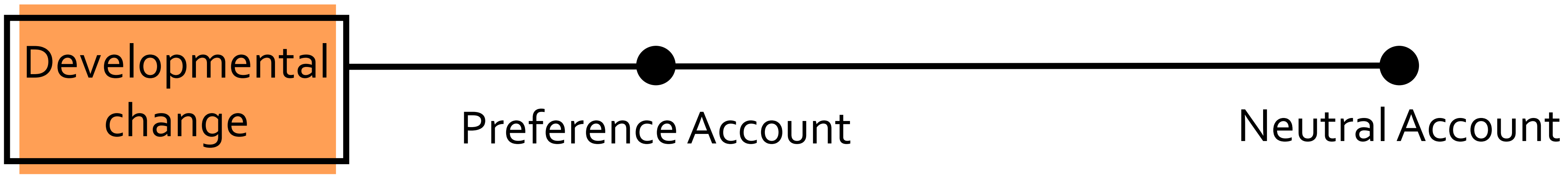
(R^2 = .96, p < .001)

Neutral Account

Translation equivalents are neither harder nor easier to learn than singlets

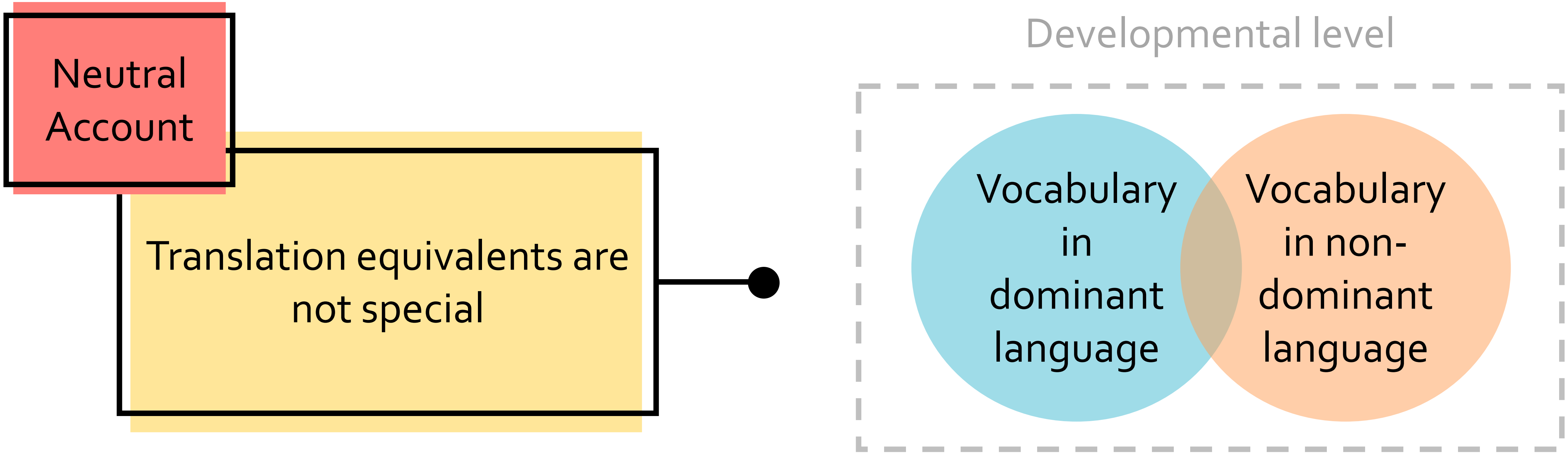


Less than 300 total vocabulary
More than 300 total vocabulary
Observed total number of TEs
Number of TEs predicted by the Bilingual Vocabulary Model
2.22
(R^2 = .88, p < .001)
1.04
(R^2 = .97, p < .001)



Developmental change
Preference Account
Neutral Account

What is the nature of translation equivalent learning?



- Vocabulary in each language develops independently (Marchman, Fernald, & Hurtado, 2010)
- Translation equivalents are the by-chance overlap between the two languages (Pearson et al., 1995)

Contributions of the Bilingual Vocabulary Model

An integrated approach

Including some quantitative factors that can predict vocabulary acquisition

Many other factors:

- A child's efficiency of processing words they hear (e.g., Hurtado et al., 2013; Weisleder & Fernald, 2013)
- Qualitative factors:
  - quality of input (e.g., Raneri et al., 2020, Rowe, 2012),
  - SES (e.g., Hoff, 2003; Fernald, Marchman, & Weisleder, 2013)

Contributions of the Bilingual Vocabulary Model

An integrated approach

Including some quantitative factors that can predict vocabulary acquisition

A simplified computation

Equal opportunities for words to be learned in each of their languages

- A high degree of commonality in the first words children produced (e.g., Braginsky et al., 2016; Tardif et al., 2008)
- Possible that bilinguals learn different words depending on linguistic contexts (Grosjean, 2016)

Translation equivalent learning does not hold a special status and emerges predictably from the word learning process.

THANK YOU!

Preprint here

Rachel Ka-Ying Tsui

rachelkytsui

Ana Maria Gonzalez-Barrero

anamaria.gonzalez@dal.ca

Esther Schott

estSchott

Krista Byers-Heinlein

Krista_BH

Université Concordia University

Concordia Infant Research Laboratory

Laboratoire de recherche sur l'enfance de Concordia