How should we teach vocabulary?
Implications from meta-analytic reviews of L2 vocabulary learning studies

Takumi Uchihara
University of Western Ontario
ALRC seminar series, Jan 26 2021

Overview

- What is meta-analysis?
- How much vocabulary can be learned from context?
- How can we support vocabulary learning from context?
  - Repetition & vocabulary learning
  - Glossing & vocabulary learning
- How much vocabulary can be learned through word-focused activities?
- Summary & Implications

What is meta-analysis?

Narrow definition

- A statistical method for calculating the mean and the variance of a collection of effect sizes (e.g., Pearson's r, Cohen's d) across studies

Example research question

- What is the average correlation between vocabulary size and TOEFL score?

➡the overall effect = 0.76

Study ID	Year	Context	N	Pearson r
Author A	1999	EFL	23	0.65
Author B	2004	EFL	132	0.77
Author C	2019	ESL	57	0.54
Author D	2001	ESL	45	0.86
Author E	2018	ESL	33	0.83
…	…	…	…	…

r = 0.76

What is meta-analysis?

Moderator analysis

- What factors explain effect-size variability across studies?

Example research question

- Does context (EFL vs. ESL) explain the variability across studies?

Study ID	Year	Context	N	Pearson r
Author A	1999	EFL	23	0.65
Author B	2004	EFL	132	0.77
Author C	2019	ESL	57	0.54
Author D	2001	ESL	45	0.86
Author E	2018	ESL	33	0.83
…	…	…	…	…

Between-study variation

What is meta-analysis?

Moderator analysis

- What factors explain effect-size variability across studies?

➢What are moderator variables?

Example research question

- Does context (EFL vs. ESL) explain the variability across studies?

- Go to Dr. Luke Plonsky's website https://lukeplonsky.wordpress.com/
- See Boers et al. (2020) for a critical review of meta-analysis

Study ID	Year	Context	N	Pearson r
Author A	1999	EFL	23	0.65
Author B	2004	EFL	132	0.77
Author C	2019	ESL	57	0.54
Author D	2001	ESL	45	0.86
Author E	2018	ESL	33	0.83
…	…	…	…	…

0.71

VS.

0.85

A list of meta-analysis on L2 vocabulary learning studies

Incidental vocabulary learning

- Montero Perez et al. (2013)
- de Vos et al. (2018)
- Nakanishi (2015)
- Uchihara et al. (2019)
- Webb et al. (under review)

Intentional vocabulary learning

- Webb et al. (2020)

Conditions contributing to learning

- Huang et al. (2012)
- Yanagisawa & Webb (2021)

Glossing

- Abraham (2008)
- Yun (2011)
- Yanagisawa et al. (2020)
- Ramezanali et al. (2020)
- Zhang & Zhang (2020)

Dictionary use

- Zhang et al. (2020)

Corpus use

- Lee et al. (2018)

Strategy training

- Plonsky (2011)

How much vocabulary can be learned from context?

Incidental vocabulary learning

- Learning words as a by-product of comprehension-based activities (e.g., reading, listening, watching television)
- No explicit focus on vocabulary learning

L1 incidental vocabulary learning from reading

- School children appear to increase L1 vocabularies by thousands of words per year through reading written texts (Nagy et al., 1985)

- A meta-analysis of 15 studies (Swanborn & de Glopper, 1999)

➢Students (Grade 5^{th} to 11^{th}) learn around 15% of the unknown words encountered while reading

L2 incidental vocabulary learning

Meta-analysis of 22 studies

- L1 = Arabic, Chinese, Thai, Japanese, Spanish …
- L2 = English and German (k = 1)

Study design

Treatment (n = 1,448)	Control (n = 1,205)
• Reading, Reading+Listening, Listening, Viewing • Exposure to target words • No forewarning of vocabulary posttests	• Test-only condition • No exposure to target words

Webb, Uchihara, & Yanagisawa (under review)

Moderator variables

Coding Scheme

Variables	Values			
Publication information				
Authors				
Year				
Source	Journal	MA/PhD thesis	Conference presentation	Book/book chapter
Learner variables				
Sample size				
Participants' L1				
Target language				
L2 proficiency	Basic	Beyond basic		
Institutional level	Primary	Secondary	University	
Material and activity features				
Text type	Narrative	Expository		
Text audience	L1 users	L2 learners		
Spacing	Spaced	Massed		
Mode of input	Reading	Listening	Reading while listening	Viewing
Methodological features				
Pre-knowledge control	Nonword use	Pilot testing	Pretest	
Test format	Form recognition	Meaning recognition	Meaning recall	Other formats

Moderator variables

Text audience (2 categories)

- Texts for L1 users
- Texts for L2 learners (e.g., graded readers)

Spacing (2 categories)

- Spaced learning condition
- Massed learning condition
  - Spaced = exposure to target words over time (more than 1 day)
  - Massed = one-time (1-day) exposure to target words

Mode of input (4 categories)

- Reading
- Reading while listening
- Listening
- Viewing

Results

- Immediate posttest: **g = 1.14 (Large)**
- Delayed posttest: **g = 1.01 (Large)**

✓Small = 0.40, Medium = 0.70, Large = 1.00 (Plonsky & Oswald, 2014)

Moderator analysis (immediate posttest)

Text audience

- Text for L2 learners (g = 1.54) > Text for L1 users (g = 0.75)

✓Text difficulty influences the rate of learning

Spacing

- Spaced learning (g = 1.51) > Massed learning (g = 0.96)

✓Exposure over time is better than a one-day intensive exposure

Mode of input

- Reading (g = 1.45)
- Listening (g = 0.97)
- Reading while listening (g = 0.78)
- Viewing (g = 0.48)

✓Reading is an important source for incidental vocabulary learning

Rate of learning (number of words learned / total number of target words)

Immediate & delayed posttest

- Reading: 17% & 15%
- Listening: 17% & 16%
- Reading+Listening: 13% & 19%
- Viewing: 4.5%

✓The amount of learning through reading in L2 is comparable to that in L1 (i.e., 15%, Swanborn & de Glopper, 1999)

✓Viewing is not as effective a source of input as other modes of input

Implications

- Learning occurs incidentally (13 to 19%, except viewing)
  - L2 readers may learn as much vocabulary as L1 students (15% gain)
- Text difficulty needs to be appropriate for learners
  - About 98% of the words in texts should be known (Hu & Nation, 2000)
- Repeated exposure needs to be spread out (see Nakata & Elgort, 2020)
  - Narrow reading and viewing => encounter the same words repeatedly
- Viewing may not be as useful as other types of input
  - Captioning supports learning: Captioned viewing > Non captioned viewing ($g = 0.87$) (Montero Perez et al., 2013)
- However, incidental learning may not be sufficient for FL adult learners (i.e., limited amount of input, limited class time, motivation etc.)

How can we support incidental vocabulary learning?

- What should we do, as a teacher or material writer, to promote vocabulary learning from context? (15% => ??)

- ☐ Dictionary use
- ☐ Inference strategy training
- ☐ Provide glosses
- ☐ Provide captions / L1 subtitles
- ☐ Increase the number of word occurrences
- ☐ Input enhancement (e.g., bold type face, highlighting, exaggeration)
- ☐ Spacing encounters
- ☐ etc

To what extend does repetition support L2 incidental vocabulary learning?

LANGUAGE LEARNING
A Journal of Research in Language Studies

SYSTEMATIC REVIEW ARTICLE | Full Access

The Effects of Repetition on Incidental Vocabulary Learning: A Meta-Analysis of Correlational Studies

Takumi Uchihara, Stuart Webb, Akifumi Yanagisawa

First published: 12 March 2019 | https://doi-org.proxy1.lib.uwo.ca/10.1111/lang.12343 | Citations: 16

This research received no specific grant from any funding agency. We are grateful to Judit Kormos, *Language Learning* reviewers, Luke Plonsky, Yo In'nami, Tatsuya Nakata, and Akira Murakami for their constructive feedback on data analysis and earlier versions of the manuscript. We also thank the following researchers who graciously provided information necessary for the current meta-analysis to be completed: Ana Pellicer-Sánchez, Marije Michel, Nina Daskalovska, Niousha Pavia, Sarvenaz Hatami, and Yanxue Feng.

Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, *69*(3), 559-599.

Repetition in incidental vocabulary learning

Goal of the study

- Explore the overall effect of repetition on incidental vocabulary learning & factors moderating the effect (k = 26 studies)

Target effect size

- Correlation between the number of encounters and learning gains

Life in London

I have **lived** in London for one year. I think **living** in an English-speaking **country** is the best way to improve English skills and enrich your life experience. Have you ever **lived** abroad? If you haven't, I highly recommend you should consider **living** in other **countries**.

"Live" x 4 encounters => larger learning gain

"Country" x 2 encounters => smaller learning gain

Results & Implications

- The overall correlation was moderate: r = .34
- Mode of input
  - Reading: r = .41
  - Listening: r = .39
  - Reading + listening: r = .28
  - Viewing: r = .22

➢Ensuring repeated encounters with L2 words is important especially in designing reading (and listening) materials

➢Repetition is not a single factor resulting in incidental vocabulary learning

To what extent does glossing support L2 vocabulary learning from reading?

Different types of glosses

1. Marginal glosses
2. Interlinear glosses
3. In-text glosses
4. Glossaries
5. Multiple-choice glosses
6. Hyperlinked glosses

1. Marginal glosses

Each year in the U.S. about 7,000 infants die in their **cribs** for no apparent reason. In 2019, the number has decreased. However, it appears that the vast majority of parents are still not …

Cribs = babies' beds (L2)

Cribs = ベビーベッド (L1)

2. Interlinear glosses

Each year in the U.S. about 7,000 infants die in their **cribs** for no apparent reason.

Cribs = babies' beds

In 2019, the number has decreased. However, it appears that the vast majority of parents are still not …

3. In-text glosses

Each year in the U.S. about 7,000 infants die in their **cribs**, babies' beds, for no apparent reason. In 2019, the number has decreased. However, it appears that the vast majority of parents are still not …

4. Glossaries

Each year in the U.S. about 7,000 infants die in their **cribs** for no apparent reason. In 2019, the number has decreased. However, it appears that the vast majority of parents are still not …

★Provided at the end of the text or as a separate paper in the form of a list

Vocabulary List

Cribs = babies‘ beds

Decrease = to go down in number or quantity

Infants = babies or very young children

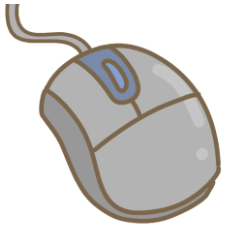
5. Multiple-choice glosses

Each year in the U.S. about 7,000 infants die in their **cribs** for no apparent reason. In 2019, the number has decreased. However, it appears that the vast majority of parents are still not …

Cribs = 1. babies' beds

2. small beds

6. Hyperlinked glosses



Each year in the U.S. about 7,000 infants die in their **cribs** for no apparent reason. In 2019, the number has decreased. However, it appears that the vast majority of parents are still not …

Cribs = babies' beds

Which types of glosses most effectively support L2 vocabulary learning from reading?

Studies in Second Language Acquisition

HOW DO DIFFERENT FORMS OF GLOSSING CONTRIBUTE TO L2 VOCABULARY LEARNING FROM READING?

A META-REGRESSION ANALYSIS

Published online by Cambridge University Press: **14 January 2020**

Akifumi Yanagisawa, Stuart Webb and Takumi Uchihara

Show author details

Article | Supplementary materials | Metrics

Save PDF | Share | Cite | Rights & Permissions

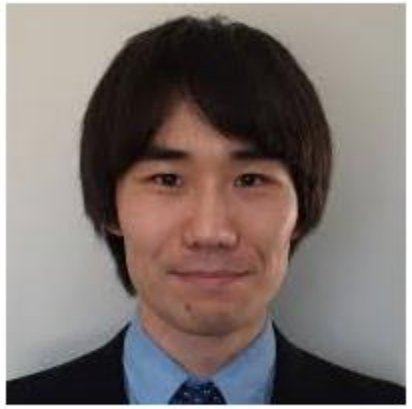
Abstract

This meta-analysis investigated the overall effects of glossing on L2 vocabulary learning from reading and the influence of potential moderator variables: gloss format (type, language, mode) and text and learner characteristics. A total of 359

Article contents

Abstract

HOW DO DIFFERENT FORMS OF GLOSSING CONTRIBUTE TO L2 VOCABULARY



Yanagisawa, A., Webb, S., & Uchihara, T. (2020). How do different forms of glossing contribute to L2 vocabulary learning from reading? A meta-regression analysis. *Studies in Second Language Acquisition*, *42*(2), 411-438.

Results

TABLE 3. The learning gain for each gloss type compared to the nonglossed condition

	Immediate					Delayed				
	k	*n*	Mean ES difference (%)	CI	*p*	*k*	*n*	Mean ES difference (%)	CI	*p*
Multiple-choice	12	31	25.2	18.5, 31.8	< .001	12	21	15.6	9.0, 22.3	< .001
Hyperlinked	11	35	18.4	5.9, 30.9	.009	11	33	15.2	3.1, 27.3	.020
Marginal	25	69	17.8	13.5, 22.0	.001	21	50	12.8	9.6, 16.0	< .001
Glossaries	2	3	17.4	–27.7, 62.5	.134	3	5	10.4	–9.9, 30.6	.147
Interlinear	1	4	16.0	8.5, 23.5	.004	0	0	–	–	–
In-text	4	12	11.0	–0.4, 22.4	.055	3	4	6.5	–19.0, 32.1	.412

Note: *k* = number of studies, *n* = number of ESs, CI = 95% confidence interval, Mean ES difference (%) = mean effect size differences between each gloss type and the nonglossed condition converted into a percentage. *p* = *p*-value for significant test.

Results & Implications

- **Multiple-choice glossing** was most effective
  - 25.2% & 15.6% higher than nonglossed reading (immediate & delayed)
- **Hyperlinked, marginal, & interlinear glosses** are moderately effective
- The least effective gloss types: **Glossaries & In-text glosses**
  - No significant differences when compared to nonglosses reading
- L1 glossing led to gains 4% higher than L2 glossing
  - L1 > L2 glossing (g = 0.33) (Kim et al., 2020)

➢Location of glosses should not be too close to (e.g., in-text glosses) or far from (e.g., glossaries) target words

➢L1 glossing is more effective than L2 glossing

Glossing promotes L2 vocabulary learning from reading

- Different modes of glosses can be presented in combination
- Mode of gloss
  - L1 or L2 texts
  - Pictures
  - Videos
  - Audio
- Commonly used combinations include:
  - Picture + Text (two modes) = **dual**
  - Picture + Text + Audio (three modes) = **triple**
- **Assumption:** multimodal glosses > single-mode glosses
  - Supported by previous meta-analyses: 2 > 1 (e.g., Yun, 2011)
- Does an additional mode of gloss lead to an additional benefit for vocabulary learning? (e.g., 3 > 2, 4 > 3, 5 > 4, etc.)

Does additional modes of input lead to more effective vocabulary learning?

tesol QUARTERLY

Full-Length Article | Full Access

Efficacy of Multimodal Glossing on Second Language Vocabulary Learning: A Meta-analysis

Nasrin Ramezanali, Takumi Uchihara, Farahnaz Faez

First published: 21 April 2020 | https://doi-org.proxy1.lib.uwo.ca/10.1002/tesq.579 | Citations: 1

Abstract

This meta-analysis examined the effectiveness of an additional gloss mode in single versus dual and dual versus triple glossing on second language (L2) learners' word learning. In total, 22 studies, providing 26 independent effect sizes, were coded, and 11 moderator variables including quality of data sample, learner variables, gloss features,

Ramezanali, N., Uchihara, T., & Faez, F. (2020). Efficacy of Multimodal Glossing on Second Language Vocabulary Learning: A Meta-analysis. *TESOL Quarterly*.

Goal of the study: Explore the effectiveness of an additional gloss mode in (1) single vs. dual & (2) dual vs. triple

Results & Implications

Overall additional mode effect

- g = 0.46 (immediate posttest)
- g = 0.28 (delayed posttest)

➢The effect may not be durable (medium => small)

2 vs. 1 & 3 vs. 2

- Dual > Single (g = 0.58, p < .05)
- Triple = Dual (g = 0.02, p > .05)

➢Adding another mode is not always facilitative

Language

- Additional mode (e.g., picture) added to **L1 gloss** **g = 0.25**
- Additional mode (e.g., picture) added to **L2 gloss** **g = 0.61**

➢More effective to add another mode to L2 textual glosses (vs. L1 glosses)

Incidental vs. intentional vocabulary learning

- **Incidental vocabulary learning** is incremental and gradual, requiring large amounts of input over a long period of time

- In immediate need of learning certain words:
  - Learning technical terms or academic words to keep up in content courses
  - Beginners learn high-frequency L2 words (e.g., like, have, take)
  - Travelers want to learn survival vocabulary (e.g., reservation, top up, delay)

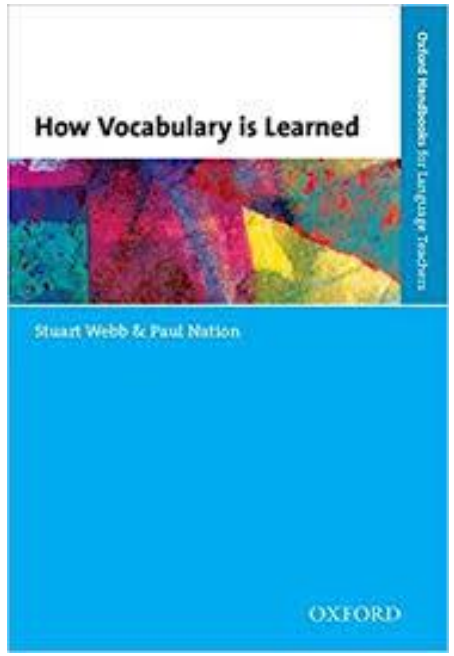
- **Intentional vocabulary learning** (= deliberate vocabulary learning, word-focused learning) is quick and efficient; learners engage in word-focused activities
  - Flashcards, word lists, word-matching, writing, crossword puzzles

- What kinds of word-focused activities are more or less effective than others?

- Which activities should we use in class or encourage learners to use?

- See Webb & Nation (2017) for

a list of 23 vocabulary learning activities



Webb, S., & Nation, I. S. P. (2017). How vocabulary is learned? Oxford: Oxford University Press

How much vocabulary can be learned through word-focused activities?

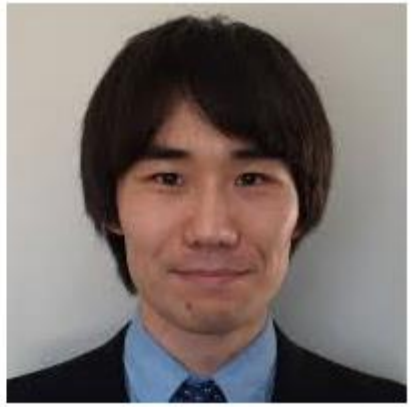
The Modern Language Journal

Original Article | Full Access

How Effective Are Intentional Vocabulary-Learning Activities? A Meta-Analysis

STUART WEBB, AKIFUMI YANAGISAWA, TAKUMI UCHIHARA

First published: 02 November 2020 | https://doi-org.proxy1.lib.uwo.ca/10.1111/modl.12671



Abstract

The present meta-analysis aimed to summarize the extent to which second language vocabulary is learned from the most frequently researched word-focused activities: flashcards, word lists, writing, and fill-in-the-blanks. One hundred effect sizes from 22

Webb, S., Yanagisawa, A., & Uchihara, T. (2020). How Effective Are Intentional Vocabulary Learning Activities? A Meta-analysis. *The Modern Language Journal*, *104*(4), 715-738.

Four commonly used activities were compared

1. **Flashcards**
   - L2 form => L1 meaning, L1 meaning => L2 form
   - *e.g., Kome (L1) => _____?*
   - *e.g., Rice (L2) => _____?*
2. **Word lists**
   - L1 meaning & L2 form presented together
   - *e.g., Kome (L1) : Rice (L2)*
3. **Writing**
   - Writing sentences using target words
4. **Fill-in-the-blanks**
   - Completing sentences with a blank using target words
   - *e.g., Children who watch a lot of TV do not _____ between reality and fantasy.*

Results (immediate posttest)

L2-to-L1 translation

L1-to-L2 translation

TABLE 2
Estimated Effect Size (ES) of Proportion of Target Words Learned on Immediate Posttests

	Meaning Recall				Form Recall			
Activity	*k*	*n*	Mean ES (SE)	CI	*k*	*n*	Mean ES (SE)	CI
Fill-in-the-blanks	8	9	0.431 (0.056)	[0.29, 0.56]	2	3	0.184 (0.042)	[–0.35, 0.72]
Writing	10	14	0.548 (0.018)	[0.43, 0.66]	4	6	0.368 (0.075)	[0.10, 0.62]
Word lists	5	11	0.732 (0.075)	[0.51, 0.94]	7	14	0.701 (0.051)	[0.57, 0.83]
Flashcards	2	6	0.770 (0.050)	[0.53, 1.00]	4	14	0.661 (0.048)	[0.50, 0.81]

Note. k = number of studies; n = number of ESs; SE = standard error; CI = 95% confidence interval adjusted with RVE. The total number of studies = 20. The total number of ESs = 77.

Results (delayed posttest: 4 days to 2 weeks after the treatment)

TABLE 3
Estimated Proportion of the Target Words Retained

	Meaning Recall				Form Recall			
Activity	*k*	*n*	Mean ES (SE)	CI	*k*	*n*	Mean ES (SE)	CI
Fill-in-the-blanks	8	12	0.239 (0.049)	[0.12, 0.35]	2	3	0.183 (0.056)	[–0.65, 1.01]
Writing	8	14	0.319 (0.047)	[0.20, 0.43]	3	14	0.180 (0.068)	[–0.15, 0.52]
Word list	2	5	0.479 (0.018)	[0.24, 0.71]	2	4	0.218 (0.024)	[–0.09, 0.53]
Flashcards	2	6	0.734 (0.012)	[0.58, 0.88]	4	4	0.320 (0.049)	[0.15, 0.48]

Note. ES = effect size; *k* = number of studies; *n* = number of ESs; CI = 95% confidence interval adjusted with RVE. The total number of studies = 15. The total number of ESs = 62.

60%, 59% (Immediate posttest) => 40%, 25% (delayed posttest)

Implications

- **Flashcards & Word Lists** are the most effective activities
  - 66-77% gains (immediate posttest)
  - But less so for delayed posttests (22-48%) (except flashcards for meaning recall, 73%)

- Intentional vocabulary learning: **40%, 25%** (delayed posttest)
  - VS. **15%** for incidental learning from reading (delayed posttest)

➢Learning from word-focused tasks is far from guaranteed

Summary: How should we teach vocabulary?

- Increase repetitions & space them
  - ➢Narrow reading, listening, and viewing

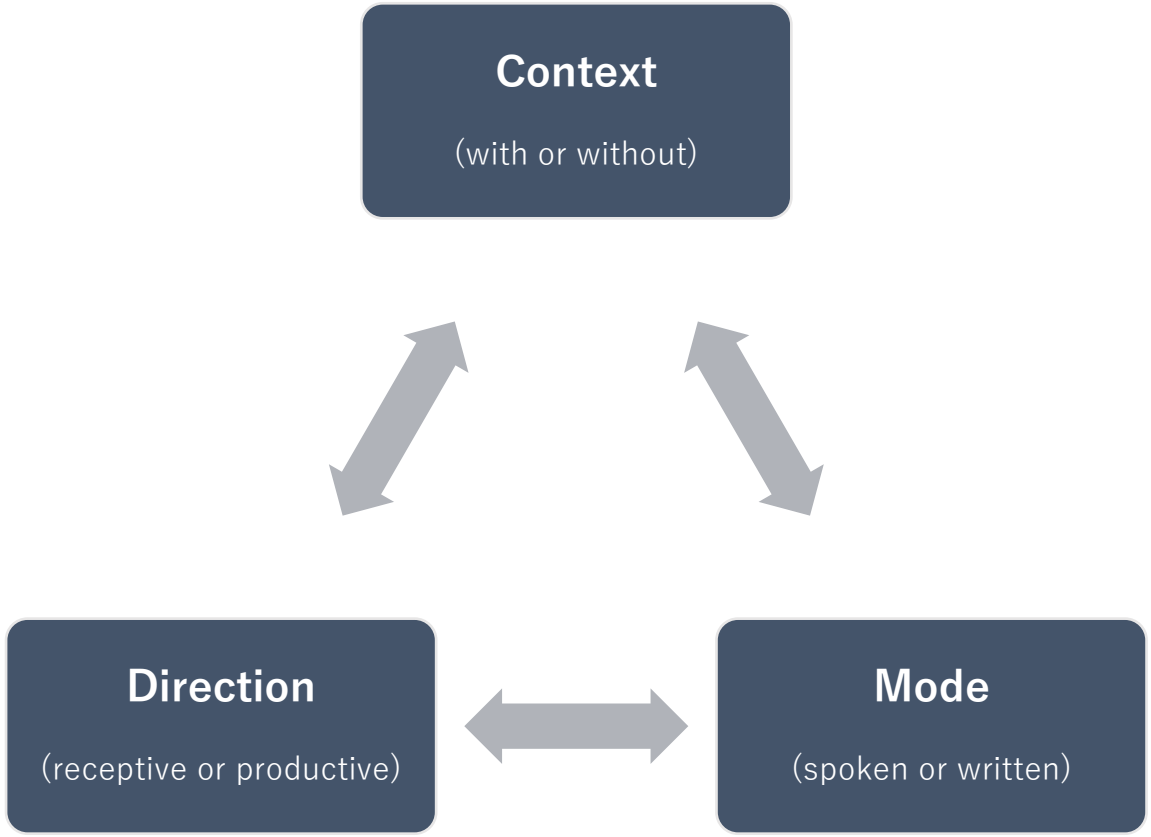
- Support meaning-focused input with captions and glosses

- In using glosses, remember that:
  - L1 glosses lead to word learning more efficiently than L2 glosses
  - Multiple-choice glosses lead to better learning, but glossaries and in-text glosses should be avoided
  - Two gloss modes are more effective than single mode glosses, but three gloss modes might not further promote learning compared to two modes
  - Adding another gloss to L2 textual glosses increases learning significantly

- Use flashcards effectively
  - Direction of learning
    - L1 meaning => L2 form, L2 form => L1 meaning
    - e.g, Kome => _____? Rice => _____?
  - The spacing of repetitions
  - Changing the order of the cards to avoid serial learning
  - Use apps and save time (see Nakata, 2011)
  - For a review, see Nakata (2020)

- Supplement word-focused activities with meaning-focused input+output activities (e.g., linked skills, Webb & Nation, 2017)
  - ➢Flashcard learning => Reading => Discussing with peers => essay writing

Rec or Pro S or W WithoutCon	Rec + W + WithCon	Rec+Pro + S + WithCon	Pro + W + WithCon



How should we teach vocabulary?

Suggestions from other meta-analysis research

- Output practice (Huang et al., 2012)
- Oral interaction (de Vos et al., 2018)
- Dictionary (Zhang et al., 2020)
- Corpus (Lee et al., 2018)
- Learning strategy (Plonsky, 2011)
- Spaced retrieval (Kim & Webb, under review)
- Avoid semantically related words (Nakata & Suzuki, 2019)
- Depth of processing (Yanagisawa & Webb, 2021)

Future directions

- Most studies focus on form-meaning connection (e.g., via cued translation or multiple choice tests)
  - Collocation
  - Pronunciation
  - Appropriateness/pragmatics
- Few replication studies
  - Same research design and materials in virtual environments

References

- de Vos, J. F., Schriefers, H., Nivard, M. G., & Lemhöfer, K. (2018). A meta - analysis and meta - regression of incidental second language word learning from spoken input. Language Learning, 68(4), 906-941.
- Huang, S., Willson, V., & Eslami, Z. (2012). The effects of task involvement load on L2 incidental vocabulary learning: A meta - analytic study. The Modern Language Journal, 96(4), 544-557.
- Kim, H. S., Lee, J. H., & Lee, H. (2020). The relative effects of L1 and L2 glosses on L2 learning: A meta-analysis. Language Teaching Research.
- Lee, H., Warschauer, M., & Lee, J. H. (2019). The effects of corpus use on second language vocabulary learning: A multilevel meta-analysis. Applied Linguistics, 40(5), 721-753.
- Nakata, T., & Suzuki, Y. (2019). Effects of massing and spacing on the learning of semantically related and unrelated words. Studies in Second Language Acquisition, 41(2), 287-311.
- Plonsky, L. (2011). The effectiveness of second language strategy instruction: A meta - analysis. *Language learning*, *61*(4), 993-1038.
- Yanagisawa, A., & Webb, S. To What Extent Does the Involvement Load Hypothesis Predict Incidental L2 Vocabulary Learning? A Meta - Analysis. *Language Learning*.
- Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning*, *24*(1), 39-58.
- Zhang, S., Xu, H., & Zhang, X. (2020). The effects of dictionary use on second language vocabulary acquisition: A meta-analysis. International Journal of Lexicography.