



Frequency of occurrence in output helps predict incidental vocabulary learning

Serge Bibauw · Thomas François · Piet Desmet

Universidad Central del Ecuador · UCLouvain · KU Leuven

AAAL 2023 Conference

Portland, OR

March 20, 2023



← Slides

 **UCLouvain**

 **KU LEUVEN**





Frequency in output and incidental vocabulary learning

interacting with a chatbot

Frequency of occurrence

Experiment: dialogue system

Results: incidental effects

Discussion & model



Frequency in output and incidental vocabulary learning

Frequency of occurrence

Experiment: dialogue system

Results: incidental effects

Discussion & model



Incidental vocabulary learning and interaction

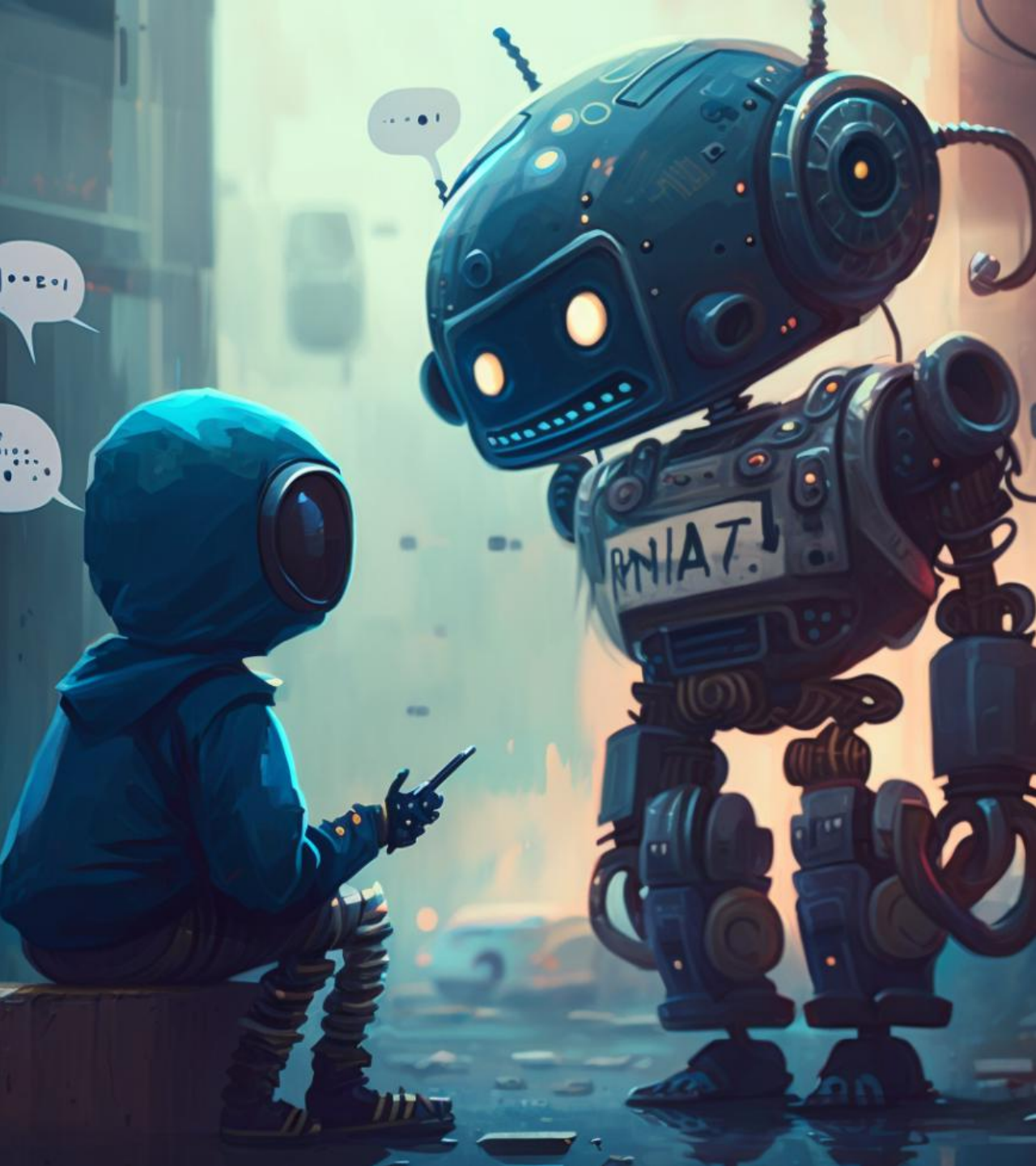
- **Involvement-Load Hypothesis** (Laufer & Hulstijn, 2001)
Factors for word retention: *need, evaluation, search* (Yanagisawa & Webb, 2021a)
 - Word use for completing the task (++) (Newton, 1995)
 - Interactionally modified output (++) (Ellis & He, 1999)
- **Involvement-Load Hypothesis Plus** (Yanagisawa & Webb, 2021b)
 - Varied use in sentence (++) and composition (+++)
- Meta-analysis of incidental word learning from spoken input (de Vos, Schriefers, Nivard & Lemhöfer, 2018)
Interactive tasks ($d + 0.73$) > Non-interactive tasks ($d + 0.10$) > Input only
- **Interactive sources of exposure:** highest predictors for vocabulary size:
online/games ($\beta = .39^{***}$, Peters *et al.*, 2019)
social media ($\beta = .29^{***}$), speaking L2 ($\beta = .13^{***}$, De Wilde *et al.*, 2019)



Frequency of occurrence (in input)

- Well-known factor influencing incidental vocabulary acquisition (Webb, 2020; Uchihara *et al.*, 2019)
- Medium predicting power in vocabulary learning, around $r = .34$ (Uchihara *et al.*, 2019)
- But no study about frequency in *output*.
- **RQ:**
How much do *productive* word uses help understand and predict incidental word learning?





Frequency in output and incidental vocabulary learning

Frequency of occurrence

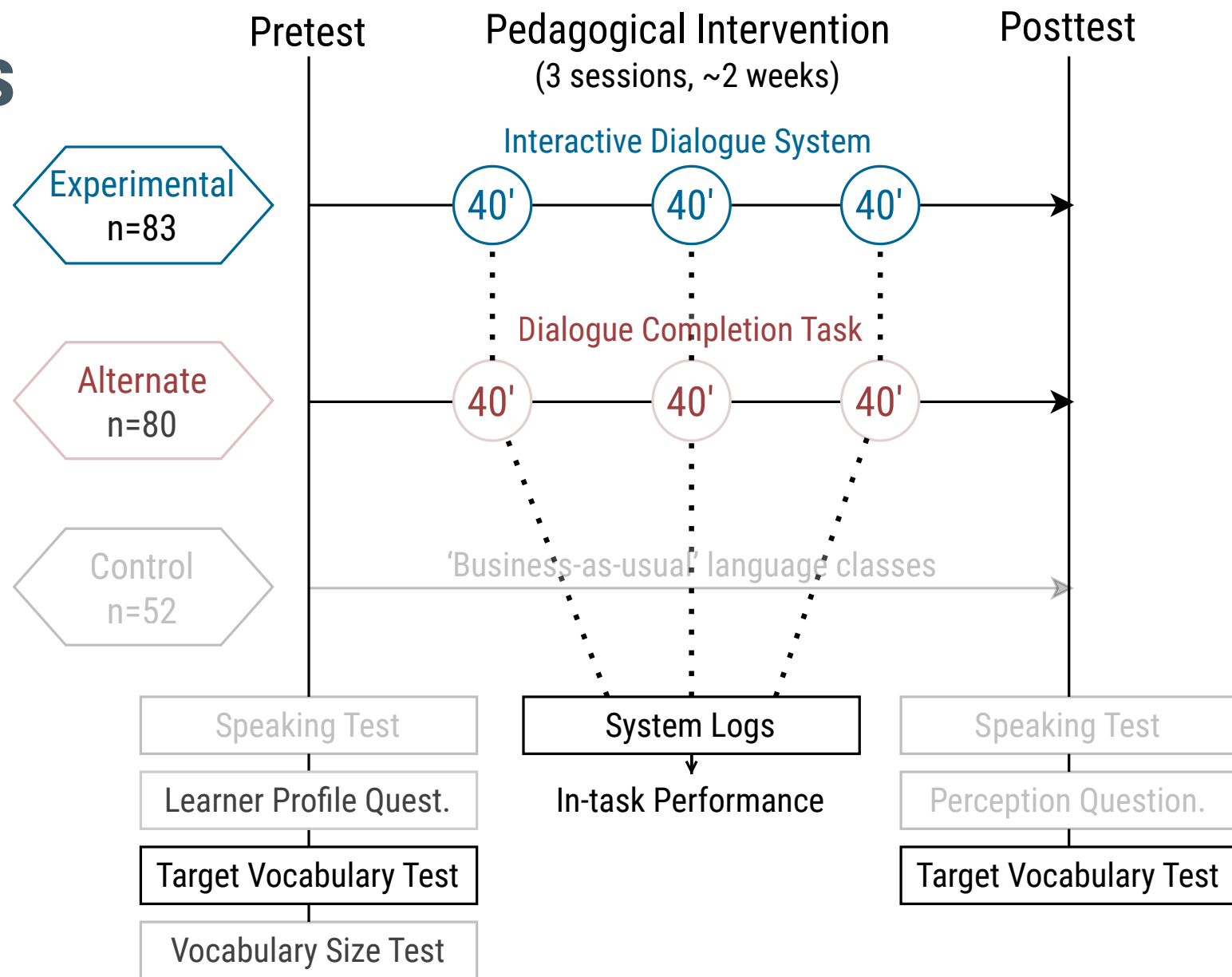
Experiment: dialogue system

Results: incidental effects

Discussion & model



Methods



Intervention: Dialogue system

- Developed in 4 BC (before ChatGPT)
- ‘Designable’ and controllable spontaneous interaction.
- Full logging of all messages read and written in the system
 - 48 353 messages
 - 577 494 words
- Alternate condition: Static dialogue completion task



Score: 424 ?
Familiarité niveau 0 : Connaissance

Tâche en cours (2/18):
Vraag waar je bent.

Lu: *Bonjour, petit hibou. Tu es enfin réveillé!*
 Vous: bonjour
 Lu: *Comment est-ce que tu vas? Tu es tombé de très haut!*
 Vous: Je suis bien
 Lu: *Ouf, c'est une bonne nouvelle! Ton aile avait l'air très mal-en-point quand je t'ai amené ici.*
 Vous: comment cava?
 Lu: *Je vais bien.*
 Vous: tu habite ou
 Correction: habite - Vérifiez l'accord entre le pronom « tu » et le verbe « habite ».
 Il ne semble pas avoir entendu...

Écrivez ou dites votre réponse:
Écrivez ici..

Envoyer réponse
Aidez-moi
Terminer la conversation



He: *Tu es dans le Grand Monastère des Montagnes éternelles, mon petit.*

qui tu es?

He: *Je suis Sensei, le moine. Je suis le gardien de ce monastère. Quel est ton nom, mon petit?*

mon name es Jean

He: *Enchanté, Rinc.*

Task: Ask what happened.

He: *Il y a eu une tempête. Une terrible tempête! Tu t'es perdu. Je t'ai trouvé près d'ici. Tu étais tout seul. Je t'ai amené ici. Dis-moi Rinc, est-ce que tu as remarqué quelque chose de bizarre pendant la tempête?*

Typ tekst..

He: *Est-ce que tu as entendu des voix dans le tonnerre?*

Send your reply
Send your reply
Quit

Participants

- **Multisite:** 4 Flemish schools, with 2-3 classes each
 - $N = 215$ $N_{\text{classes}} = 11$
 - Here, focus on the dialogue system condition:
 $n_{\text{ds}} = 81$
- Teenagers: 12-13 y.o. (2nd grade of sec. school, ~8th grade)
- **L1** = 95% Dutch
- **L2** = French (first L2 learned at school)
M = 3.1 years of French instruction
Mostly at A1 level (beginners)
- In “classroom” (school computer lab)



Instrument: Target vocabulary test

- Target words: selected because of (expected) occurrence, but no focus of instruction
→ **Incidental learning** only
 - At pre- & posttest, randomized, identical tests
 - **50 target items**
1. **Receptive** part: **meaning recognition**
25 isolated words *potager*
 2. **Productive** part: **form recall**
25 'formulaic sequences' in gap-filling

beaucoup d'imagination

le membre

lid

vlies

gebouw

hetzelfde

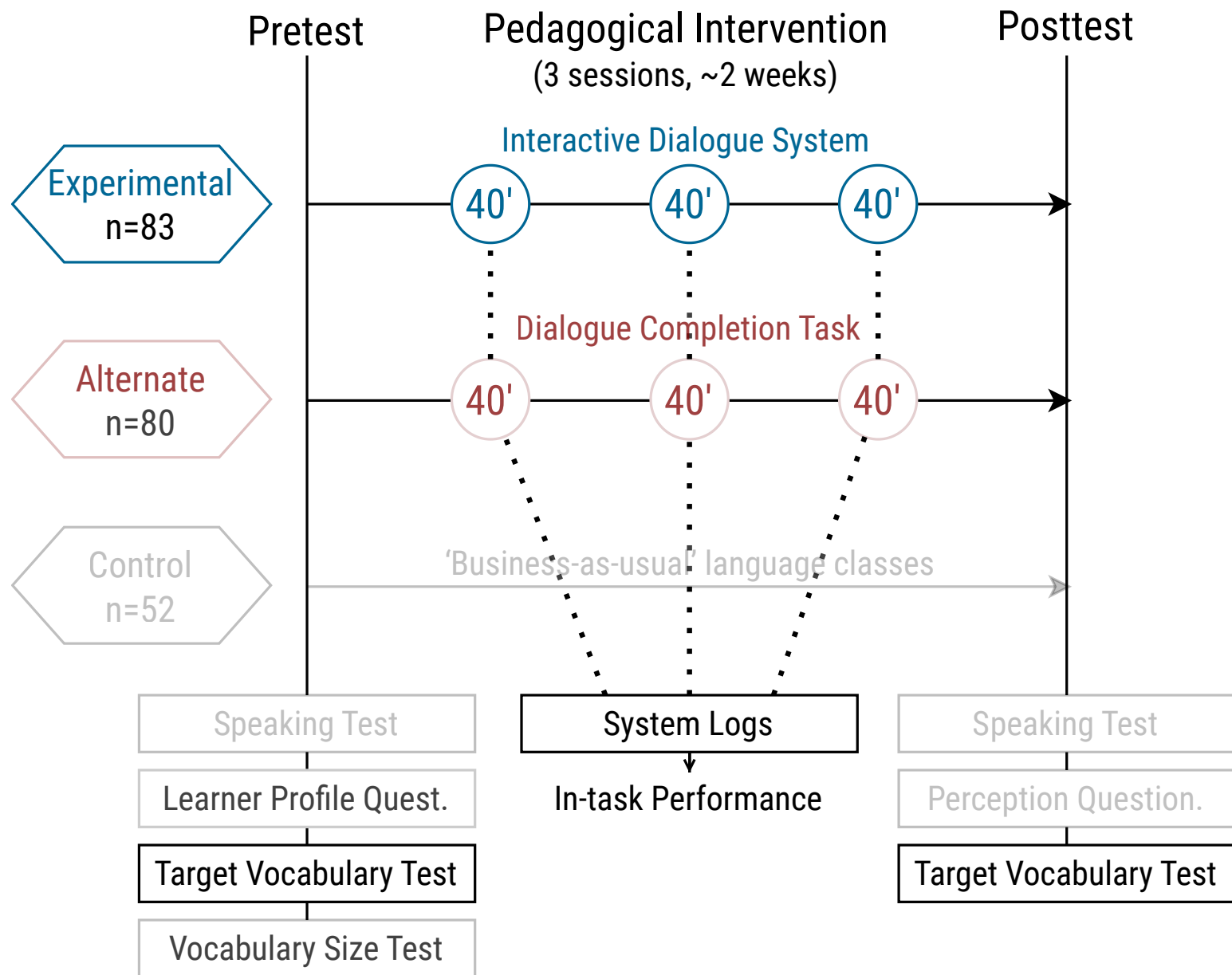
Ik weet het niet.



Allô ? Jean ? Il est 22h00 et tu n'es toujours pas à la maison. J'es _____ que tout va bien...

espère







Frequency in output and incidental vocabulary learning

Frequency of occurrence

Experiment: dialogue system

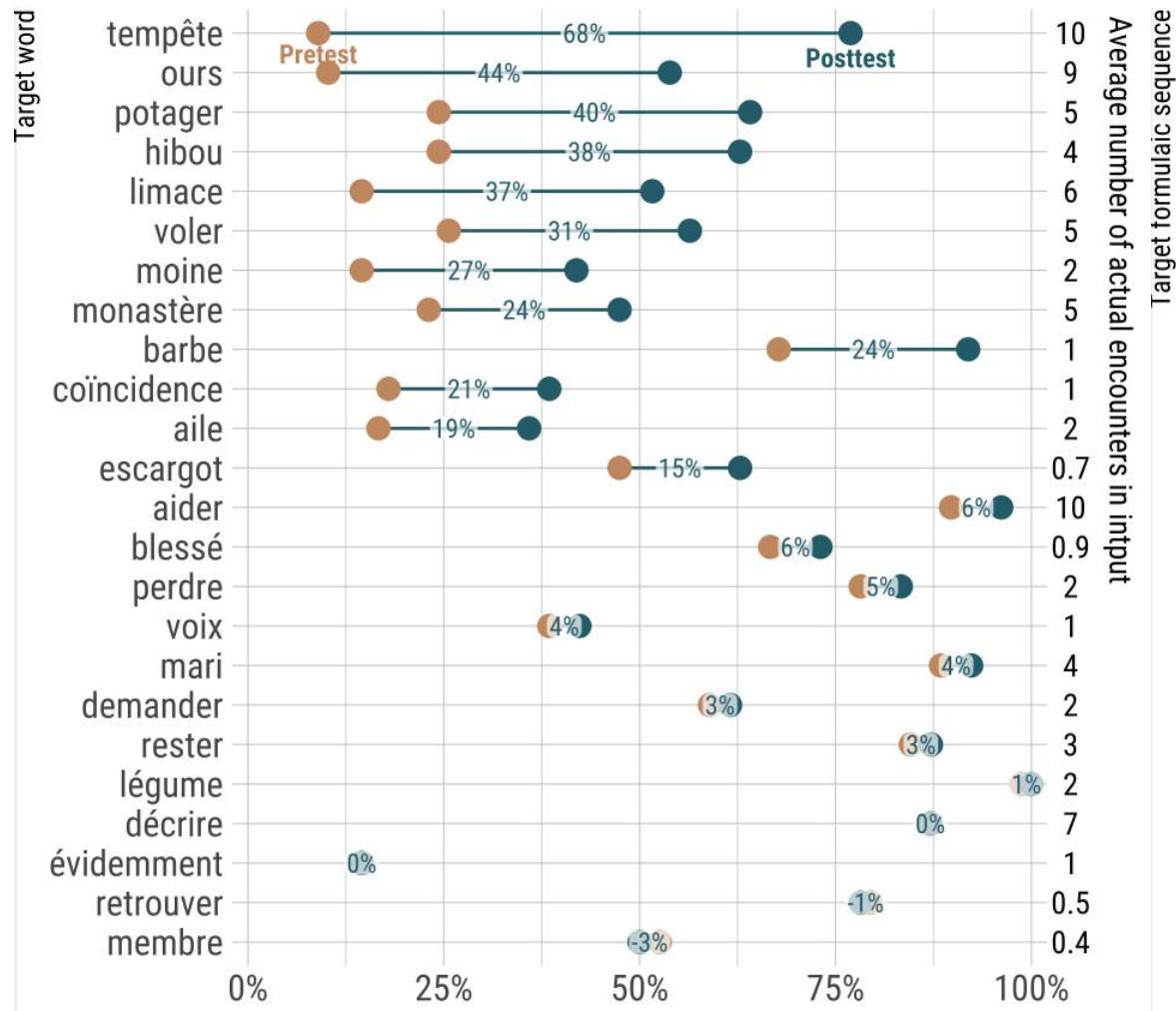
Results: incidental effects

Discussion & model

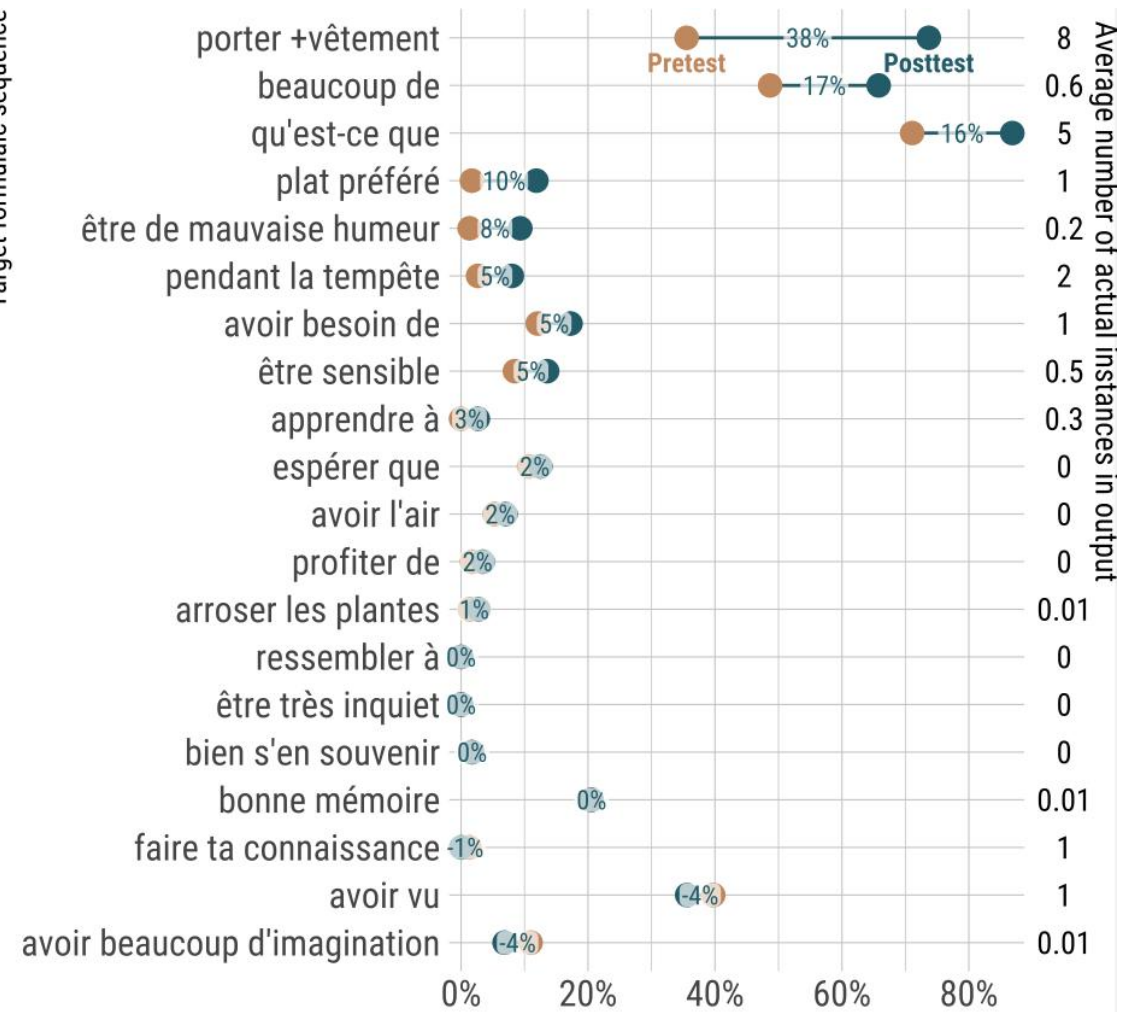


Large variations in learning gains

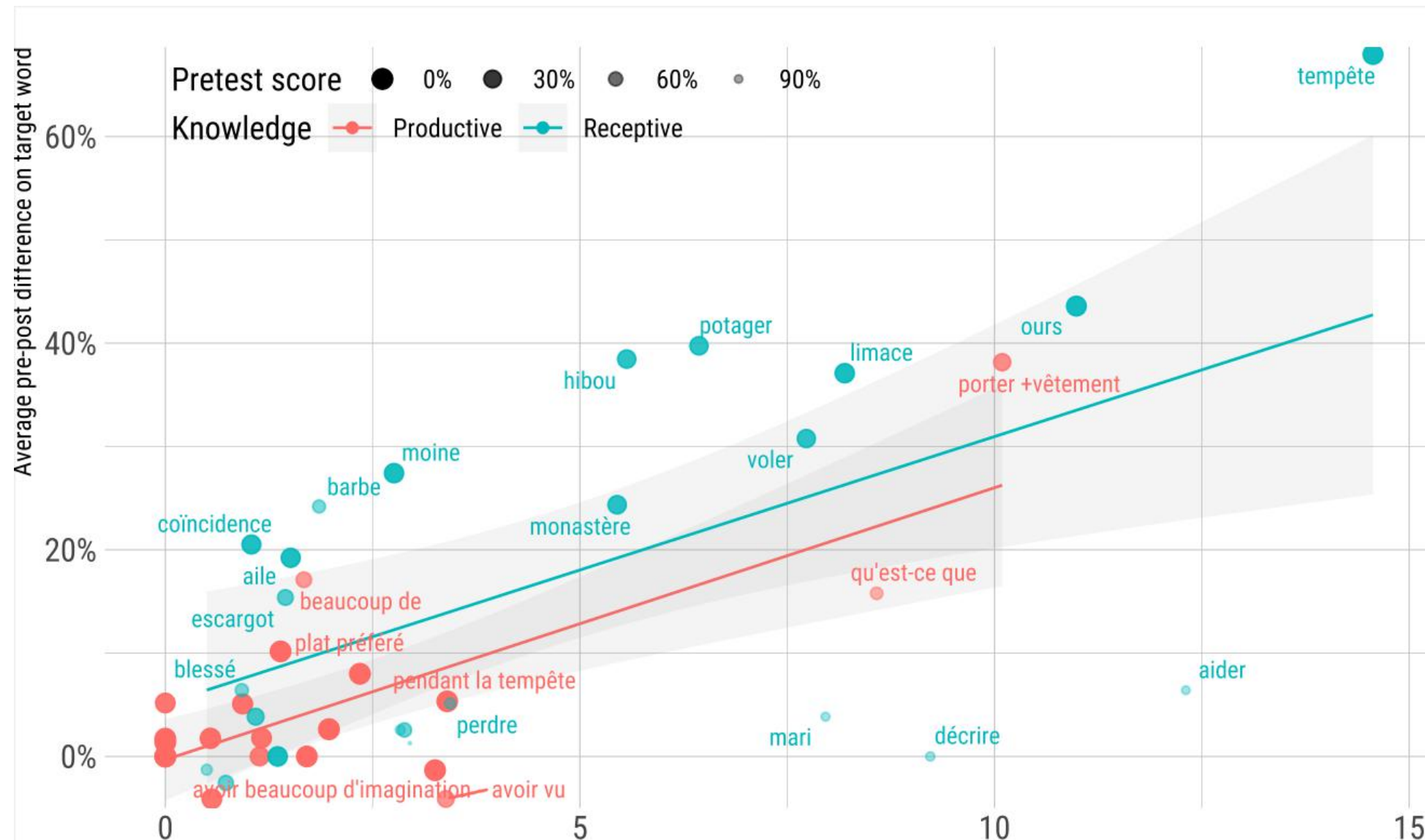
Receptive test



Productive test



Input encounters contribute to both receptive and productive knowl.



Mixed-effects model

$$R^2 = .66$$

FIXED EFFECTS	LOG-ODDS	SE	95% CI	<i>z</i>	<i>p</i>
(Intercept)	-2.56	0.35	[-3.25, -1.87]	-7.26	<.001 ***
Dialogue-based CALL ^d = 1	0.15	0.26	[-0.36, 0.67]	0.58	.563
Dialogue System ^e = 1	-0.18	0.21	[-0.60, 0.24]	-0.86	.393
Knowledge ^a = Productive	-1.49	0.30	[-2.07, -0.90]	-5.00	<.001 ***
<u>Frequency occurrence in input^b</u>	0.14	0.02	[0.09, 0.18]	6.01	<.001 ***
<u>Frequency occurrence in output^c</u>	0.26	0.03	[0.20, 0.32]	8.02	<.001 ***
Pretest score	2.34	0.09	[2.16, 2.53]	24.98	<.001 ***
Vocabulary size score	4.69	0.99	[2.75, 6.63]	4.74	<.001 ***
Knowledge ^a × Occurrence output ^c	-0.16	0.04	[-0.24, -0.07]	-3.67	<.001 ***
RANDOM EFFECTS	SD	GROUPING		<i>n</i>	ICC
Item	1.33	Items		44	.33
Subject:(Class:School)	0.47	Subjects		164	.04
Class:School	0.25	Classes		11	.01
School	0.21	Schools		4	.01
Residual	1.00				

Output instances best predict productive learning

OCCURRENCE COUNT	M	SD	$r_{\text{Receptive}}$	$r_{\text{Productive}}$
Hypothetical input encounters	5.82	5.63	.24 ***	.05 n.s.
Actual input encounters (dialogue)	2.55	3.15	.26 ***	.10 ***
Actual input encounters (dialogue + models)	3.58	4.51	.27 ***	.24 ***
Hypothetical output opportunities	2.05	2.76	.17 ***	.17 ***
Actual output opportunities (models)	1.03	2.19	.21 ***	.25 ***
Actual output instances (messages)	1.16	2.12	.25 ***	.28 ***

*** $p < .001$. n.s. = $p > .05$.

$r_{\text{Receptive}}$ = correlation with meaning recognition test.

$r_{\text{Productive}}$ = correlation with form recall test.



Frequency in output and incidental vocabulary learning

Frequency of occurrence

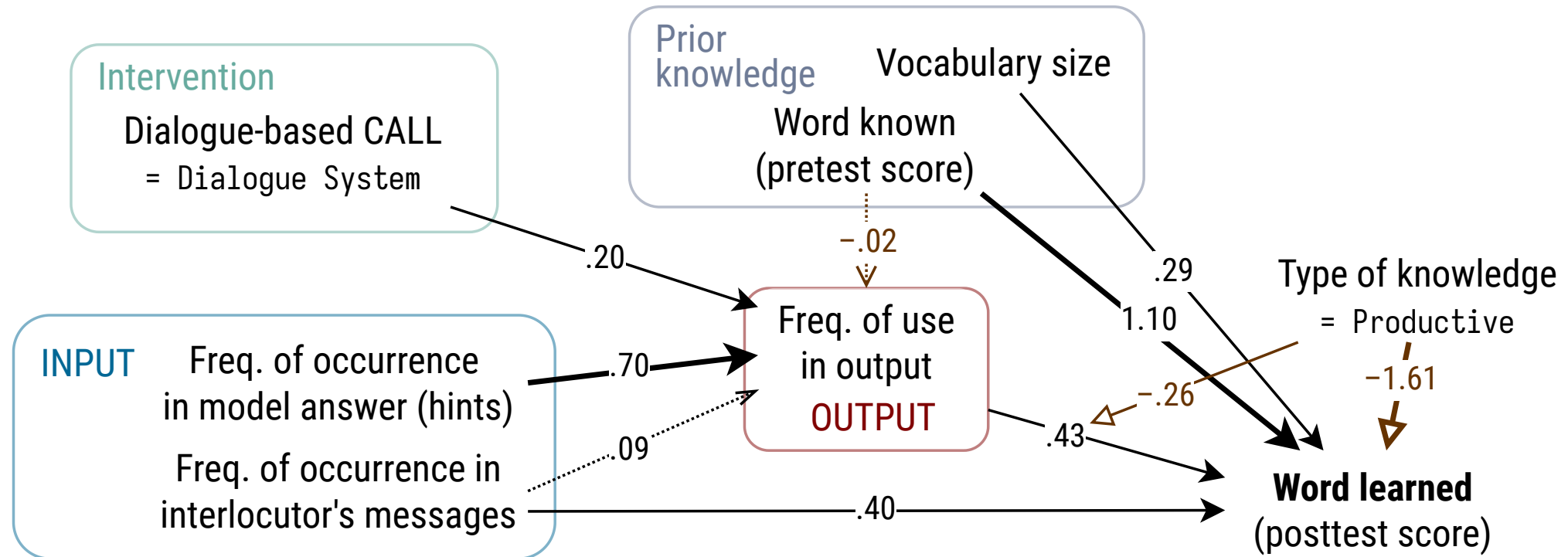
Experiment: dialogue system

Results: incidental effects

Discussion & model



Model of incidental productive vocabulary learning



Conclusions

- Productive exposure is complementary to receptive exposure.
- Confirms the need for productive practice for productive knowledge. (e.g., Yanagisawa & Webb, 2021)
- In summary:
 - Frequency in *input* \Rightarrow Word learned
 - but even more:
Frequency in *input* \Rightarrow Frequency in *output* \Rightarrow Word learned

Limitations

- **System:** story-first design
 - → explicit microtask prompts
 - → reduced user control
 - → limited difference across conditions (exp. condition less interactive than foreseen)
- **Instruments:**
 - *I don't know* option: probably better to avoid (Stoeckel et al., 2016)
 - many items with too limited/no exposure in the end
 - excessive difficulty of target vocabulary test (form recall)

Thanks! Questions & suggestions?

Serge Bibauw

serge.bibauw@uclouvain.be

Tw: @SergeBibauw

↓ Slides in PDF:

