1. ABOUT THE DATASET

------------

Title: Vocabulary learning through multimodal input

Creator(s): Pengchong Zhang (https://orcid.org/0000-0002-2136-4984)

Organisation(s): University of Reading

Rights-holder(s): University of Reading

Publication Year: 2025

Description: The dataset consists of second language learning data from 40 French learners of English (aged 11 to 12) at two time points. At Time 1, quantitative data were collected through a French vocabulary size test, a target vocabulary pre-test, and a phonological short-term memory test. At Time 2, quantitative eye-tracking data included average dwell time, average number of fixations, and average fixation duration for two types of Areas of Interest: one for non-verbal information and the other for verbal information. Additionally, quantitative test data were collected from a vocabulary post-test and a comprehension test.

Cite as: Zhang, Pengchong (2025): Vocabulary learning through multimodal input. University of Reading. Dataset. https://doi.org/10.17864/1947.001343

Related publication: Zhang, P., & Zhang, S. (2024). Multimedia enhanced vocabulary learning: the role of input condition and learner-related factors. System, 122, 103275. https://doi.org/10.1016/j.system.2024.103275

Zhang, P., & Zhang, S. (Under review). Attention and learning in L2 multimodality: A webcam-based eye-tracking study. Language Learning & Technology.

Contact: anthony.zhang@reading.ac.uk

2. TERMS OF USE

------------

This dataset is licensed under a Creative Commons Attribution 4.0 International Licence: https://creativecommons.org/licenses/by/4.0/.

3. PROJECT AND FUNDING INFORMATION

------------

Title: Improving foreign language in the UK: Transforming classroom language teaching through multimedia

Dates: 09/2022-09/2023

Funding organisation: University of Reading

Grant no.: A368700

4. CONTENTS

------------

File listing

Vocab\_multimodality.csv

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable name** | **Description** | **Unit** | **Value labels** |
| Rec\_Session\_Id | Participation session ID | Numeric | N/A |
| Subject | Participant ID | Numeric | N/A |
| Block\_Name | Experiment block | Binary | block 1 vs. block 2 |
| Time | Test time points | Binary | N/A |
| Task\_Name | Name of the task | String | N/A |
| Task\_Nr | Task number | Numeric | N/A |
| Trial\_Id | Trial ID | Numeric | N/A |
| Trial\_Nr | Trial number | Numeric | N/A |
| teaching\_method | Teaching methods | Categorical | control vs. content vs. paralinguistic |
| video\_clip | Name of the video clip | String | N/A |
| answer\_clip\_recognition | Video clip recognition | Numeric | N/A |
| answer\_clip\_comprehension | Video clip comprehension | Numeric | N/A |
| answer\_word\_used | Form recognition | Numeric | N/A |
| answer\_word\_recognition | Meaning recognition | Numeric | N/A |
| answer\_word\_translation | Meaning recall | Numeric | N/A |
| answer\_WM\_practice | Phonological short-term memory test practice | Numeric | N/A |
| answer\_WM\_formal | Phonological short-term memory formal test | Numeric | N/A |
| score\_clip\_recognition | Overall score for clip recognition | Numeric | N/A |
| score\_clip\_comprehension | Overall score for clip comprehension | Numeric | N/A |
| score\_word\_used\_PRE | Pretest form recognition | Numeric | N/A |
| score\_word\_used\_POST | Posttest form recognition | Numeric | N/A |
| score\_word\_recognition\_PRE | Pretest meaning recognition | Numeric | N/A |
| score\_word\_recognition\_POST | Posttest meaning recognition | Numeric | N/A |
| score\_word\_translation\_PRE | Pretest meaning recall | Numeric | N/A |
| score\_word\_translation\_POST | Posttest meaning recall | Numeric | N/A |
| score\_WM\_practice | Overall score for phonological short-term memory practice test | Numeric | N/A |
| score\_WM\_formal | Overall score for phonological short-term memory formal test | Numeric | N/A |
| score\_xlex | Overall score for vocabulary size | Numeric | N/A |
| condition\_post | Experimental conditions | Categorical | Written+Speaker-Video vs. Written+Audio+Picture vs. Written+Audio |
| mean\_fixtime\_nonverbal | Mean fixation time for non-verbal AOI | Numeric | N/A |
| mean\_fixtime\_verbal | Mean fixation time for verbal AOI | Numeric | N/A |
| mean\_dwelltime\_nonverbal | Mean dwelltime for non-verbal AOI | Numeric | N/A |
| mean\_dwelltime\_verbal | Mean dwelltime for verbal AOI | Numeric | N/A |
| mean\_fixnumber\_nonverbal | Mean number of fixations for non-verbal AOI | Numeric | N/A |
| mean\_fixnumber\_verbal | Mean number of fixations for verbal AOI | Numeric | N/A |
| mean\_firstfix\_nonverbal | Mean time to first fix to non-verbal AOI | Numeric | N/A |
| mean\_firstfix\_verbal | Mean time to first fix to verbal AOI | Numeric | N/A |
| mean\_dwelltime\_nv2v | Mean dwelltime non-verbal to verbal ratio | Numeric | N/A |
| centered\_mean\_dwelltime\_nv2v | Mean dwelltime non-verbal to verbal ratio z-score | Numeric | N/A |
| mean\_fixtime\_nv2v | Mean fixation time for non-verbal to verbal ratio | Numeric | N/A |
| centered\_mean\_fixtime\_nv2v | Mean fixation time for non-verbal to verbal ratio z-score | Numeric | N/A |
| mean\_fixnumber\_nv2v | Mean number of fixations for non-verbal to verbal ratio | Numeric | N/A |
| centered\_mean\_fixnumber\_nv2v | Mean number of fixations for non-verbal to verbal ratio z-score | Numeric | N/A |
| mean\_firstfix\_nv2v | Mean time to first fix to non-verbal to verbal ratio | Numeric | N/A |
| mean\_firstfix\_aoi | Mean time to first fix to AOI | Numeric | N/A |
| centered\_mean\_firstfix\_aoi | Mean time to first fix to AOI z-score | Numeric | N/A |

5. METHODS

-----------

Detailed information about methods is available in the related publication:

Zhang, P., & Zhang, S. (2024). Multimedia enhanced vocabulary learning: the role of input condition and learner-related factors. System, 122, 103275. https://doi.org/10.1016/j.system.2024.103275

- Experimental procedures/protocols

The study investigated the effects of different types of multimodal input on vocabulary learning. Participants were 40 young English learners of French who attended two experiment sessions. In session one, they completed a French vocabulary size test, a target vocabulary pre-test, and a phonological short-term memory test. In session two, they watched three sets of multimodal input, each representing one input condition: Written+Audio (Control), Written+Audio+Picture (Content), and Written+Speaker-Video (Paralinguistic). They completed a target vocabulary post-test and a comprehension test.

- Environmental/experimental conditions.

The study had three experimental conditions, each involving a different type of multimodal input:

1. Written+Audio (Control): verbal information about the target words

2. Written+Audio+Picture (Content): verbal plus static pictures demonstrating the meaning of the target words

3. Written+Speaker-Video (Paralinguistic): verbal plus speaker videos demonstrating the meaning of the target words

- Instruments used

The data were collected through Labvanced online experiment builder(https://www.labvanced.com), which features an online eye-tracking software.

- Methods used for processing the data

We adopted Bayesian mixed-effects models for data analysis, implemented in R (R Core Team, 2024) using the brms package (B√ºrkner, 2021).