



*Hiroshima Lexical Research Forum*

# **H-LRF 2025**

*September 14th & 14th, 2025*

## **Book of Abstracts**

# H-LRF 2025

## How to Participate

---

This year's H-LRF will be held on Saturday, September 13th and Sunday, September 14th. All of the day's of the conference will be held using the same Zoom session. You will be able to join the Zoom session by clicking on the link below during the conference times.

### Zoom Link:

<https://bit.ly/hlrf2025zoom>

### Full Zoom link

[https://us02web.zoom.us/j/86234382468?  
pwd=e93RN2GDQo79mkNEpz7aioRzcY4b9m.1](https://us02web.zoom.us/j/86234382468?pwd=e93RN2GDQo79mkNEpz7aioRzcY4b9m.1)



Click on the link or scan  
the QR Code to join the  
Zoom session.

### The conference schedule can be found here:

[https://h-lrf.org/files/2025/HLRF\\_2025\\_Schedule.pdf](https://h-lrf.org/files/2025/HLRF_2025_Schedule.pdf)

### Zoom Etiquette:

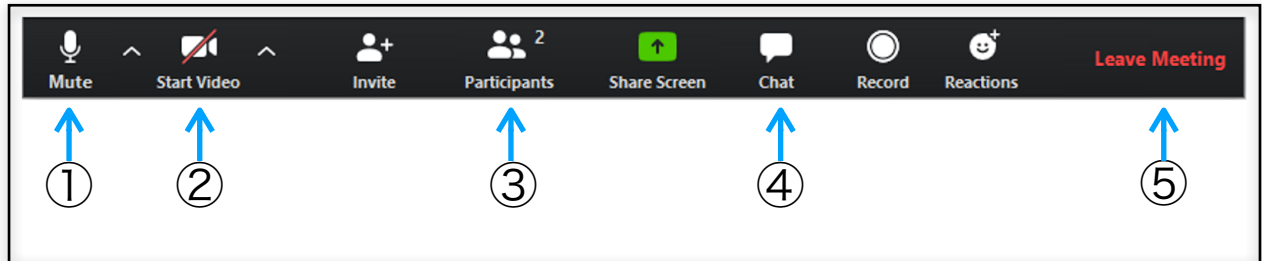
To ensure that everyone gets the most out of this year's conference, we have included a few requests regarding Zoom etiquette below.

- Please ensure that your microphone is muted when the presenter is talking
- We want the discussions to be as interactive as possible, so please feel free to use the Chat feature to ask and answer questions or make comments during the presentation.
- While the speaker will (probably) not be able to respond to your question during their presentation, we will have a question and answer session at the end of each talk.
- If you have a question, you would like to ask, please use the “raise your hand” feature of Zoom after the speaker has finished presenting.
- When asking your questions during the question and answer session, please ask your questions orally and ensure that both your microphone and camera are turned on.

## Using Zoom:

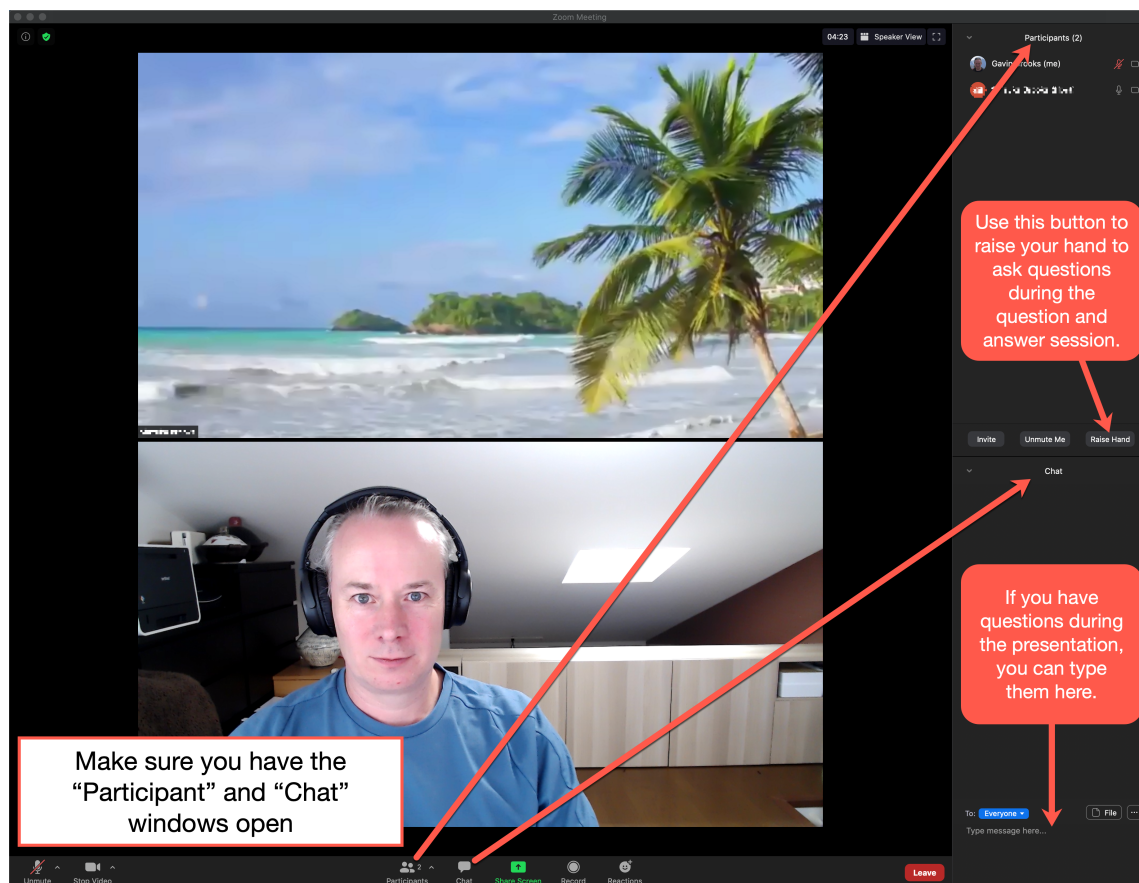
---

This year's H-LRF will be delivered using Zoom. While I am sure that everyone is very familiar with Zoom at this point in time, we have included a few simple instructions and requests to ensure that the conference runs smoothly.



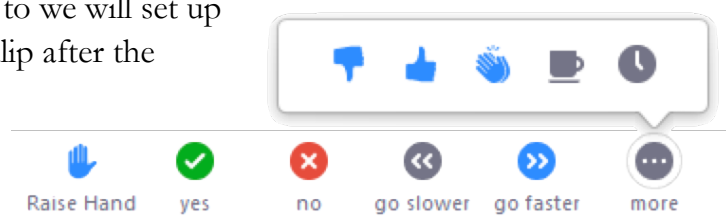
### When Joining:

- As soon as you login, make sure your video ② and microphone ① are turned on. To make things easier, we would also appreciate it if you could mute your microphone when the presenters are talking, as background noise from the microphones of audience members could cause Zoom to shift the focus away from the presenter.
- You should also open the participant ③ and chat windows ④ so that you can raise your hand if you have any questions or send a message to someone in the session.
- Make sure that your name is visible in the participant's window, as this is the name that the session chair will see when you are asking questions.
- If you need to leave the presentation, you can do so by clicking on Leave Meeting ⑤.



### Asking questions:

- Please ensure that your microphones are muted when the presenters are talking.
- You can ask questions during the presentation using the Chat feature of Zoom.
- We want these sessions to be as interactive as possible, so please feel free to respond to other audience members' comments or questions in the Chat box.
- After the presenter finishes, there will be time for questions and answers. These will be done orally. To ask a question:
- Raise your hand using the "Raise Hand" button in Zoom.
- The session chair will call on the audience members in the order in which they raised their hands.
- When your name is called, please turn on your microphone and ask your question to the presenter.
- Due to time constraints, we may not get to all of the questions. If you have a question that you wanted to ask but were not able to we will set up a question channel in the H-LRF Zulip after the conference and continue the discussion there.





## Multiword Units in L2 Vocabulary Testing: A Critical Examination

**Presenter:** Ayako Aizawa, *Rikkyo University*

**Contributing authors:** Kohei Fukuda, *Tokyo University of Foreign Studies*; Yusuke Kuroki, *Michigan State University*

The use of multiword units (MWUs) has been considered an indicator of second language (L2) proficiency (e.g., Boers et al, 2006; Tavakoli & Uchihara, 2020). Over the past decades, a growing number of studies have focused on MWUs, leading to the development and use of various measures to assess L2 learners' knowledge of MWUs (e.g., Gyllstad, 2020). However, the field has yet to reach a consensus on how best to assess MWUs, and no assessment has been widely accepted as the standard (Boers & Lindstromberg, 2009; Gyllstad & Schmitt, 2018). This current situation suggests a lack of comprehensive understanding regarding the test design, what the tests actually measure, and their validity and reliability. To address this gap, this study will critically examine existing MWU tests. This examination will focus on their design and administration, and the reported evidence for their validity and reliability, which we will analyze through a thorough review of the accompanying literature for each test. In addition, as an extension of our critical review, we will conduct a corpus-driven analysis of the linguistic features. This will involve employing consistent association measures to analyze the representativeness and commonality of MWUs across different tests. By integrating the test items into a unified corpus, we aim to uncover the extent to which these tests differentiate their assessment of MWU knowledge, addressing concerns about potential redundancy and the ecological validity of the tested items. This research aims to contribute to a more comprehensive understanding of current MWU test characteristics and the scope of their supporting documentation, highlighting areas for improvement. Our findings are expected to inform the design and reporting of future MWU tests, enabling test users to more effectively select instruments that align with their specific measurement goals.

### References

- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, 10(3), 245–261. <https://doi.org/10.1191/1362168806lr195oa>
- Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition*. Palgrave Macmillan UK.
- Gyllstad, H. (2020). Measuring knowledge of multiword units. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 387–405). Routledge.

- Gyllstad, H., & Schmitt, N. (2018). Testing formulaic language. In A. Siyanova-Chanturia, & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 174–191). Routledge.
- Tavakoli, P., & Uchihara, T. (2020). To what extent are multiword sequences associated with oral fluency? *Language Learning*, 70(2), 506–547. <https://doi.org/10.1111/lang.12384>

## Can Lexical Diversity Measures Predict L2 Attrition in Collocational Knowledge?

**Presenters:** Hadil Alraddadi, *University of Reading*; Jeanine Treffers-Daller, *University of Reading*

**Contributing author:** Scott Jarvis, *Northern Arizona University*

This study aims to establish to what extent sophisticated measures of lexical diversity (LD) outperform traditional measures in the analysis of between-group differences in L2 English attrition. We also compare the performance of LD measures with a standardized vocabulary test (PPVT, Dunn & Dunn, 2007) in identifying these differences. Finally, we explore how well LD measures and the PPVT can predict accuracy and reaction times on tasks measuring receptive and productive knowledge of verb-noun English collocations. Participants included 44 adult returnees who had returned to their country of origin (Saudi Arabia) at different ages in childhood and 20 adult heritage speakers who continued to live in the United States. We found that LD measures could reveal existing overall between-group differences in L2 attrition, although the amount of explained variance was low by comparison with analyses which used the PPVT as a predictor. Analyses using mixed effects modelling, indicate that LD measures may not good predictors of respondents' ability to process collocations. Instead, a direct measure of vocabulary size, specifically the PPVT, emerged as a more robust predictor of collocational processing.

### References

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test–IV (PPVT–IV)* [Measurement instrument]. Circle Pines, MN: American Guidance Service.

# AI-Augmented Corrective Feedback: Effects on Multiword Expressions in Writing

**Presenters:** Vahid Asadi, *University of Milan*; Mojtaba Tadayonifar, *Victoria University of Wellington*

Written Corrective Feedback (WCF) plays a significant role in addressing errors and enhancing learners' linguistic abilities, and it remains a crucial component in second language (L2) writing instruction (Balla et al., 2025; Hongxia & Razali, 2025). However, delivering detailed and personalized feedback can be labour-intensive and resource-demanding, especially in large classroom settings (Mulenga & Shilongo, 2025). Recent advancements in artificial intelligence, particularly the advent of large language models (LLMs) like ChatGPT, offer promising avenues for automating and supplementing traditional feedback mechanisms (Gao et al., 2023). Prior research indicates that AI-mediated feedback supports learners' linguistic development by offering timely corrections that complement teacher input. Hybrid approaches combining human and AI feedback have shown promise in improving outcomes and motivation. However, studies exploring effects of AI-driven feedback on lexical diversity in learners' writing remain limited. To bridge this gap, the current study investigates the effects of hybrid and teacher-only feedback approaches on the use of multiword expressions in learners' writing. Forty adults Iranian EFL learners are randomly assigned to two groups: an experimental group (n=20) receiving hybrid feedback and a control group (n=20) receiving only teacher feedback. Data collection includes pre-test, post-test, and delayed post-test writing assessments. This research contributes to strategies for integrating AI into language education, enhancing learner engagement and scalable feedback mechanisms.

## References

- Balla, A. A. S., AbdAlgane, M., Ahmed, A. O. A., & Osman, E. (2025). AI-driven innovations in adult EFL learning: Exploring potentials and practicalities. *International Journal of Interactive Mobile Technologies*, 19(8). <https://doi.org/10.3991/ijim.v19i08.52295>
- Hongxia, H., & Razali, A. B. (2025). Impact of ChatGPT on English academic writing ability and engagement of Chinese EFL undergraduates. *International Journal of Instruction*, 18(2), 323–346. <https://e-iji.net/ats/index.php/pub/article/view/738>
- Mulenga, R., & Shilongo, H. (2025). Hybrid and blended learning models: Innovations, challenges, and future directions in education. *Acta Pedagogica Asiana*, 4(1), 1–13. <https://doi.org/10.53623/apga.v4i1.495>

## Impact of Bilingual Glosses and Reading Activities on Vocabulary Acquisition

**Presenter:** Xiaoke Bai, *Faculty of Education, University of Macau*

**Contributing author:** Barry Lee Reynolds, *Faculty of Education, University of Macau, Macau SAR, China; Centre for Cognitive and Brain Sciences, University of Macau, Macau SAR, China*

The effectiveness of utilizing first language (L1) glosses versus second language (L2) glosses in L2 reading and vocabulary acquisition remains a topic of ongoing academic debate. Building upon it, this study employs a between-subject design to examine the effects of bilingual glosses—comprising both L1 and L2 glosses—on incidental vocabulary learning through different reading activities from the lens of bilingual education. It considers several factors that have been insufficiently investigated in prior glossing studies, including the processing of glosses, differences in gloss processing across different reading activities, and the role of learners' prior vocabulary knowledge. Fifty-one L2 learners participated in three randomly assigned reading activities: a reading-only activity (n=16), a meaning-oriented reading activity (n=16), and a form-oriented reading activity (n=17), along with a control group (n=24). Participants accessed bilingual glosses by clicking on target words, while their eye movements were tracked to measure the time spent on and the frequency of visits to each gloss. To gauge vocabulary learning and retention, participants completed pre-tests, immediate post-tests, and delayed post-tests assessing their vocabulary knowledge of receptive form, productive meaning, and receptive meaning. The results revealed complex interactions among gloss processing, reading activities, and vocabulary acquisition. Additionally, learners' prior vocabulary knowledge was found to independently affect both vocabulary learning and retention. These findings highlight the importance of bilingual glosses, considering factors related to reading conditions and individual learner differences. This study contributes to the contentious discussion about effective glosses for L2 vocabulary learning, emphasizing the potential benefits of integrating both L1 and L2 glosses in educational settings.

# Enhancing Learner Corpus Creation through LLMs and Natural Language Processing

**Presenter:** Gavin Brooks, *Kyoto Sangyo University*

While recent studies have explored how learner corpora can help teachers develop materials that meet learners' needs (Brezina et al., 2022), there remains a need for targeted learner corpora that provide insights into specific groups of learners (Götz & Granger, 2024). However, cleaning a corpus for analysis is time-consuming (Gablasova et al., 2019). This presentation demonstrates an LLM-powered corpus cleaning workflow that can help address this by using advances in LLMs and NLP tools like spaCy and Stanza to streamline the process.

Our approach addresses this by integrating LLMs with NLP libraries to identify spelling errors, classify words (e.g., proper nouns, technical terms, foreign words), and apply structured markup. By leveraging API-based processing from modern LLMs like Claude or ChatGPT, this approach allows these LLMs to assist with the systematic analysis and cleaning of a corpus.

This presentation showcases the workflow in action. By using a subset of texts from our existing learner corpus, along with a cleaned and annotated gold-standard version of these texts, we will illustrate how LLMs facilitate preprocessing and structuring learner corpora. The results suggest that this method enhances efficiency and consistency, allowing researchers to focus on linguistic analysis rather than data cleaning.

## References

- Brezina, V., Gablasova, D., & McEnery, T. (2022). Corpus-based approaches to spoken L2 production: Evidence from the Trinity Lancaster Corpus. *International Journal of Learner Corpus Research*, 5(2), 119–125. <https://doi.org/10.1075/ijlcr.00008.int>
- Gablasova, D., Brezina, V., & McEnery, T. (2019). The Trinity Lancaster Corpus: Development, description and application. *International Journal of Learner Corpus Research*, 5(2), 126–158. <https://doi.org/10.1075/ijlcr.19001.gab>
- Götz, S., & Granger, S. (2024). Learner corpus research for pedagogical purposes: An overview and some research perspectives. *International Journal of Learner Corpus Research*, 10(1), 1–38. <https://doi.org/10.1075/ijlcr.00039.got>



## Assessing Collocations in Global Gateway Textbooks: A Corpus-Based Comparison

**Presenter:** Chi Cuong Chau, *The University of Macau*

This study investigates the pedagogical relevance of lexical collocations (LCs) in the Global Gateway (GG) textbook series for Grades 10 and 11, which are aligned with Vietnam's 2018 General Education Programme. Drawing on a corpus-based approach, it addresses persistent concerns regarding the limited representativeness of textbook language and the inconsistent application of statistical Association Measures (AMs) in collocation research. A specialised corpus (29,522 tokens) was compiled from the GG series, and LCs were extracted using AntConc and manually refined following phraseological principles. Four AMs, including MI-score, t-score, Log-Likelihood, and Log-Dice, were applied to identify statistically salient collocations based on the British National Corpus. This refinement process reduced 4,197 raw collocations to 693 statistically salient entries.

To validate the pedagogical appropriateness of these entries, expert judgment was incorporated from two experienced EFL professionals using a five-point Likert scale. Results showed that over 93% of collocations were rated "very appropriate," with an Intraclass Correlation Coefficient of 0.704, indicating good inter-rater reliability. The findings revealed that Adjective-Noun and Verb-Noun collocations were the most prevalent and statistically robust. However, most LCs occurred only once in the textbooks, suggesting a lack of recycling, while others were overused in specific units. The study underscores the importance of balancing frequency, diversity, and context in collocation presentation.

This research contributes to both corpus linguistics and applied pedagogy by demonstrating how combining multiple AMs with expert review enhances the identification of pedagogically valuable collocations. It offers theoretical implications for collocational strength and lexical diversity, and practical guidance for textbook writers and language instructors. Future research is recommended to expand the corpus scope, diversify expert panels, and explore cultural and semantic prosody dimensions in collocation use.

# Developmental Trajectories of Multiword Expressions in Mandarin School-Age Children's Narratives

**Presenters:** Zixing Fan, *Victoria University of Wellington*; Anna Siyanova-Chanturia, *Victoria University of Wellington*

**Contributing author:** Shuhui Huang, *Ma Gang Fourth Primary School*

Narrative competence is increasingly recognized as central to children's language and literacy development, with multiword expressions (MWEs) playing a key role in fluent and proficient language use (Christiansen & Arnon, 2017; Gillam et al., 2023; Jiang et al., 2020). Based on the usage-based theory of language acquisition, which views linguistic knowledge as emerging from repeated exposure to and use of language patterns (Ellis & Ogden, 2017; Wulff, 2018), this study focuses on how children acquire MWEs through authentic narrative production. In the field of Mandarin research, many studies have already highlighted the importance of written narratives in revealing children's expressive language ability and the key role of MWEs in this process (Ding et al., 2023; Jiang et al., 2020; Xue et al., 2022). However, little is known about the longitudinal development of MWEs in Mandarin-speaking school-age children's written narratives.

To this aim, this study tracks 74 children from Grades 2-5 across four time points over two years, analyzing their use of adjective-noun and verb-noun phrases in narrative writing produced in standardized language tests. Mixed-effects models are employed to investigate the extent to which phrase frequency, association strength (mutual information), lexical gravity, Delta P forward, Delta P backward vary as a function of time, grade, gender, the frequency and length of the first and second words, lexical diversity (Type-Token Ratio, vocD, Guiraud index, MTLTD), and the interactions among these predictors.

This research addresses a key gap in understanding how children acquire, develop, and employ Mandarin multiword expressions over time in narrative production. The findings will provide a foundation for more inclusive, discourse-based approaches to studying their expressive language development. In addition, the findings will offer first normative benchmarks, informing educators and clinicians about expected growth patterns in children's usage of MWEs, supporting the development of assessment tools and evidence-based interventions.

## References

Christiansen, M. H., & Arnon, I. (2017). *More than words: The role of multiword sequences in language learning and use*. In (Vol. 9, pp. 542-551): Wiley Online Library.

- Ding, W., Wu, M., Ye, Z., & He, W. (2023). Narrative Development of Mandarin-Speaking Children in Grades 2–5 at a Mainland Chinese Primary School. *Reading & Writing Quarterly*, 1-15. <https://doi.org/10.1080/10573569.2023.2165204>
- Ellis, N. C., & Ogden, D. C. (2017). Thinking about multiword constructions: Usage-based approaches to acquisition and processing. *Topics in Cognitive Science*, 9(3), 604-620. <https://doi.org/10.1111/tops.12256>
- Gillam, S. L., Vaughn, S., Roberts, G., Capin, P., Fall, A.-M., Israelsen-Augenstein, M., Holbrook, S., Wada, R., Hancock, A., & Fox, C. (2023). Improving oral and written narration and reading comprehension of children at-risk for language and literacy difficulties: Results of a randomized clinical trial. *Journal of Educational Psychology*, 115(1), 99. <https://doi.org/10.1037/edu0000766>
- Jiang, S., Jiang, X., & Siyanova-Chanturia, A. (2020). The processing of multiword expressions in children and adults: An eye-tracking study of Chinese. *Applied Psycholinguistics*, 41(4), 901-931. <https://doi.org/10.1017/S0142716420000296>
- Wulff, S. (2018). Acquisition of Formulaic Language from a Usage-Based Perspective. In A. Siyanova-Chanturia & A. Pellicer-Sanchez (Eds.), *Understanding formulaic language: A second language acquisition perspective* (pp. 19-37). Routledge.
- Xue, J., Zhuo, J., Li, P., Liu, J., & Zhao, J. (2022). Characterizing macro-and micro-structures of narrative skills for Mandarin-speaking school-age children with specific language impairment. *Journal of Communication Disorders*, 96, 106199. <https://doi.org/10.1016/j.jcomdis.2022.106199>

# A Study Investigating Word Association Behaviour in People with Acquired Language and Communication Disorders

**Presenter:** Angela Maria Fenu, *Swansea University*

While extensive research on word association has advanced understanding of the mental lexicon in neurotypical individuals, considerably less attention has been paid to populations with language impairments, such as aphasia. This study contributes to closing that gap by examining word association behaviour in individuals with Broca's aphasia, a non-fluent syndrome marked by effortful speech and agrammatism. Building on earlier work, the current research investigates how aphasia influences lexical retrieval and the nature of associative responses, with particular attention to grammatical class, concreteness, abstractness and word frequency. The research consists of three experiments. The first was a conceptual replication of Gewirth et al. (1984), adapted to the linguistic and clinical profiles of the present sample. Data from five test sessions were analysed to determine whether changes in response types indicated therapeutic progress and to assess individual variability. Findings were compared to those from non-aphasic controls to highlight similarities and differences in associative patterns. The second experiment, based on the same participant cohort, investigated whether changes in word association responses over five testing sessions reflected gradual therapeutic progress. It also assessed whether individuals with comparable aphasic profiles and severity levels demonstrated consistent patterns of associative behavior, or whether individual differences emerged as a significant factor. The third experiment, currently ongoing, involves a larger participant sample and employs a newly developed set of 100 cue words, constructed according to specific linguistic criteria. This investigation seeks to identify which aspects of word association remain intact in aphasia and which are most vulnerable, aiming to inform clinical diagnosis and intervention. The results hold promise for enhancing understanding of lexical-semantic impairments in aphasia and for guiding targeted speech-language therapy.

## References

- Fitzpatrick, T., & Izura, C. (2011). Word association in L1 and L2: An exploratory study of response types, response times and inter-language mediation. *Studies in Second Language Acquisition*, 33(3), 373–398. <https://doi.org/10.1017/S0272263111000027>
- Fitzpatrick, T., Playfoot, D., Wray, A., & Wright, M. J. (2015). Establishing the reliability of word association data for investigating individual and group differences. *Applied Linguistics*, 36(1), 23–50. <https://doi.org/10.1093/applin/amt020>

- Fitzpatrick, T., & Thwaites, P. (2020). Word association research and the L2 lexicon. *Language Teaching*, 53(2), 180–200. <https://doi.org/10.1017/S0261444820000105>
- Gewirth, L. R., Schindler, A. G., & Hier, D. B. (1984). Altered patterns of word associations in dementia and aphasia. *Brain and Language*, 21(2), 307–317. [https://doi.org/10.1016/0093-934X\(84\)90084-4](https://doi.org/10.1016/0093-934X(84)90084-4)
- Gollan, T. H., Salmon, D. P., & Paxton, J. L. (2006). Word association in early Alzheimer's disease. *Brain and Language*, 99(3), 289–303. <https://doi.org/10.1016/j.bandl.2005.07.003>

## AI-Assisted Corpus-Based Approach: Formulaic Language and L2 Oral Proficiency

**Presenter:** Akihiro Kotoh, *Victoria University of Wellington*

Formulaic language (FL) has been a central topic in second-language (L2) speaking research, and its pivotal role in L2 oral production has been well documented (Hougham, Clenton, Uchihara, & Higginbotham, 2024). While many corpus-driven studies of lexical bundles have explored the relationship between FL use and L2 oral proficiency, large-scale, corpus-based analyses of overall FL usage patterns remain limited, due to methodological challenges such as compiling look-up lists for target FL and conducting extensive manual data collation (Moon, 1998).

To address these issues, the present study leveraged two AI large-language models. First, Claude Sonnet 4 assisted in translating 735 items from the Oxford Phrase List (OPL) into Corpus Query Language (CQL). Second, ChatGPT-4o assisted in collating 48,344 concordance lines into a confusion matrix of 1,301 speakers by 575 OPL items, extracted from the NICT JLE Corpus (Izumi, Uchimoto, & Isahara, 2004), which comprises oral-proficiency interviews with 1,281 Japanese EFL learners and 20 native speakers.

Using both type coverage and relative token frequency of the identified phrases, a Random Forest classifier with repeated  $10 \times 5$  cross-validation distinguished five proficiency levels with 72.56 % accuracy (95 % CI = 71.8 – 73.3 %). Finally, a lexical-progression analysis pinpointed mismatches between the assumed CEFR levels of OPL items and their actual usage by speakers in the NICT JLE Corpus.

Overall, type coverage and relative token frequency of FL serve as strong indicators of oral proficiency, highlighting a clear relationship between FL use and L2 speaking ability. Furthermore, the resulting inventory—linking AI prompts and CQL, per-item frequencies for Japanese EFL learners versus native speakers, and CEFR-level gaps—offers a valuable resource as: (1) AI-assisted corpus-based research tools, (2) psycholinguistic item pools, and (3) pedagogical materials, aligned with the OPL's original intent but enriched by authentic learner and native usage data, to inform teachability and support valid L2 proficiency assessment.

### References

Hougham, D., Clenton, J., Uchihara, T., & Higginbotham, G. (2024). The impact of lexical bundle length on L2 oral proficiency. *Languages*, 9(7), 232. <https://doi.org/10.3390/languages9070232>



- Izumi, E., Uchimoto, K., & Isahara, H. (2004). *The NICT JLE Corpus (Version 4.1)* [Data set]. National Institute of Information and Communications Technology. [https://alaginrc.nict.go.jp/nict\\_jle/](https://alaginrc.nict.go.jp/nict_jle/)
- Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford University Press.

# Acquiring Productive Collocation Knowledge: The Effects of Teaching Method and Incongruence Type

**Presenters:** Batia Laufer, *University of Haifa*; Ronit Breslaw, *University of Haifa* and *The Hebrew University of Jerusalem*

Research indicates that learners, even advanced ones, have difficulties with productive knowledge of collocations, especially L1-L2 incongruent ones (e.g., Saito & Liu, 2022; Laufer & Waldman, 2011; Peters, 2016). To address this challenge, form-focused activities have been recommended for collocation learning as they yield better results than exposure through input (Peters, 2012; Szudarski, 2012). Even though the most effective form-focused tasks require output, retrieval, and the use of L1 (Laufer & Girsai, 2008; Laufer & Hulstijn, 2001; Webb & Nation, 2011), not many studies have investigated such tasks.

We investigate how the acquisition of productive knowledge of incongruent collocations achieved through pushed output tasks is affected by two factors: 1) teaching method (L2 only vs L1-L2 contrastive analysis) and 2) collocation incongruence type (lexical vs syntactic). We define lexically (LEX) incongruent collocations as collocations with a similar structure in L1 (Hebrew), but a different collocate, e.g., 'follow instructions' ('obey instructions' in Hebrew). Syntactically (SYN) incongruent collocations have different phrase structures in L1 and L2 and are translated by a syntactic paraphrase in L1, e.g. 'ridiculously expensive' (expensive in a ridiculous way' in Hebrew).

Seventy-five EFL university students studied sixty collocations, thirty lexically incongruent and thirty syntactically incongruent, in four output activities. Thirty were taught and practiced via L2, and thirty via L1-L2 contrastive analysis. In a counterbalanced within-subject design, each student studied fifteen different collocations in four conditions: LEX/L2, LEX/L1-L2, SYN/L2, SYN/L1-L2. Learning was assessed by immediate and delayed form recall posttests. A Mixed Effects analysis of the results revealed effects for incongruence type and the teaching method with the LEX/L1-L2 condition showing the highest recall and SYN/L2 the lowest. We relate the results to L1-L2 awareness raising, Pushed Output Hypothesis, Involvement Load Hypothesis, and Technique Feature Analysis.

## References

- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29(4), 694-716. <https://doi.org/10.1093/applin/amn018>
- Laufer, B., & Hulstijn, J.H. (2001). Incidental vocabulary acquisition in a second language. The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26. <https://doi.org/10.1093/applin/22.1.1>

- Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners. *English Language Learning*, 61(2), 647–672. <https://doi.org/10.1111/j.1467-9922.2010.00621.x>
- Peters, E. (2012). Learning German formulaic sequences: the effect of two attention-drawing techniques. *The Language Learning Journal*, 40(1), 65-79. <https://doi.org/10.1080/09571736.2012.658224>
- Peters, E. (2016). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research*, 20(1), 113–138. <https://doi.org/10.1177/1362168814568131>
- Saito, K., & Liu, Y. (2022). Roles of collocation in L2 oral proficiency revisited: Different tasks, L1 vs. L2 raters, and cross-sectional vs. longitudinal analyses. *Second Language Research*, 38(3), 531–554. <https://doi.org/10.1177/0267658320988055>
- Szudarski, P. (2012). Effects of meaning-and form-focused instruction on the acquisition of verb-noun collocations in L2 English. *Journal of Second Language Teaching & Research*, 1(2), 3–37. <https://pops.uclan.ac.uk/index.php/jsltr/article/view/32>
- Webb, S., & Nation, I. S. P. (2011). *Researching and analyzing vocabulary*. Boston: Heinle Cengage Learning

# The Effectiveness of Semantic Clustering on Vocabulary Learning: A Meta-Analysis

**Presenters:** Muhang Li, *University of Oxford*; Takumi Uchihara, *Tohoku University*

**Contributing author:** Tatsuya Nakata, *Rikkyo University*

How to organize and introduce novel vocabulary to L2 learners in a more effective manner has sparked intense discussion. The semantic clustering method, which involves grouping words based on their semantic relatedness/categories (Nation & Webb, 2011), has been witnessed as a prevalent practice of presenting words in both L2 classrooms and textbooks (McDonald & Reynolds, 2023). For example, language teachers may introduce words: hippo, tiger, snake and panda together, whose meanings all fall under one superordinate concept "animals". Recognizing the potential pedagogical value of the semantic clustering method and theoretically motivated by semantic field theory, scholars have invested extensive research efforts into investigating its efficiency in enhancing vocabulary learning outcomes, but generated highly contradictory results. The absence of systematic reviews and meta-analyses synthesizing the findings of this line of research motivated the current meta-analytic study to investigate the overall effectiveness of semantic clustering, as well as the potential moderators that might help account for the inconsistency of previous findings. A three-level meta-analytic model was fitted to the dataset of 51 effect sizes obtained from 28 eligible primary studies. The result of the overall effect reveals a statistically significant detrimental effect of semantic clustering in the initial learning stage, with a medium-to-large effect size, as measured by trials-to-criterion metrics, in line with the predictions of the Distinctiveness Hypothesis and Interference Theory. However, such an adverse effect does not seem to persist, and no significant difference was found in the aggregated effect size of studies with posttest designs. The subsequent moderator analysis examined six potential moderators to account for the within- and between-studies heterogeneity. The significance of this meta-analysis lies in advancing our understanding of the overall effectiveness of presenting vocabulary in a semantically related manner, so that research-informed suggestions could be provided to material developers and language teachers.

## References

- McDonald, J. A., & Reynolds, B. L. (2023). Learning semantic and thematic vocabulary clusters through embedded instruction: Effects on very young English learners' vocabulary acquisition and retention. *Applied Linguistics Review*, 14(5), 1129–1156. <https://doi.org/10.1515/applirev-2020-0102>
- Nation, I. S. P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle: Cengage Learning.

## Word Part Technique: Roles of Prior Vocabulary and Semantic Transparency

**Presenters:** Gaia Oikawa, *Toboku University*; Masato Sugawara, *Toboku University*

**Contributing author:** Takumi Uchihara, *Toboku University*

To master adequate vocabulary for L2 use, learners must employ effective vocabulary learning strategies that facilitate deeper lexical processing. This study examined the effectiveness of the word part technique (WPT), a morphological mnemonic that strengthens form–meaning mapping through knowledge of affixes and stems (Nation, 2022). Through the WPT, learners leverage their morphological knowledge (i.e., affixes and stems) for learning new words. Although some previous studies have explored the effectiveness of WPT (e.g., Suzuki, 2022; Wei, 2015), they have not examined the influences of marginal learner-related (i.e., vocabulary size, word part knowledge) and item-related variables (i.e., semantic transparency). Thus, this study attempted to bridge the gap. Forty-eight Japanese university EFL learners used the WPT to study 16 low-frequency English words, while rating each word's semantic transparency. Learners' vocabulary size and word part knowledge were also measured via VST-NJ8 (Hamada et al., 2021) and the word parts levels test (Sasao & Webb, 2017). Meaning recall and meaning recognition of the target words were assessed immediately and two weeks later. Results revealed a complex interplay of learner- and item-related variables associated with learning different aspects of form-meaning knowledge: (a) word part knowledge consistently predicted scores in the meaning recognition test at the immediate and delayed posttests; (b) vocabulary size predicted meaning recall and recognition at the immediate and delayed posttests respectively; and (c) semantic transparency significantly enhanced meaning recall performance. Based on these findings, we discuss the theoretical explanations underlying the results and pedagogical implications for how to best optimize the effects of WPT on L2 vocabulary learning.

### References

- Hamada, A., Iso, T., Kojima, M., Aizawa, K., Hoshino, Y., Sato, K., Sato, R., Chujo, J., & Yamauchi, Y. (2021). Development of a vocabulary size test for Japanese EFL learners using the new JACET list of 8,000 basic words. *JACET Journal*, 65, 23–45. [https://doi.org/10.32234/jacetjournal.65.0\\_23](https://doi.org/10.32234/jacetjournal.65.0_23)
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.
- Sasao, Y., & Webb, S. (2017). The word part levels test. *Language Teaching Research*, 21(1), 12–30. <https://doi.org/10.1177/1362168815586083>

- Suzuki, K. (2022). Effects of semantic transparency and knowledge of morphemes on the process and products of the word part technique in EFL vocabulary learning. *ARELE: Annual Review of English Language Education in Japan*, 33, 81–96. [https://doi.org/10.20581/arele.33.0\\_81](https://doi.org/10.20581/arele.33.0_81)
- Wei, Z. (2015). Does teaching mnemonics for vocabulary learning make a difference? Putting the keyword method and the word part technique to the test. *Language Teaching Research*, 19(1), 43–69. <https://doi.org/10.1177/1362168814541734>



# Exploring Factors That Influence Vocabulary Acquisition in Single-Session Flashcard Learning

**Presenter:** Thomas Stones, *Kwansei Gakuin University*

**Contributing authors:** Jon Clenton, *Hiroshima University*; Tatsuya Nakata, *Rikkyo University*; Thomas Boutorwick, *Sanyo Gakuen University*; George Higginbotham, *Eikei University of Hiroshima*

Flashcards are a widely used tool to facilitate vocabulary acquisition. Research into flashcard use has explored various factors that may influence flashcard learning, such as practice distribution (massing vs spacing), absence or presence of context, or the type of knowledge acquisition (e.g. explicit vs tacit) (Nakata & Elgort, 2021). It has also been suggested that vocabulary strategy use (Uchihara et al., 2022) and L2 proficiency level (Webb et al., 2020) are also factors that can influence vocabulary acquisition. However, as yet, no study has combined these factors into a single project. Thus, this presentation, aims to report on a study that explored the how massing and spacing influenced acquisition of explicit and tacit knowledge vocabulary in single-session flashcard learning. Additional factors such as preferred vocabulary learning strategies and language proficiency and their influence on acquisition are also analysed.

83 L1 Japanese learners of L2 English (CEFR A2 - B1) study 48 pseudoword-Japanese word pairs using flashcard software in a single session. The words are studied under two conditions, massed learning and spaced learning. Furthermore, all participants respond to a validated vocabulary learning strategy questionnaire (Gu, 2018) to ascertain the preferred vocabulary strategies of participants and the Vocabulary Size Test (Nation & Beglar, 2007) to ascertain levels of vocabulary knowledge.

Preliminary analysis of results suggests that spacing significantly outperforms massing for both tacit and explicit knowledge, but there is a limited effect for receptive vocabulary knowledge. Strategy use appeared not to impact the effectiveness of vocabulary acquisition through flashcard learning. Implications for pedagogy and avenues for future research are discussed.

## References

- Nakata, T. & Elgort, I. (2021). Effects of spacing on contextual vocabulary learning: Spacing facilitates the acquisition of explicit, but not tacit, vocabulary knowledge. *Second Language Research*, 37(2), 233–260. <https://doi.org/10.1177/0267658320927764>
- Uchihara, T., Eguchi, M., & Clenton, J. (2022). The contribution of guessing from context and dictionary use to receptive and productive vocabulary knowledge: A structural

equation modelling approach. *Language Teaching Research*, 29(6), 2583-2606. <https://doi.org/10.1177/13621688221122138>

## Degree Adverb-Adjective Collocations Across L2 Levels: A Corpus Study

**Presenter:** Alexandra Terashima, *University of Tokyo*

The use of adverbs to amplify or attenuate the intensity of an utterance is employed in both to add emphasis or nuance to the meaning of one's words. Such adverbs of degree can modify various parts of speech, but this study focused specifically on the use of degree adverbs to modify adjectives (Beltrama & Bochnak, 2015; Zhiber & Korotina, 2019). Although many studies have explored the use of intensifying and attenuating adverbs in L1 speech and writing (Lorenz, 1998; Indhiarti & Chaerunnisa, 2020), far fewer have examined how L2 users of English approach adjective modification, and fewer still have looked at degree adverbs in L2 spoken production (Hasselgård, 2022; Makhatadze, 2023). Thus, this study examined the use of degree adverb + adjective pairings across several CEFR proficiency levels in L2 spoken language by examining the Trinity Lancaster Corpus (Gablasova et al., 2019) and comparing these with the L1 data drawn from the Spoken BNC2014 corpus (Love et al., 2017; Brezina & Fox, 2021). The results suggest that while less proficient learners use fewer adverbs and adjectives than more proficient and L1 speakers, they tend to overuse the degree adverb + adjective pairing as well as specific adverbs such as *very*. Additionally, although the most frequent adjectives are similar between L1 and L2 speakers, L2 speakers draw on a narrower selection of degree adverbs to modify these adjectives and almost never use evaluative adverbs such as *horribly* or *splendidly* compared to L1 speakers. While limitations of this study should be taken into consideration, such as the differences in the nature of the spoken interactions and interpersonal relationships of the speakers, the findings of this study have pedagogical implications in the EFL context (Makhatadze, 2023) and support the claim that learners of English need more explicit instruction in the use of degree adverbs (Pérez-Paredes & Díez-Bedmar, 2019).

### References

- Beltrama, A., & Bochnak, M. R. (2015). Intensification without degrees cross-linguistically. *Natural Language & Linguistic Theory*, 33(3), 843–879. <https://doi.org/10.1007/s11049-015-9294-8>
- Brezina, V., Fox, L., & Granger, S. (2021). Adjective + noun collocations in L2 and L1 speech: Evidence from the Trinity Lancaster Corpus and the Spoken BNC2014. In S. Granger, V. Brezina, & M. Paquot (Eds.), *Perspectives on the L2 Phrasicon: The view from learner corpora* (pp. 152–177). Multilingual Matters.

- Gablasova, D., Brezina, V., & McEnery, T. (2019). The Trinity Lancaster Corpus: Development, description and application. *International Journal of Learner Corpus Research*, 5(2), 126–158. <https://doi.org/10.1075/ijlcr.19001.gab>
- Hasselgård, H. (2022). Adverb-adjective combinations in young writers' English (EL1 and EL2). *Nordic Journal of Language Teaching and Learning*, 10(2), 383–403.
- Indhiarti, T. R., & Chaerunnisa, E. R. (2020). A corpus-driven collocation analysis of degree adverbs very, really, quite, and pretty. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya*, 6(2), 226–242. <https://doi.org/10.22219/kembara.v6i2.12312>
- Love, R., Dembry, C., Hardie, A., Brezina, V., & McEnery, T. (2017). The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics*, 22(3), 319–344. <https://doi.org/10.1075/ijcl.22.3.02lov>
- Lorenz, G. (2014). Overstatement in advanced learners' writing: Stylistic aspects of adjective intensification. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Learner English on computer* (pp. 53–66). Routledge.
- Makhatadze, M. (2023). Perfectly perfect adverbs: A learner corpus study of the amplifier collocations by Georgian learners of English. *Online Journal of Humanities ETAGTSU*, (8), 1–14. <https://doi.org/10.52340/PUTK.2023.2346-8149.08>
- Pérez-Paredes, P., & Díez-Bedmar, M. B. (2019). Certainty adverbs in spoken learner language: The role of tasks and proficiency. *International Journal of Learner Corpus Research*, 5(2), 253–279. <https://doi.org/10.1075/ijlcr.18010.per>
- Zhiber, E. V., & Korotina, L. V. (2019). Intensifying adverbs in the English language. *Training, Language and Culture*, 3(3), 70–88. <https://doi.org/10.29366/2019tlc.3.3.5>

# Phrase Versus Subpart Frequency Effects in Production of Multi-Word Expressions

**Presenter:** Qi Yi, *Victoria University of Wellington*

**Contributing author:** Anna Siyanova-Chanturia, *Victoria University of Wellington*

Whether or not multi-word expressions (MWEs) are stored and retrieved holistically in the mental lexicon has been a longstanding debate in language processing. While Wray (2002) proposed holistic storage of MWEs in her seminal work, studies have shown that MWE processing is unlikely to be holistic. Arnon and Cohen Priva (2014) showed that phrase frequency effects increased, while word frequency effects decreased but did not disappear, during the production of trigrams by L1 English speakers. Studies in this area have predominantly focused on L1 speakers and/or have relied on elicited data. L2 phrase frequency effects during language production, especially subpart frequency effects, have so far received very little attention, with no study to date looking at naturalistically produced (rather than laboratory elicited) data.

The present study uses L1 and L2 spoken corpora to investigate whether and how phrase frequency and subpart frequency influence articulatory durations of verb-preposition-noun (VPN, e.g., ask for help) sequences produced by L1 and L2 English speakers in spontaneous speech. The corpora consist of naturalistically elicited telephone conversations and interviews, which were transcribed with word-level timestamps. VPN sequences and their durations were extracted from the transcribed data. Mixed-effect models were used to analyse the data, with several factors being considered during modeling, such as, L2 speaker's language proficiency, phrase frequency, unigram frequency, bigram frequency, length of phrase, etc. The present study provides novel insights into L1 and L2 speaker sensitivity to subpart versus phrase frequency during language production, and further contributes to the ongoing debate as to the nature of MWE representation and processing.

(This work is submitted as a work-in-progress study)

## References

- Arnon, I., & Cohen Priva, U. (2014). Time and again: The changing effect of word and multi-word frequency on phonetic duration for highly frequent sequences. *The Mental Lexicon*, 9(3), 377–400. <https://doi.org/10.1075/ml.9.3.01arn>
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

## Assessing the Validity of Lexical Indices Using Direct Judgements across Prepared Speech Tasks

**Presenters:** Xinyu Zhao, *Nanjing University*; Kelly Kendro, *North Arizona University*

Over the past century, lexical indices have been widely used to assess L2 learners' lexical proficiency (Jarvis, 2017), focusing on sub-constructs such as word sophistication, diversity, and density. While these indices have been applied extensively to L2 writing, few studies have examined their validity for non-spontaneous L2 speech, particularly in relation to human judgment across different tasks. This study investigates the relationship between indices of word sophistication, diversity, and density and human ratings of lexical competence in 1,152 prepared speeches by Chinese EFL learners (576 informative, 576 persuasive) from six public universities in China. Results show that lexical diversity was the strongest predictor of lexical proficiency ratings, followed by lexical density and academic word frequency. Multivariate models showed that lexical diversity, density, and sophistication collectively explained 55% of the variance in lexical proficiency ratings. Task type significantly influenced the predictive power of both diversity and density measures, suggesting that distinct tasks may require different lexical qualities to achieve higher performance levels. These results underscore the importance of task specificity in assessing lexical proficiency for oral presentations.

### References

Jarvis, S. (2017). Grounding lexical diversity in human judgments. *Language Testing*, 34(4), 537–553. <https://doi.org/10.1177/0265532217710632>