

A Critical Review of Think-Aloud Methodology in Second Language Vocabulary Acquisition and Comprehension

Alya Alghannai Abu Dirbala*

Department of English, Faculty of Arts, University of Sirte, Sirte, Libya

مراجعة نقدية لمنهجية التفكير بصوت عالٍ في اكتساب وفهم مفردات اللغة الثانية

عليا الغناي محمد ابودربالة*

قسم اللغة الانجليزية، كلية الآداب، جامعة سرت، سرت، ليبيا

*Corresponding author: alya.abudirbala@su.edu.ly

Received: April 07, 2026

Accepted: May 19, 2026

Published: June 02, 2026

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract:

This paper presents a critical review of Parry's (1997) "Vocabulary and Comprehension: Two Portraits". The review opens with the study context and findings, followed by reviewing other studies conducted in the same research area. Finally, the paper evaluates the relevance of the research method employed by the researcher. Parry's study combines a mixed research approach of both quantitative and qualitative as data-gathering procedures; however, the current paper places greater emphasis on the qualitative research design-precisely, it examines the relevance of utilizing case study in general and the think-aloud protocol in particular as a research method. Results from the current study showed that Parry's study has offered significant qualitative data into L2 vocabulary learning strategies using think-aloud protocols; however, a lack of presenting background knowledge of the study, insufficient research tools, and inconsistent coding restrict its generalizability. This review paper is expected to provide insightful implications for those who are interested in conducting similar research. In addition, presenting this kind of paper helps students develop critical thinking skills, including analysis, synthesis, and research evaluation.

Keywords: Analytical Review, Parry's Study, Second Language Vocabulary, Think-aloud Protocol.

المخلص:

تقدم هذه الورقة مراجعة نقدية لدراسة باري (1997) بعنوان "المفردات والفهم: دراسة لمتعلمين". تبدأ المراجعة بعرض سياق الدراسة ونتائجها، ثم تستعرض دراسات أخرى أجريت في نفس المجال البحثي. وأخيراً، تُقيم الورقة مدى ملاءمة المنهج البحثي الذي استخدمه الباحث. تجمع دراسة باري بين منهج بحثي مختلط يجمع بين الأساليب الكمية والنوعية في جمع البيانات؛ إلا أن هذه الورقة تُركز بشكل أكبر على تصميم البحث النوعي، وتحديدًا، تُحلل مدى ملاءمة استخدام دراسة الحالة بشكل عام، وبروتوكول التفكير بصوت عالٍ بشكل خاص، كمنهج بحثي. أظهرت نتائج هذه الدراسة أن دراسة باري قدّمت بيانات نوعية قيّمة حول استراتيجيات تعلم مفردات اللغة الثانية باستخدام بروتوكولات التفكير بصوت عالٍ؛ ومع ذلك، فإن نقص المعلومات الأساسية عن الدراسة، وعدم كفاية أدوات البحث، وعدم اتساق الترميز، تُحد من إمكانية تعميم نتائجها. من المتوقع أن تُقدم هذه الورقة معلومات مفيدة للباحثين المهتمين بإجراء بحوث مماثلة. بالإضافة إلى ذلك، فإن تقديم هذا النوع من الأوراق البحثية يساعد الطلاب على تطوير مهارات التفكير النقدي، بما في ذلك التحليل وتقييم البحوث.

الكلمات المفتاحية: مراجعة نقدية، دراسة باري، مفردات اللغة الثانية، بروتوكول التفكير بصوت عالٍ.

Introduction:

Conducting critical review of empirical studies is essential to develop students' research competence and enhance their evaluative reading skills. Aboutaha (2018) stated that the ability to analyse and criticize studies from a methodological perspective would help learners think in a logical and organised manner and equip them with applicable knowledge of research construction and validity concepts. Knowles & Gray (2011) further argued that the analytical presentation of a study's research design reveals critical methodological elements that researchers and students must consider when publishing their research or critiquing published studies. In light of this, the present critical review aims to examine a research paper conducted by Parry (1997) in the area of second language (L2) vocabulary acquisition and learning, assessing its methodological strengths and limitations.

The rationale for choosing this study is that it employs multiple methods of data collection, offering a rich case for methodological analysis. This will help other students who might be interested in this kind of research strengthen their research practices and develop their greater awareness of conducting research papers (Murad, 2014). Moreover, evaluating research papers published in the educational field could be regarded significant, as very few research studies have been developed to evaluate the effectiveness of conducting research papers from a methodological perspective.

The current paper is organized into five sections. Section one and two provide an overview of Parry's article and its key findings. The third section is devoted to discussing the literature review related to the topic of the article. Section four evaluates the qualitative research design employed by Parry. Finally, section five introduces the conclusions reached from evaluating the article.

Background and Context of Parry's Study:

The primary focus of Parry's paper is to investigate how advanced English as second language (ESL) learners acquire and understand new vocabulary that they might encounter while learning another language. She addresses the dilemma confronted by most L2 learners who intend to complete their degrees in foreign universities. Parry argues that most L2 learners face a problem when they learn words which are linked to their study major. As the researcher notes, L2 learners are expected to acquire these words (either common or rare) independently, since they are not explicitly taught in content courses or in any language course the learners might attend. Another focus of Parry's research is to shed light on the strategies learners might employ to learn these words. The study also provides a pedagogical reference for language teachers of the essential teaching methods that would assist students develop their learning strategies while acquiring new words.

Parry used a case study methodology to collect data from multiple sources. Two students- one Greek and one Korean, both proficient English speakers- first completed a standardized vocabulary test from the Michigan Test of English Language Proficiency (MTELP). They were then asked to identify difficult vocabulary from an anthropology textbook and maintain a word list. For each item, the students registered details like the page number, their own guesses at the meaning, and the dictionary definition for the lexical item. After six weeks of compiling these lists, both students took part in a think-aloud protocol in which they also had to keep word lists from a second anthropology text. During each session, students verbalized their own inference of unknown words or consulted a dictionary. At the end of the study period, each student was tested on their personal word list. The target words were presented both in isolation and embedded in sentences. Students were asked to write the meaning of each word in its given context and were allowed to use their first language (L1) in their interpretations throughout the study. Glosses and translations were then put back into English by two Greek and Korean proficient speakers of English.

Findings of Parry's Study:

Based on the vocabulary lists each student created, Dimitri (the Greek student) appears to be more successful; he read more pages and made mostly accurate guesses. However, the post-test scores revealed reversed results. Ae Young (the Korean student) demonstrated a recognizable ability in both identifying and recalling the meanings of the words she had listed than Dimitri; whether the words were presented in isolation or in sentence contexts.

The think-aloud protocols and translation data help explain these contrasted outcomes. In terms of the number of words listed, Ae Young recorded more words than Dimitri during think-aloud. Dimitri's proportion was slightly greater compared to his initial list. Regarding word accuracy, both participants produced more 'partly correct' glosses than they did on their own lists. Notably, Ae Young's guesses were generally less accurate than Dimitri's. A remarkable finding emerged in processing strategies. In other words, Dimitri appears to deal with the target words as part of the sentence in which they exist, often paraphrasing the whole sentence to understand them. The researcher describes this as a 'holistic' or 'top down' approach. Ae Young, by contrast, prefers to isolate the words and provide a general meaning or a substitute word/phrase, referring to Ae Young's word processing approach as being 'analytic' or 'bottom-up'. Parry concludes that Ae Young's analytic approach negatively affected the

speed of her reading, but it enables long-term retention and more accurate word comprehension. Dimitri's holistic approach, conversely, allows for faster reading but with lower retention and less precise comprehension.

Differences in word processing approaches are well-documented in the field of L2 learning and acquisition. For example, the concepts of holistic and analytic are related to the term 'cognitive style', which is specifically linked to the studies of individual differences in L2 learning.

Dörnyzei (2005) defines cognitive styles as a learner's preferred and habitual way of perceiving, processing, and organising information. Cognitive styles include two key dimensions: holistic and analytic. Holistic learners process information as an integrated whole and use unstructured and personalised approaches. Analytic learners break information into discrete parts and follow structured, organised paths. In the domain of L2 vocabulary acquisition, cognitive styles appear to reflect learners' approaches to unfamiliar lexis. Studies showed that analytic learners exhibit a preference for information presented in a coherent and organised format. In contrast, holistic learners rely on broader and less structured strategies (Sadler-Smith & Riding, 1999).

The distinction between Dimitri as a holistic learner and Ae Young as an analytic learner appears to stem from differences in their strategy use rather than from the dimensions themselves. For instance, Nezhad and Shokrpour (2012) investigated the effect of cognitive styles on vocabulary learning performance. They argue that the main reason for the dissimilarity between field dependent and field independent- earlier terms for holistic and analytic styles, respectively- may be attributed to the variation "in the kind of reasoning, the restructuring ability and the use of strategies in processing information" (2012, p.21) employed within each dimension.

A Review of Previous Studies in the Same Research Area:

Chang et al. (2013) carried out a study which aimed to examine the type of strategies each dimension might use to learn new vocabulary. The strategies they included in their study were the ones discussed in Schmitt's taxonomy (1997), which are related to social strategies (SOCs), determination strategies (DETs), strategies (COGs), memory strategies (MEMs), cognitive, and metacognitive strategies (METs). The results showed that field-dependents (holistic) were more likely to adopt SOCs, whereas field-independent (analytic) learners tend to use more METs and MEMs in their approach to learn L2 words.

Another study was carried out by Liu and Leed (1994) (as cited in Chang et al, 2013) in which their results, to some extent, resembled Parry's findings. In their study, examining detailed information of the target words was found to be the preferred learning approach of field independent learners. These details are concerned with knowledge of vocabulary relationships and parts of speech; this appears to correspond with Ae Young's technique of word analysis. In contrast, field-dependents tend to focus on broad information of words from a general perspective, resembling Dimitri's way of approaching and learning the unknown words.

In her study, Littlemore's (2001) asked some learners, who were categorised as holistic and analytics, to describe a picture for which a vocabulary was not given. Her study mainly aimed to investigate the relationship between learners' cognitive styles and their use of communication strategies. The participants employ the communicative strategies that significantly replicate their cognitive styles. For example, learners who have been categorised as analytics tend to use analytic strategies, which include description-based strategies, whereas holistics employ more holistic strategies, such as comparison-based strategies.

Relatively similar research to Parry's was conducted by Lawson and Hogben (1996). They used a think-aloud protocol to examine the strategies students use to learn unfamiliar Italian words. The target words were presented to the participants on flashcards with their translations and semantically related words on the reverse side. Learners were then tested on these words. The results indicated a positive correlation between accurate recall of target word and students' use of specific strategies.

Another similar study was conducted by Barcroft (2009) who used a retrospective think-aloud protocol to examine strategies L2 students use to acquire Spanish words. After the learning stage, the researcher administered a recall test. The results showed that some strategies were significantly more effective than others.

Analysis of the Qualitative Research Design of Parry's Study:

Throughout her study, Parry provided a convincing justification for administering case study as a methodology for data collection. She explained that this procedure allowed her to provide sufficient details of the strategies employed by each student when encountering unfamiliar words and of the accuracy of their textual interpretations. Moreover, it yielded a clear profile of students' lexical store and documented the stages of vocabulary development over the study period. Azhar & Kernagaran (2025) pointed out that the case study approach enables comprehensive analyses of complex problems in their actual environments, as it explores a current phenomenon in detail and within its real-life

environment, particularly when the boundaries between the phenomenon and setting are not immediately apparent. By accumulating data from various facets during the case study (i.e., pre-test/post-test, think-aloud protocol, and students' translations), it can be argued that it assists the researcher in forming a detailed representation, which works together to confirm and support the results.

Parry's use of the think-aloud protocol further strengthened her study in achieving the study aims she identifies. Specifically, the protocol functions as an effective instrument to reveal an immediate and direct representation of learners' cognitive processes and of the strategies they prefer to acquire new vocabulary. For instance, think-aloud reports allowed Parry to trace the developmental stages of vocabulary acquisition of both learners without disrupting the natural sequence of thought narration. This linear improvement would not be effectively reflected if alternative research tools were used. This observation is also addressed by Lawson and Hogben, who noted that the using think-aloud protocol enabled them "to undertake detailed observations of the strategies students spontaneously and gradually employed as they engaged in the vocabulary acquisition task we set for them (i.e., students)" (1996, p. 108).

Participants were assessed in an environment that reflects their real learning processes, which constitutes a further methodical strength of Parry's research. The test materials were part of the students' learning course, and they were permitted to consult dictionaries and use their L1 as assistive tools during the think-aloud task. According to Polio (2003), using think-aloud protocols can pose some difficulties for L2 learners when required to verbalize their thoughts. Additionally, the use of the L2 can be challenging and might not guarantee a spontaneous flow of students' ideas, as they are not used to such metacognitive behaviours in their L1. Arguably, allowing participants to draw on both their L1 and L2 as required has mitigated this limitation and enabled them to describe and express their thoughts more freely.

In addition to the think-aloud protocol, questionnaires and retrospective think-aloud are other two research methods administered in several studies to ask students about their use of strategies to learn unknown words (e.g., Chang et al., 2013; Barcroft, 2009; Hogben & Lawson, 1993). Nonetheless, these research instruments have been criticised for being less effective compared to think-aloud procedures. For instance, Lawson and Hogben (1993) found that retrospective think-aloud failed to represent students' spontaneous use of learning strategies. Moreover, within retrospective think-aloud, learners are found to include additional analysis and interpretation in their responses (Barcroft, 2009). Regarding the use of interviews, they have been criticised for including some degree of retrieval from long-term memory, which can affect the original cognitive processes.

The use of questionnaires to ask participants about their strategies while learning new words is also found to have certain limitations. For example, Chang et al. (2013) argued that questionnaires may not accurately reflect participants' actual learning behaviours, since they rely on self-reported data. In addition, the use of closed-ended items in questionnaires often restricts the amount of responses learners can express. Tseng, Dörnyzei, and Schmitt (2006) further argue that questionnaire statements usually require participants to verbalize their actions across general learning situations, rather than referring to specific learning events.

Some researchers have identified several weaknesses of administering think-aloud protocols. They argue that think-aloud protocols may not be a suitable research instrument for collecting data on learning strategies, as learners are often unable to report all the strategies they actually employ in vocabulary learning tasks. For example, Ericsson and Simon (1980, cited in Charters, 2003) stated that data gathered from think-aloud procedures are usually incomplete and may exclude some cognitive processes that are not present in working memory and therefore are not verbalised. This limitation might result from learners' unfamiliarity with such behaviours in their learning process. Charters (2003) further argues that thinking-aloud cannot completely reveal the complexity of the underlying cognitive processes.

In addition to issues of incomplete data, risks of reactivity during think-aloud tasks raise further concerns about the validity of this method. Learners may become self-conscious during this process, which can lead to either over-use or under-use of learning strategies. Parry herself observed this effect when examining Dimitri during think-aloud; she noted that the word lists he created while verbalising thoughts significantly outnumbered the ones he compiled independently.

Another area of weakness concerns the use of post-tests. Hulstijn (2001) argued that administering tests in association with the think-aloud protocols can negatively affect the collection of spontaneous and valid responses from the participants. This explains Ae Young's remarkable ability to understand and recall target words, which could have resulted from her use of additional vocabulary learning strategies such as memorisation, rather than from the strategies she verbalised during think-aloud. Reliance on such strategies allowed her to have better performance on subsequent memory tests.

Some scholars have proposed remedies to address weaknesses that may arise from using think-aloud protocols in this type of research. Charters (2003), for example, recommends using interviews or retrospective questions as follow-up procedures to collect detailed data on participants' learning strategies. Rankin (1988) similarly argues that retrospective questioning serves as a useful supplementary instrument, particularly for learners who struggle with concurrent think-aloud protocols. Interviews and retrospective questions can be effective tools for eliciting more detailed information on the procedures Ae Young and Dimitri followed to translate the reading passages, or to ask about the challenges they encountered when completing the vocabulary task. However, as stated earlier, retrospective methods may involve retrieval from long-term memory and introduce biased responses which, in turn, conflict with the core objectives and nature of think-aloud procedures (Charters, 2003). To eliminate this risk, Nunan (1992) recommends that combining data from both retrospective and concurrent thinking-aloud reports could provide a more complete picture of learners' cognitive processes.

Concerning Parry research design, it appears that several fundamental elements were missing from the study. For example, the study provided little background on how L2 learners acquire and comprehend new lexical items, nor did it review prior literature regarding the vocabulary learning strategies L2 learners use. Instead, Parry briefly stated the problem, outlined her methodology and research questions, and moved directly to present her results. This is an unfamiliar research design. While this untraditional structure may serve as an elicitation strategy to encourage readers' critical and creative reasoning, as unexpected findings are continually unfolded throughout the study, it can be problematic for readers who do not have prior knowledge of certain terms used in the study, or of the strategies associated with each dimension. For instance, readers may struggle to distinguish between 'holistic' and 'analytic', and the introduction of alternative notions of 'bottom up' and 'top down' does not help in understanding these concepts.

To address this issue, some researchers recommend including coding as a key element of data analysis. Gu (2014) offers a useful discussion of coding when using think-aloud protocols. Based on his experience, he argues that if the primary aim of think-aloud is to uncover the learners' strategic behaviours, coding is unnecessary and would not be as effective as using narratives or discourse analysis. At first glance, this seems applicable to Parry's study, which appears descriptive and narrative in nature and only aims to identify the strategies learners use to acquire L2 vocabulary. Parry clearly stated this aim in her first research question. However, when presenting her results, Parry made a direct comparison between Dimitri and Ae Young in terms of their approach and strategy use for learning new words. In this case, coding becomes essential. Gu (2014) states that coding is required when the research design aims to identify learning patterns or make systematic comparisons and contrasts.

Conclusion:

Parry's research aimed to examine the learning strategies used by two proficient L2 learners to acquire and comprehend new vocabulary. Her choice of case study as a research method clearly enabled her to achieve her research objectives, which included tracing the development of the students' strategy use and assessing their accuracy level in dealing with unfamiliar words. The think-aloud protocol, in particular, allowed Parry to reveal variation in the learning approaches adopted by most L2 learners when encountering new vocabulary. However, the study could be further improved by providing more background information on the research area. Moreover, the results could have been presented in greater detail if additional research instruments were included. For instance, supplementing the study with qualitative methods such as interviews could significantly enhance its validity and provide deeper insights into the challenges L2 learners face in acquiring and comprehending new words.

Limitation of the Present Review:

It should be acknowledged that this critical review is limited by its reliance on Parry's published report, without access to raw think-aloud transcripts or participant data. This may restrict the depth of the methodological critique analysis. Additionally, this review focuses primarily on the strengths and limitations of Parry's qualitative design, excluding other factors such as the statistical reliability of her comprehension tests. Finally, because the conclusions drawn from a single study, their generalizability to broader trends in think-aloud research is limited. Future research could address these limitations by conducting comparative critiques of multiple think-aloud studies or replicating Parry's research design to validate its findings.

References:

1. Coady, J. & Huckin, T. (1997). *Second Language Vocabulary Acquisition*. Cambridge: Cambridge University Press. 55-68.
2. Aboutaha, A. S., (2018). The Critical Importance of Learning to Critique Research Articles. LinkedIn. Retrieved from: <https://www.linkedin.com/pulse/critical-importance-learning-critique-research-abir-s-aboutaha-ph-d>
3. Azhar, A., & Kernagaran, K. K. (2025). A critical review of journal articles in educational research with case study research design.
4. Barcroft, J. (2009). Strategies and performance in intentional L2 vocabulary. *Language Awareness*. 18(1), 74-89.
5. Charters, E. (2003). The use of think-aloud methods in qualitative research: an introduction to think-aloud methods. *Block Education*. 12(2), 68-82.
6. Chang, M., Weng, C., & Zakhavora, A., (2013). Vocabulary learning strategies and cognitive styles among junior high school students in Taiwan. *The Asian Conference on Language Learning*, 443-455.
7. Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Erlbaum.
8. Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*. 87(3), 215-251.
9. Gu, Y. (2014). To code or not to code: Dilemmas in analyzing think-aloud protocols in learning strategies research. *System*, 1-8.
10. Hogben, D., & Lawson, M. J. (1993). Elaborated keyword strategies for foreign language acquisition. Paper presented at the annual conference of the Australian Association for Research in Education, Fremantle, Western Australia.
11. Hulstijn, J.H. 2001. Intentional and incidental second language vocabulary learning: a reappraisal of elaboration, rehearsal and automaticity. In Robinson, P. (ed.) *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, 258-286.
12. Knowles, J. M., & Gray, M. A., (2011). The experience of criticizing published research: learning from student and researcher perspective. *Nurse education in Practice*. 11(6), 390-394
13. Lawson, M. J. & Hogben, D. (1996). The vocabulary learning strategies of foreign language students. *Language Learning*. 46(1), 101-135.
14. Liu, M., & Reed, W. M. (1994). The relationship between the learning strategies and learning styles in a Hypermedia environment. *Computer in Human Behavior*, 10(4), 293-312.
15. Littlemore, J. (2001). An empirical study of the relationship between cognitive style and the use of communication strategy. *Applied Linguistics* 22(2), 241-265.
16. Murad, I., H. (2014). A Critical Review Study Conducted on Two Academic Articles Published in the Educational Field: From a Research Perspective. *Education*. 4(6), 148-155
17. Nunan, D. (1992). *Research Methods in Language Learning*. Cambridge University Press.
18. Polio, C. (2003). Research on second language writing: An overview of what we investigate and how. In B. Kroll (ed.), *Exploring the Dynamics of Second Language Writing*. Cambridge: Cambridge University Press, 35-65.
19. Rankin, J. M. (1988). Designing thinking-aloud strategies in ESL reading. *Reading in a Foreign Language*. 4(2), 119-132.
20. Sadler-Smith, E., Riding, R.J., (1999). Cognitive style and instructional preferences. *Instructional Science* 27, 355-371.
21. Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied linguistics* 27(1), 78-102.