

Game-based TPR Vocabulary Programme Without a Shared Language in Indonesia

Sathyavathi Yarlagadda

Azim Premji University, Bengaluru, India

sathyavathi.yarlagadda23ug@apu.edu.in

Abstract: Most game-based language teaching research assumes that the teacher and learners share at least one common language. This assumption shapes how interventions are designed and leaves a gap when it does not hold. This practitioner reflection examines that assumption through a one-month English vocabulary programme at the Speak and Speak English Training Centre in Temanggung, Central Java, Indonesia, with 25 students aged 3.5 to 13. Students came from a trilingual context where Javanese was the home language, Bahasa Indonesia the language of schooling, and English an additional language with minimal classroom exposure. The students and teacher shared no common language at all. The programme drew on Total Physical Response, embodied cognition, and game-based learning, and used no reading, writing, or homework. This paper distinguishes between two categories of activity: games, defined as rule-governed activities with a competitive goal and immediate visible feedback tied to vocabulary performance, and playful instructional routines, which are structured physical activities without competitive orientation. Building on this distinction, vocabulary was developed through games including dog and bone, musical chairs, hopscotch, find the noun, and colour identification, alongside playful routines such as call-and-response and embodied action pairing. All meaning was conveyed through gesture, demonstration, and physical modelling, with no spoken instruction in a shared language at any point. Outcomes were assessed using flashcard-based tasks. Students viewed an image and produced the English word orally to test receptive knowledge, or viewed a written word and performed the corresponding physical action to test productive knowledge. From a baseline at which most students could not produce common greetings without prompting, students demonstrated consistent use of approximately 120 words by the end of the programme. Notably, only five of the twenty-two activities in the framework qualify as games by the criteria defined above; the remaining activities are playful instructional routines that provided the input and modelling on which the games depended. The findings suggest that shared language is not a precondition for effective game-based vocabulary instruction. The paper offers a low-resource, replicable activity framework for practitioners working in multilingual, no-shared-language contexts.

Keywords: Total physical response, Game-based learning, Foundational vocabulary, Multilingual contexts, Embodied cognition

1. Introduction

Children across the Global South often grow up between languages. Many speak a home language, learn through a national language at school, and encounter English as an additional instructional language. Children in Temanggung speak Javanese at home, use Bahasa Indonesia at school, and encounter English with almost no exposure outside the classroom. English remains peripheral to their daily life, even as its importance for education and employment increases. This creates a gap between limited exposure and rising expectations. How can English vocabulary be taught when there is little exposure and no shared language between the teacher and learners?

Stanovich (1986) describes how early vocabulary differences widen over time through the Matthew Effect. Learners who begin with stronger vocabularies encounter more language and continue to advance, while others fall further behind. Vocabulary, in this sense, is not just one skill among many. It is the gateway through which every other school subject is accessed. Closing foundational vocabulary gaps early is, therefore, central to educational equity.

This paper presents a one-month game-based English vocabulary programme implemented in Temanggung, Indonesia. The programme used TPR, embodied cognition, and game-based learning, and relied entirely on movement, gesture, and play. There was no shared language between the teacher and the students. The paper argues that shared language is not a precondition for effective game-based vocabulary instruction. The activity framework documented here shows that meaning, motivation, and lexical practice can be sustained through action, physical engagement, and play.

2. Literature Review

Three bodies of research informed the programme. Total Physical Response (TPR), developed by Asher (1969), links language input to physical movement and argues that learners retain language better when they act on it. Yang (2024) reports consistent gains in vocabulary retention and motivation, while Trung Cang, Diem and Thien

(2021) show that students taught through TPR retained vocabulary more reliably than those taught through conventional methods.

Embodied cognition offers a theoretical account of why physical engagement supports vocabulary retention. Macrine and Fugate (2021) argue that sensorimotor experience shapes how concepts are encoded and retrieved. Macedonia and Klimesch (2014) show that pairing vocabulary with gesture improves recall, including over time.

Game-based learning adds a third layer. Hafiza and Pratolo (2024) report gains in engagement and vocabulary learning, as games create repeated, low-stakes practice. What counts as a game in that literature, however, requires clarification before the activities in this programme can be situated within it. The term "game-based learning" carries different meanings across research communities, and situating this paper within that landscape requires some precision. Prensky (2001), who is widely credited with popularising the concept, defined games through twelve structural elements including rules, goals, feedback, competition, and interaction, and applied this framework primarily to digital environments. His work is concerned with how video and computer games can be harnessed for learning. Gee (2003) extended this into a theory of learning, identifying thirty-six principles embedded in good video games and arguing that these principles represent what effective learning environments should look like more broadly. Plass, Homer and Kinzer (2015) built on this to propose cognitive, motivational, affective, and sociocultural foundations for game-based learning, again with digital games as the primary referent. Taken together, these accounts have shaped a body of GBL research that rests on infrastructural assumptions about device access, reliable connectivity, and software availability that do not hold across much of the Global South, where mobile penetration is uneven, devices are shared or unavailable, and internet access is expensive or unreliable (Trucano, 2010). In these settings, digital GBL is not a realistic default. The present programme did not use digital tools of any kind, and this was not a limitation of the design. It was a design choice appropriate to the context.

A different and older body of work treats game-based learning as a broader category that encompasses adapted physical activities with rules, goals, and feedback mechanisms in language teaching contexts. Wright, Betteridge and Buckby (2006), whose work on games for language learning has been in continuous use since 1979 and spans hundreds of classroom-tested physical activities, treat the game as a pedagogical structure rather than a technological one. What makes something a game in this account is not its medium but its architecture: a task with a rule, a goal, and a consequence for performance. By this definition, five activities in this programme qualify as educational games: find the noun, colour identification, dog and bone, musical chairs, and hopscotch. Each has explicit rules demonstrable through action alone, a competitive goal, and immediate visible feedback tied directly to vocabulary performance. Critically, each was adapted so that a correct vocabulary response was the condition for performing the game mechanic at all. The game and the vocabulary assessment were not separate activities. The game was the assessment.

Not all activities in this programme meet this definition equally, and the distinction matters. Call-and-response routines, embodied gesture pairing, and marching sequences are better described as playful instructional routines: structured, purposeful, and physically engaging, but without the competitive goal orientation that defines a game proper. Both types served complementary functions within the programme. Playful routines provided initial input and repeated modelling, building the sensorimotor traces that embodied cognition research predicts will support later recall. Rule-based games then activated that vocabulary under time pressure, competitive motivation, and social stakes. Neither type alone would have been sufficient. Future work should examine how these two categories function differently, and whether their interplay is itself a meaningful design variable in low-resource, no-shared-language contexts. This is a question the current GBL literature, focused as it is on digital environments, has not addressed.

Across all three bodies of research, a shared linguistic resource between teacher and students is assumed. Translation, explanation, and translanguageing all depend on this. Pardede (2018) documents how teachers and students in Indonesian EFL classrooms routinely draw on shared languages to scaffold instruction. While methods such as TPR reduce reliance on translation, the literature does not directly examine teaching in contexts where no common language exists between teacher and learners. Yet this is not a marginal scenario. Volunteer teachers, NGO educators, travelling instructors, and short-term programmes across the Global South routinely operate in this condition. The programme addresses this gap.

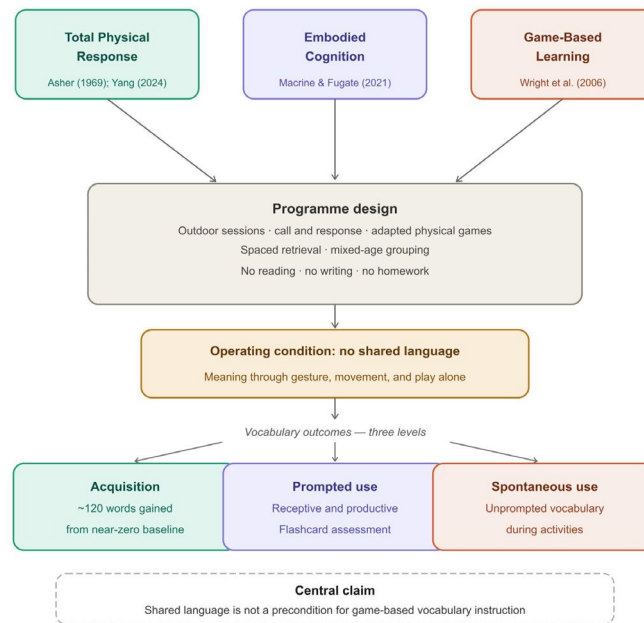


Figure 1: Conceptual framework showing the three theoretical pillars informing programme design, the no-shared-language operating condition, and the three levels of vocabulary outcome assessed

3. Programme Design, Implementation, and Assessment

The programme was implemented in December 2024 at the Speak and Speak (SAS) English Training Centre in Temanggung, Central Java, Indonesia. Twenty-five students participated, aged 3.5 to 13, in mixed-age groups. Students came from Javanese-speaking homes, used Bahasa Indonesia in formal schooling, and had minimal prior English exposure. At baseline, most students could not produce common greetings such as "good morning" or "hello" without prompting.

This paper reports on practitioner inquiry, a form of systematic reflection in which the teacher documents, analyses, and theorises their own practice (Cochran-Smith and Lytle, 1999). The data presented here are practitioner observations collected during and immediately after the programme rather than through standardised instruments. This is consistent with the work-in-progress framing of the paper: the programme was designed as teaching, not as a formal experiment, and the findings reported here are the teacher's structured reflection on what that teaching produced.

The programme ran as daily one-hour outdoor sessions over four weeks, covering approximately 120 target words across seven lexical domains: action verbs, spatial terms, body parts, numbers, colours, descriptive terms, and social expressions. Learning took place through TPR, where students used their whole body to understand and respond to language. Reading, writing, and homework were deliberately excluded. Research consistently shows that performance anxiety inhibits language acquisition, particularly in young learners and beginners (Krashen, 1982), and that literacy tasks carry cognitive demands that can overwhelm students whose attention should be on the target language itself (Swain, 1985). The programme was designed to feel categorically different from formal schooling, lighter, physical, and playful, so that students associated English with movement and enjoyment rather than with the pressure and cognitive load that reading and writing tasks tend to introduce at the earliest stages of exposure.

Before any new word was introduced, every session opened with a physical warm-up recalling vocabulary from previous sessions. Students moved through words they already knew, jumping while calling them out, running through spatial sequences, responding to body-part calls, sometimes led by the teacher, sometimes by a student. Hesitation on a word during the warm-up flagged it for re-teaching: that word returned later in the session through a different activity. This sequencing reflects spaced retrieval, the reactivation of prior knowledge at intervals to strengthen long-term retention (Roediger and Butler, 2011). New vocabulary was introduced through call and response: the teacher initiated a word or action, the whole group repeated it in unison, and meaning was built through the body before it was built through the mind. Session planning also drew on Lemov's (2021) high-engagement principles. Cold calling ensured younger and quieter students were assessed rather

than overlooked. Activities shifted every eight to ten minutes because when attention drops in an outdoor, no-shared-language setting, there is no way to verbally redirect the group.

New words were introduced through action rather than explanation. When a verb such as "jump" or "stop" was introduced, students acted immediately. For spatial terms, a chair was used as a reference point, with instructions such as "stand in front of the chair" or "go behind the chair." For colours, students located matching objects in the environment. For descriptive words such as fast and slow, students adjusted their movement accordingly. Meaning developed through action rather than through translation or verbal explanation.

Recall was practised through activities that required quick, physical responses. In body-part tasks, students used flour and placed it on the body part called out, making their responses immediate and visible. In games like dog and bone, each student listened for their word and ran to retrieve an object, so recall happened under time pressure. Movement-based sequences, such as marching or hopscotch, combined instructions like "three steps forward" or "turn left," requiring students to process numbers, direction, and verbs together. Musical chairs focused on language use: students followed changing commands like jump, skip, hop, go back, or turn while moving, and still had to respond when the music stopped. Across these activities, recall was practised through action in different situations rather than through repetition of the same form.

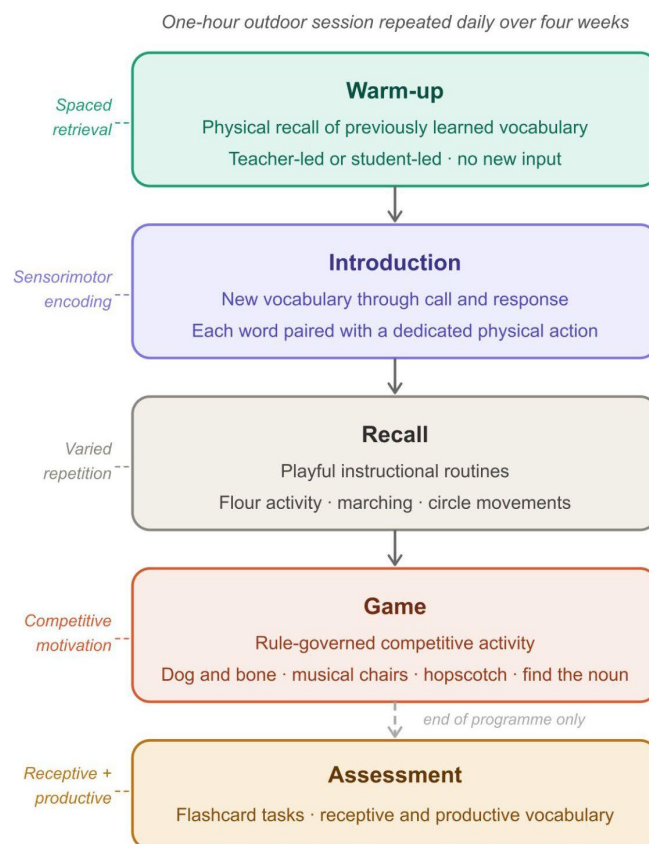


Figure 2: Session structure showing the five phases of each one-hour outdoor session, repeated daily over four weeks

Vocabulary development was tracked through ongoing observation, focusing on participation, response accuracy, and spontaneous use of target words during activities. End-of-programme assessment used flashcard-based tasks testing both receptive and productive vocabulary knowledge. Students saw an image and produced the word, or they saw a word and performed the action. Some cards were paired as opposites (fast/slow, up/down) to check relational understanding. Flashcards were also used for simple sentence work: students responded to prompts such as "My name is..." and, in reverse tasks, they were given a response and had to produce the appropriate question ("What is your name?"). Older students who could read English text read the words aloud for younger students, who then performed the corresponding action. This cross-age mediation was

encouraged, as it reflected authentic mixed-age classroom dynamics and allowed assessment of comprehension even among pre-literate learners.

Many activities were adapted based on space, time, group size, and student level. Games worked only if rules could be shown through simple actions and minimal language, so instructions stayed short and clear. Over the hour, activities kept changing, using different approaches to work with the same vocabulary. This variation kept students engaged while strengthening recall through repeated use in different forms.

Table 1 presents the activity framework, grouping activities by phase and type. The five activities listed as games meet the full criteria of rule-governed structure, competitive goal, and immediate visible feedback established in Section 2.

Table 1: Session activity framework showing phase/type, representative activities, pedagogical function, and classification rationale

Phase	Activity	Pedagogical function	Type
Warm-up	Musical warm-up	Physical and cognitive transition into the session; signals the start of English without verbal instruction	Warm-up
	Physical recall sequences	Spaced retrieval of prior vocabulary; activates stored knowledge before new input is introduced	Warm-up
	Student-led recall	Peer-initiated retrieval; reveals retention gaps without teacher-directed testing	Warm-up
Introduction	Call and response	First exposure to target vocabulary; whole-group sensorimotor encoding through simultaneous sound and movement	Playful routine
	Embodied learning	Durable sensorimotor memory trace for each new lexical item; tactile engagement deepens encoding	Playful routine
	Group call and response	Introduces question-and-answer structures for functional vocabulary; builds syntactic as well as lexical knowledge	Playful routine
	Act the adjective	Embodied introduction of emotional and descriptive vocabulary; meaning built through physical enactment rather than definition	Playful routine
	Chair prepositions	Embodied introduction of prepositional and spatial vocabulary through physical positioning relative to a fixed reference point	Playful routine
Recall	Body-part flour activity	Rapid receptive recall with tactile reinforcement; whole-group response visible simultaneously, allowing continuous informal assessment	Playful routine
	Circle movements	Simultaneous recall of number, spatial, and action vocabulary in coordinated group movement sequences	Playful routine
	Verb-and-action	High-frequency productive recall of action vocabulary; builds automaticity through progressively increasing pace	Playful routine

Phase	Activity	Pedagogical function	Type
Game	Find the noun	Competitive receptive recall of concrete nouns under time pressure; students race to touch the named object (see Section 2 for game criteria)	Game
	Colour identification	Competitive receptive recall of colour vocabulary through environmental scanning; first to touch the correct colour wins (see Section 2)	Game
	Dog and bone	Competitive receptive recall under time pressure; lexical discrimination required to retrieve correctly before the opposing student (see Section 2)	Game
	Musical chairs (adapted)	Productive recall integrated into the game mechanic as condition for dignified elimination; motion descriptors introduced through play (see Section 2)	Game
	Hopscotch	Sequential receptive recall of numerical and target vocabulary embedded in physical movement; rule-governed structure with competitive goal (see Section 2)	Game
Assessment	Flashcard image-to-word	Receptive vocabulary knowledge across all seven target domains; student views image and produces English word orally	Assessment task
	Flashcard word-to-action	Productive vocabulary knowledge; student views written word and performs the corresponding physical action, separating comprehension from literacy	Assessment task
	Semantic opposite pairs	Relational vocabulary understanding; conceptual grasp tested beyond rote word-action pairing through response to opposite pairs	Assessment task
	Sentence frames	Functional and syntactic vocabulary use beyond single-word production; students complete prompted frames and produce matching questions	Assessment task

4. Findings

From minimal prior English at the start, students demonstrated consistent use of approximately 120 words by the end of the programme. Over time, participation increased, responses became more accurate, and students began using vocabulary spontaneously during activities. Younger learners (under 6) required more frequent modelling and physical cueing before responding independently, while older learners showed faster independent recall, though all age groups demonstrated observable gains. Students were also able to respond to opposites and simple sentence prompts during flashcard tasks, suggesting the development of relational and functional use of vocabulary. The absence of a shared instructional language did not appear to hinder learning. Reliance on gesture and action appeared to support vocabulary recall, consistent with embodied cognition research (Macrine and Fugate, 2021).

The programme is not defined by the specific games used but by how those games are structured. Activities only worked when three conditions were met: the word was tied to a visible action, students responded physically rather than verbally, and understanding could be seen through what they did. Because these conditions remained constant, different activities could be adapted or replaced without changing how vocabulary was learned.

5. Discussion

The programme supports the central claim of this paper: shared language is not a precondition for game-based vocabulary instruction. Twenty-five children with minimal prior English and no common language with their

teacher acquired around 120 functional words in one month through outdoor games, gestures, and physical interaction. The figure of approximately 120 words is based on systematic practitioner observation of production and comprehension during flashcard tasks and classroom activities, rather than on standardised testing. While this limits claims of precise measurement, it is consistent with the practitioner-inquiry design of the study and reflects the functional vocabulary students were observed to use with consistency by the end of the programme. This challenges the assumption that translation or shared linguistic resources are necessary for early vocabulary learning.

Three implications follow. First, the activity framework offers a low-resource model that teachers can adapt across contexts, using simple materials such as chalk, a stick, chairs, and flashcards, and working across vocabulary domains and age groups. Second, the findings challenge a key assumption in game-based learning research, which often relies on shared language and overlooks common contexts such as volunteer programmes, NGO classrooms, and short-term teaching settings across the Global South. Third, removing translation shifts instruction toward gesture, demonstration, and movement to carry meaning, aligning with how children acquire their first language: through exposure, physical context, and repetition rather than through translation.

The mixed-age grouping, outdoor setting, and deliberate exclusion of reading and writing may also have contributed to a low affective filter (Krashen, 1982). By removing literacy demands and formal classroom pressure, the design allowed students to engage with English primarily through movement and play, which appeared to reduce anxiety and support willingness to participate from the earliest sessions.

Notably, only five of the twenty-two activities in the framework qualify as games by the criteria established in Section 2. The remaining activities are playful instructional routines that provided the input and modelling on which the games depended. This suggests that the ratio of games to routines, and the sequencing between them, may be as important a design variable as the games themselves.

Game-based vocabulary learning works without a shared language. It should not be treated as a default condition in programme design or evaluation.

Because the programme was conducted as practitioner inquiry rather than as a controlled experimental study, the findings should be interpreted as evidence from structured classroom observation rather than as statistically measured learning gains. The absence of baseline and post-programme quantitative scores prevents claims about the magnitude of vocabulary improvement.

6. Limitations and Future Work

The programme has several limitations. Assessment was informal and practitioner-administered, with no independent verification or standardised instruments. Long-term retention was not tracked, leaving it unclear whether vocabulary gains persisted over time. The wide age range (3.5 to 13) meant that age-disaggregated analysis was not conducted, despite likely differences in how learners responded to specific activities. The programme was implemented by a single teacher at a single site, which limits its generalisability.

Future research should extend the implementation period, incorporate retention tracking, and include age-disaggregated analyses and independent assessments across multiple sites. Comparative studies examining no-shared-language instruction alongside shared-language and translanguaging approaches would allow for a more rigorous test of the central claim.

Acknowledgements

The author would like to thank Indrotomo and the staff of the Speak and Speak English Training Centre, Temanggung, Central Java, for their institutional support and their willingness to allow this work to be documented and shared. Thanks are also due to the students and their families for their participation.

Ethics Declaration: This work was conducted as an independent practitioner-led teaching initiative at the Speak and Speak (SAS) English Training Centre, Temanggung, Indonesia, in December 2024. The programme was implemented with the knowledge and written approval of the centre's institutional head. Parents and guardians enrolled their children voluntarily in the English programme. No identifying information about individual students is reported in this paper. Written institutional consent for the use of teaching observations and assessment outcomes for academic publication has been obtained from the centre's head.

AI Declaration: AI tools were used for editing this paper. The design, implementation, data collection, intellectual content, and central argument are the author's own.

References

- Asher, J.J. (1969) "The total physical response approach to second language learning", **The Modern Language Journal**, Vol. 53, No. 1, pp. 3–17.
- Cochran-Smith, M. and Lytle, S.L. (1999) "The teacher research movement: A decade later", **Educational Researcher**, Vol. 28, No. 7, pp. 15–25.
- Gee, J.P. (2003) **What Video Games Have to Teach Us About Learning and Literacy**, Palgrave Macmillan, New York.
- Hafiza, A.F. and Pratolo, B.W. (2024) "A systematic review of the effectiveness of game-based learning in English language teaching", **International Journal of Education and Learning**, Vol. 6, No. 2, pp. 56–64. doi:10.31763/ijele.v6i2.1312
- Krashen, S.D. (1982) **Principles and Practice in Second Language Acquisition**, Pergamon, Oxford.
- Lemov, D. (2021) **Teach Like a Champion 3.0**, Jossey-Bass, San Francisco.
- Macedonia, M. and Klimesch, W. (2014) "Long-term effects of gestures on memory for foreign language words trained in the classroom", **Mind, Brain, and Education**, Vol. 8, No. 2, pp. 74–88.
- Macrine, S.L. and Fugate, J.M.B. (2021) "Translating embodied cognition for embodied learning in the classroom", **Frontiers in Education**, Vol. 6. doi:10.3389/educ.2021.712626
- Pardede, P. (2018) "Use of mother tongue in EFL classes of secondary schools in Jabodetabek: Students' and teachers' perception", **JET (Journal of English Teaching)**, Vol. 4, No. 2, pp. 62–80.
- Plass, J.L., Homer, B.D. and Kinzer, C.K. (2015) "Foundations of game-based learning", **Educational Psychologist**, Vol. 50, No. 4, pp. 258–283.
- Prensky, M. (2001) **Digital Game-Based Learning**, McGraw-Hill, New York.
- Roediger, H.L. and Butler, A.C. (2011) "The critical role of retrieval practice in long-term retention", **Trends in Cognitive Sciences**, Vol. 15, No. 1, pp. 20–27.
- Stanovich, K.E. (1986) "Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy", **Reading Research Quarterly**, Vol. 21, No. 4, pp. 360–407.
- Swain, M. (1985) "Communicative competence: Some roles of comprehensible input and comprehensible output in its development", in Gass, S. and Madden, C. (eds.) **Input in Second Language Acquisition**, Newbury House, Rowley, MA, pp. 235–253.
- Trucano, M. (2010) "Worst practice in ICT use in education", *World Bank EduTech Blog*. Available at: <https://blogs.worldbank.org/en/education/worst-practice> (Accessed: 15 January 2026).
- Trung Cang, N., Diem, L.T.K. and Thien, L.Q. (2021) "Impacts of total physical response on young learners' vocabulary ability", **VNU Journal of Science: Education Research**, Vol. 37, No. 4. doi:10.25073/2588-1159/vnuer.4573
- Wright, A., Betteridge, D. and Buckby, M. (2006) **Games for Language Learning**, 3rd edn, Cambridge University Press, Cambridge.
- Yang, L. (2024) "A comprehensive review of total physical response", **Frontiers in Sustainable Development**, Vol. 4, No. 7, pp. 90–98. doi:10.54691/hmtwrm67