

SPEAKING TO LEARN: AN EXPLORATORY MIXED-METHODS STUDY OF RECIPROCAL MOMENTUM IN SECOND-LANGUAGE PRAGMATIC DEVELOPMENT

Desiree Dawn P. Justol-Saraspe¹, Josan C. Fermano² and Renato Base³

^{1,2} Department of Teaching Languages, University of Science and Technology of Southern Philippines, Cagayan de Oro City, Philippines;

³Department of Educational Planning and Administration, University of Science and Technology of Southern Philippines, Cagayan de Oro City, Philippines

Corresponding author: Desiree Dawn P. Justol-Saraspe; e-mail: desireedawn.saraspe@ustp.edu.ph

Date of publication; 28 July, 2026

ABSTRACT: Learning-by-teaching research indicates that preparing to explain material can activate generative processing, but its relevance to second-language (L2) pragmatic communication and the role of reciprocal peer response remain insufficiently established. This exploratory convergent mixed-methods classroom study examined a proposed Educational Reciprocal Momentum (ERM) framework among 32 Filipino senior high school students. Students completed a situational willingness-to-communicate (WTC) measure before and after a prepare-to-teach and peer-teaching activity on polite academic requests, a post-preparation teaching self-efficacy scale, and a delayed pragmatic production task 48-72 hours later. Audio-recorded teaching episodes, open-ended responses, and interviews with a purposive subsample ($n = 20$) were analyzed using hybrid codebook thematic analysis. Situational WTC increased from a median of 3.0 to 4.5 ($Z = -4.72$, $p < .001$, $r = .83$). Participants who subsequently reported higher peer engagement ($n = 24$) showed larger WTC change (median = 1.3 versus 0.3, $U = 28.0$, $p = .002$, $r = .52$), higher pre-teaching self-efficacy (4.0 versus 2.5, $U = 22.5$, $p < .001$, $r = .61$), and higher delayed pragmatic scores (4.0 versus 2.8, $U = 30.0$, $p = .003$, $r = .49$) than those reporting lower engagement ($n = 8$). Qualitative themes indicated audience-oriented planning, peer response as evidence of comprehension and relational safety, and culturally meaningful negotiation of *hiya*, *kapwa*, *pakikisama*, and *pagmamalasakit*. Because the study lacked random assignment and used post hoc engagement groups, the findings demonstrate association rather than causal amplification. ERM is therefore advanced as a provisional framework for experimental testing rather than a confirmed law.

Keywords: learning by teaching; willingness to communicate; L2 pragmatics; peer engagement; mixed methods; Filipino learners

INTRODUCTION

Learning by teaching is a generative activity in which learners reorganize content, select relevant information, and construct explanations for an audience. Experimental studies distinguish the benefits of merely expecting to teach from the additional benefits of actually teaching, while meta-analytic evidence indicates that preparing to teach and interactive teaching can improve learning when tasks require retrieval, explanation, and audience-oriented knowledge organization [1-3]. These effects are commonly attributed to generative processing, self-explanation, metacognitive monitoring, and the responsibility to make knowledge intelligible to another person. Translating such principles into classroom practice, however, also depends on teachers' pedagogical beliefs, perceived norms, and assessment intentions, which can shape the learning activities they select and the evidence they accept as demonstrating learning [20]. Existing learning-by-teaching research focuses mainly on retention of academic content; less is known about whether a prepare-to-teach sequence changes learners' readiness to use a second language in socially consequential situations.

This question matters because willingness to communicate (WTC) is a proximal determinant of whether L2 knowledge is converted into actual language use. WTC is situational rather than purely dispositional: it varies with perceived competence, task demands, interlocutor behavior, and social safety [4,5]. L2 pragmatics adds a further challenge. Learners must not only produce grammatically acceptable utterances but also select forms that are socially appropriate for a particular relationship, purpose, and degree of imposition [6]. Polite academic requests, such as asking a teacher for a deadline extension, therefore require linguistic

resources, confidence, perspective taking, and sensitivity to face and institutional roles.

Output and interaction research offers plausible mechanisms for improvement. Producing language can prompt learners to notice gaps, test linguistic hypotheses, and reflect on form-meaning relationships [7]. Collaborative dialogue can distribute problem solving across peers [8], and peer interaction is more productive when learners are prepared to respond to one another rather than merely take turns [9,10]. However, these traditions typically examine output, corrective feedback, or task engagement separately. They do not directly specify how anticipation of explaining, actual transmission, the listener's response, and the speaker's subsequent communicative readiness may form a reciprocal sequence.

The Philippine classroom provides a theoretically relevant setting because learners may interpret speaking and feedback through relational meanings that are not captured by individual competence measures alone. Concepts discussed in Filipino psychology, including *kapwa* (shared identity or relational selfhood), *hiya* (a context-sensitive concern for propriety and social evaluation), and *pakikisama* (maintaining smooth interpersonal relations), offer an emic vocabulary for examining how students describe communicative risk and support [11]. These concepts should not be treated as fixed traits shared uniformly by all Filipino learners. Rather, they are interpretive resources whose relevance must be demonstrated in participants' accounts.

To organize this problem, the present study proposes Educational Reciprocal Momentum (ERM) as a provisional heuristic: anticipating an explanation may orient a learner toward an audience; explanation may require restructuring and pushed output; and an engaged listener may provide

informational and relational signals that sustain further communication. ERM is a law and framed peer response as a moderator. Those claims exceed what the available design can establish. The present analysis therefore treats ERM as an exploratory model and asks whether its proposed elements co-occur in a classroom task, while explicitly considering competing explanations such as initial confidence, task repetition, and self-selection.

The study addressed three questions: (1) Does situational WTC change from before to after a combined prepare-to-teach and dyadic peer-teaching activity focused on a polite academic request? (2) How do students who retrospectively report higher versus lower peer engagement differ in WTC change, post-preparation teaching self-efficacy, and delayed pragmatic performance? (3) How do students interpret anticipation, peer response, and culturally meaningful relational dynamics during the peer-teaching activity?

MATERIALS AND METHODS

Research design

The study used an exploratory convergent mixed-methods design [12,13] because the research questions required quantitative evidence of change and subgroup patterns together with qualitative interpretation of anticipation, peer response, and culturally situated communicative risk. Philippine educational research likewise demonstrates the value of integrating measured outcomes with participants' accounts to explain how educational change is experienced in context [19]. Unlike the explanatory sequential design used in previous study, the present strands were collected within the same instructional sequence and integrated through joint displays and narrative comparison. The quantitative strand comprised a one-group pretest-posttest WTC assessment, post-preparation teaching self-efficacy, and delayed pragmatic performance; the qualitative strand comprised teaching recordings, open-ended notes, and interviews. Because all participants received the same activity and engagement groups were formed retrospectively, subgroup differences are interpreted as exploratory associations rather than treatment effects.

Participants and context

The study was conducted in an intact Oral Communication class in a Philippine public senior high school. All 35 students participated in the regular classroom activity; 32 students aged 16-18 years provided assent and parental or guardian consent for research participation and were included in the dataset. The class included speakers of Cebuano, Tagalog, Ilocano, and other Philippine languages. Institutional assessments placed students broadly within the B1-B2 range of the Common European Framework of Reference. The sample represents one naturally occurring class rather than the wider population of Filipino senior high school learners.

Instructional Task and Data-Collection Sequence

The target pragmatic routine was a polite academic request, illustrated by asking a teacher for an extension after a family emergency. At T0, students completed a six-item situational WTC scale. They then prepared to explain the request routine to a peer, considering appropriate wording, mitigating expressions, reasons, and possible listener questions. At T1, immediately after preparation but before teaching, they completed a four-item L2 teaching self-efficacy scale. Each student then delivered a 3-5 minute

explanation to a classmate. At T2, the WTC scale was administered again, together with the prompt, "How did your classmate react? Did they seem to understand?" Teaching episodes were audio-recorded with permission. Within 24 hours, 20 purposively selected participants completed brief semi-structured interviews. At T3, 48-72 hours after the activity, participants recorded a 30-60 second response to a new but closely related academic-extension scenario.

Measures

Situational WTC was measured with six five-point items adapted from the situational framework of MacIntyre et al. [4] and contextualized to academic requests. Scores represent task-specific readiness to speak rather than a stable WTC trait. Teaching self-efficacy was assessed with four five-point items concerning clarity of explanation, provision of examples, adjustment to listener needs, and ability to support understanding. Because self-efficacy was measured only after preparation, the study can describe its level and subgroup association but cannot estimate change in self-efficacy.

The delayed pragmatic task was rated on pragmatic appropriateness, linguistic complexity, fluency, and coherence using five-point dimensions. Two raters who were blind to the perceived-engagement classification independently scored all responses after calibration with five anchor samples. Intraclass correlation coefficients using a two-way random-effects, single-measure specification were .85 for pragmatic appropriateness, .82 for linguistic complexity, .88 for fluency, and .84 for coherence, indicating good reliability under commonly used guidelines [14]. Scores were averaged across dimensions and raters.

Perceived peer-engagement classification

After data collection, students were classified from their T2 reaction notes as reporting higher perceived peer engagement ($n = 24$) when they described verbal uptake, confirmation of understanding, collaborative elaboration, or explicit validation. Students were classified as reporting lower perceived engagement ($n = 8$) when notes described minimal uptake, continued reluctance, uncertainty, or absence of an observable response. The labels "higher" and "lower perceived engagement" replace the original "Engaged" and "Passive conditions" because no engagement condition was assigned or experimentally manipulated. Classification was based on each speaker's perception, not an independent behavioral rating of the listener.

Data analysis and integration

Given the sample size and ordinal response scales, the reported nonparametric analyses were retained [15]. A Wilcoxon signed-rank test assessed T0-T2 WTC change. Three Mann-Whitney U tests compared the post hoc engagement subgroups on WTC change, T1 teaching self-efficacy, and T3 pragmatic performance. Effect size was reported as $r = Z/\text{square root of } N$, with .10, .30, and .50 used as approximate small, medium, and large benchmarks [16]. Because the subgroup analyses were exploratory, exact p values and effect sizes are emphasized. The three reported subgroup findings remain below .01 after a Holm adjustment (adjusted $p \leq .004$).

Interview transcripts and teaching-episode records, totaling approximately 90 pages, were analyzed using a hybrid

codebook thematic approach [17]. Deductive codes represented anticipation, explanatory restructuring, peer engagement, confirming response, and culturally salient terms; inductive codes were added for recurrent patterns outside the initial framework. A second coder independently coded 20% of the corpus, yielding Cohen's kappa = .84, which was treated as a codebook-consistency check rather than proof that themes were objectively fixed [18]. Five participants reviewed preliminary interpretations, and the teacher-researcher maintained reflexive notes. Integration examined convergence, complementarity, and dissonance [13]. Joint displays connected each statistical result with qualitative evidence, plausible alternative explanations, and limits of inference so that final meta-inferences rested on both strands.

Ethical considerations

Participation was voluntary and did not affect students' grades. Students could withdraw without academic penalty. Research records were anonymized with participant codes, stored on password-protected devices, and scheduled for secure deletion after publication. Nonconsenting students completed the instructional activity but were excluded from research observation and analysis. Ethical review information must be completed before submission

RESULTS

Participant inclusion and overall WTC change

Thirty-two of the 35 students in the intact class were included in the research dataset. Situational WTC increased

across the combined preparation and peer-teaching sequence: the median score rose from 3.0 at T0 to 4.5 at T2, Wilcoxon $Z = -4.72$, $p < .001$, $r = .83$. The effect size is large, but the design cannot isolate whether the change arose from anticipation, preparation, actual speaking, peer response, repeated measurement, or their combination. Post-preparation teaching self-efficacy was high overall (median = 4.0).

Associations with perceived peer engagement

Students in the higher-perceived-engagement subgroup reported a larger median WTC increase than those in the lower-engagement subgroup (1.3 versus 0.3), $U = 28.0$, $p = .002$, $r = .52$. They also obtained higher delayed pragmatic scores (4.0 versus 2.8), $U = 30.0$, $p = .003$, $r = .49$. These results are consistent with an association between perceived reciprocal response and subsequent communicative readiness or task performance. However, the higher-engagement subgroup already had higher teaching self-efficacy at T1, before peer teaching occurred (4.0 versus 2.5), $U = 22.5$, $p < .001$, $r = .61$. This temporal detail is important: later peer response cannot have caused the earlier self-efficacy difference. More self-efficacious students may have prepared more effectively, explained more clearly, elicited stronger listener responses, or interpreted ambiguous reactions more positively.

Table 1. Summary of quantitative findings

Outcome or comparison	First value	Second value	Test	Effect size r
Situational WTC, whole sample	T0 Mdn = 3.0	T2 Mdn = 4.5	$Z = -4.72$; $p < .001$	.83
WTC change by perceived engagement	Higher: Mdn = 1.3	Lower: Mdn = 0.3	$U = 28.0$; $p = .002$	.52
Teaching self-efficacy at T1	Higher: Mdn = 4.0	Lower: Mdn = 2.5	$U = 22.5$; $p < .001$	.61
Delayed pragmatic score at T3	Higher: Mdn = 4.0	Lower: Mdn = 2.8	$U = 30.0$; $p = .003$	.49

Note: Higher perceived engagement $n = 24$; lower perceived engagement $n = 8$. T1 occurred after preparation but before peer teaching. Mdn = median; WTC = willingness to communicate.

Qualitative themes

Three themes clarified how students experienced the activity. First, anticipation prompted audience-oriented planning. Participants described selecting polite forms, anticipating confusion, and adjusting examples rather than merely memorizing a script. One student stated, "I don't just want to explain - I want them to feel safe asking questions." Another reported becoming "more confident speaking" after identifying a socially safe way to formulate the request. These accounts are consistent with generative preparation, although the study did not include an independent pre-task measure of explanation quality.

Second, students interpreted peer response as both informational and relational evidence. Verbal uptake ("She used the phrase back - that's when I knew it worked"), collaborative elaboration, and expressions such as *salamat* signaled that the explanation had been understood and socially received. In contrast, silence or minimal response

was described as uncertainty or possible rejection: "Baka mali grammar ko... nakakahiya" ("Maybe my grammar is wrong... it is embarrassing"). These accounts help explain why WTC change was larger among students who perceived stronger engagement, but they also show that perceived engagement and speaker confidence may be mutually reinforcing.

Third, some students used Filipino relational concepts to make sense of teaching. *Kapwa* appeared in statements that framed improvement as shared rather than competitive; *pakikisama* was invoked when helping a classmate preserved group harmony; and *pagmamalasakit* described care for the listener's dignity and understanding. *Hiya* was not uniformly a barrier: it initially constrained risk taking, but supportive uptake sometimes converted evaluative concern into careful, other-oriented communication. These patterns are presented as situated interpretations from this class, not as universal properties of Filipino learners.

Table 2. Integrated interpretation of quantitative patterns and qualitative themes

Analytic focus	Quantitative pattern	Qualitative evidence	Integrated inference
Preparation and WTC	WTC rose from Mdn 3.0 to 4.5 across the full activity.	Students described audience-aware planning, mitigation, examples, and concern for listener safety.	Preparation may support communicative readiness, but the effect cannot be separated from teaching and repeated measurement.
Perceived engagement	Higher engagement was associated with larger WTC change and higher T3 scores.	Uptake, collaboration, and “ <i>salamat</i> ” were read as evidence that transmission succeeded.	Peer response may sustain communication; perception, elicited response, and prior confidence are plausible competing pathways.
Cultural meaning	No quantitative cultural measure was administered.	Hiya, kapwa, pakikisama, and pagmamalasakit organized some students' accounts of risk, care, and shared learning.	Culture explains how feedback was interpreted in this class; it was not tested as a causal moderator.

DISCUSSION

A prepare-to-teach sequence and situational WTC

The large pre-post WTC difference provides initial evidence that a brief prepare-to-teach and peer-teaching sequence can coincide with increased readiness to use English for a socially demanding academic request. This pattern is compatible with learning-by-teaching research: preparing an explanation can focus attention on conceptual structure and audience needs, whereas actual teaching adds retrieval, elaboration, explanation, and interaction [1-3]. It also accords with situational models of WTC, in which communicative readiness changes with task-specific self-efficacy and the immediate social environment [4,5]. The result is educationally meaningful because the observed shift was not limited to knowledge of a request formula; it concerned students' reported readiness to use English in a face-sensitive academic situation. Nevertheless, the finding should be framed as change across the whole instructional sequence. Without study-only, preparation-only, or speaking-without-teaching comparison groups, the data do not identify anticipation as the unique cause. The effectiveness of implementing such tasks may also depend on how teachers frame, facilitate, and assess them, because teachers' pedagogical beliefs and perceived norms can shape their assessment choices and instructional behavior [20].

The nature of the target task may also explain the magnitude of the change. A polite request gives learners a concrete communicative purpose and a reusable discourse frame. Preparing to teach this routine may have reduced uncertainty about both what to say and how to protect the interlocutor's face. In pragmatic learning, this dual reduction of linguistic and social uncertainty is important because appropriate performance depends on the relationship between speaker, hearer, imposition, and context [6]. Re-administering the same WTC items after successful task completion may also have increased scores through familiarity or perceived mastery. Future experiments should therefore separate content knowledge, communicative practice, and reciprocal response.

Peer engagement as a reciprocal, not one-way, process

The subgroup findings are consistent with ERM's central proposition that a listener's response matters. Students reporting stronger engagement showed larger WTC change and higher delayed pragmatic performance. L2 interaction research similarly indicates that collaborative uptake can encourage noticing, elaboration, and continued participation [7-10]. The qualitative accounts add specificity: students did

not experience all positive responses as generic praise. Repetition of the target expression, a follow-up question, collaborative completion, or a culturally meaningful acknowledgment such as “*salamat*” served as evidence that the explanation had reached another person.

At the same time, the self-efficacy result requires a more complex model than “feedback causes momentum.” The higher-engagement subgroup differed in teaching self-efficacy before the peer-teaching episode. A confident and well-prepared speaker may produce clearer explanations, which elicit more responsive listening; confident students may also interpret equivocal cues more positively. Conversely, a receptive peer may strengthen WTC after the interaction. The plausible process is therefore reciprocal and bidirectional: preparation, explanation quality, listener behavior, and perceived success may influence one another. This interpretation is theoretically more defensible than a unidirectional moderation claim and identifies a testable mechanism for future dyadic research.

Pragmatic performance and delayed transfer

Higher delayed pragmatic scores among students reporting greater engagement suggest that the activity may have supported transfer to a related request scenario 48-72 hours later. Producing an explanation can require learners to verbalize why a form is appropriate, while interaction can expose gaps and prompt reformulation [7,8]. These processes could strengthen access to mitigation strategies and discourse organization beyond the original example. However, no baseline pragmatic-production task was administered. The subgroup difference therefore demonstrates differential delayed performance, not improvement from a known starting point. It may partly reflect pre-existing proficiency, self-efficacy, or explanation skill. A stronger study would obtain baseline and delayed pragmatic samples, use multiple request types, and model change while controlling for initial proficiency.

Cultural-relational interpretation

The cultural findings deepen rather than replace the cognitive and interactional explanation. Participants' references to *hiya*, *kapwa*, *pakikisama*, and *pagmamalasakit* show how communicative events were evaluated in relational terms. Filipino psychology emphasizes that selfhood and social action may be understood through relationships rather than through isolated individual attributes [11]. In this study, a response such as “*salamat*” could simultaneously convey comprehension, acceptance, and appreciation of the speaker's effort. Such meanings may

help create the psychological safety associated with situational WTC.

Caution is essential. Cultural concepts were elicited qualitatively and were not measured as moderators; the sample included diverse linguistic backgrounds, and participants differed in how prominently they invoked these ideas. It would therefore be inappropriate to claim that culture caused the quantitative effects or that all Filipino classrooms function identically. The defensible conclusion is that local relational meanings shaped how some students interpreted feedback. This supports culturally responsive task design while preserving within-group diversity and avoiding essentialist accounts.

Theoretical contribution of the ERM framework

The integrated findings provide partial, exploratory support for a sequence linking audience-oriented anticipation, explanation, perceived reception, and subsequent communicative readiness. ERM contributes by placing the listener's response inside the learning-by-teaching process rather than treating teaching as a purely generative act performed by an isolated learner. It also connects learning-by-teaching to L2 WTC and pragmatic performance, domains in which social reception is intrinsic to the outcome. The evidence does not justify calling ERM a law. A law implies invariant or strongly established relations, whereas the present study is small, nonrandomized, and based partly on post hoc self-report classification. ERM is better represented as a provisional explanatory framework whose pathways require prospective measurement and experimental testing.

Limitations and directions for research

Several limitations constrain inference. First, the sample comprised one intact class of 32 students, with an unbalanced post hoc subgroup division of 24 and 8; estimates may be unstable and are not population-representative. Second, the one-group design lacked randomization and comparison conditions, so maturation within the session, repeated measurement, task familiarity, speaking practice, and teacher expectancy cannot be separated from the proposed ERM process. Third, perceived engagement was coded from the speaker's post-task note rather than independently observed listener behavior. This post-treatment classification creates risks of common-method bias, reverse causality, and confounding by initial confidence or proficiency.

Fourth, self-efficacy and pragmatic performance lacked baseline measures, preventing change analysis. Fifth, the WTC and self-efficacy measures were adapted to a specific routine; full psychometric evidence and distributional summaries should be reported from the raw dataset before submission. Sixth, the delayed task occurred after only 48-72 hours and involved a closely related scenario, so longer-term retention and far transfer remain unknown. Seventh, students were nested in peer interactions, but the available analyses treated individuals as independent. Finally, interviews were conducted by an educator-researcher in a classroom context, making social desirability and interpretive positionality possible despite reflexive and coding safeguards.

A rigorous next test may preregister an ERM model and randomly assign learners to preparation-only, peer-teaching with neutral response, and peer-teaching with standardized

engaged response conditions. Likewise, it is also suggested that measures may be collected before preparation, after preparation, immediately after teaching, and at delayed follow-ups. Independent coding of listener behavior, explanation quality, and pragmatic performance would permit mediation or cross-lagged analyses. Larger multisite samples and dyadic or multilevel models could examine whether speaker confidence elicits engagement, engagement changes WTC, or both processes operate reciprocally. Cultural meanings should be investigated through mixed methods without assuming uniformity across Filipino learners or extending claims beyond the evidence.

Pedagogical implications

Within the limits of an exploratory study, the findings support low-stakes tasks in which students prepare to explain a pragmatic routine, teach it to a peer, and receive response cues that demonstrate understanding. Listeners can be trained to paraphrase, ask a clarifying question, or acknowledge the usefulness of an explanation rather than provide vague praise. Teachers may align these tasks with explicit criteria for pragmatic appropriateness, clarity, responsiveness to the listener, and reflective self-assessment. Such alignment is important because pedagogical beliefs, social expectations, and assessment intentions influence how teachers translate instructional principles into actual classroom and assessment practices [20]. Professional development should therefore address not only the mechanics of peer teaching but also the beliefs and assessment assumptions that determine whether reciprocal interaction is treated as substantive evidence of learning. Teachers may normalize linguistic risk and permit strategic use of familiar languages during planning when this supports conceptual clarity. These practices may strengthen explanation and social safety, but their effects should be evaluated with objective performance measures rather than assumed to produce momentum automatically.

CONCLUSION

In one Philippine senior high school class, situational WTC increased after a combined prepare-to-teach and peer-teaching activity on polite academic requests. Students who perceived stronger peer engagement showed larger WTC change and higher delayed pragmatic performance, while qualitative accounts linked audience-oriented preparation and reciprocal response to feelings of clarity, safety, and shared responsibility. The finding that these students already reported higher pre-teaching self-efficacy demonstrates that the process is likely bidirectional and cannot be reduced to a simple feedback effect. ERM consequently offers a useful provisional framework for studying how preparation, explanation, listener response, and communicative readiness interact; the present evidence does not establish a universal law or causal mechanism. Carefully controlled, longitudinal, and dyadic research is required before stronger theoretical or policy claims are warranted.

REFERENCES

- [1] Fiorella, L. and Mayer, R. E., "The relative benefits of learning by teaching and teaching expectancy," *Contemporary Educational Psychology*, 38(4): 281-288 (2013). <https://doi.org/10.1016/j.cedpsych.2013.06.001>

- [2] Nestojko, J. F., Bui, D. C., Kornell, N. and Bjork, E. L., "Expecting to teach enhances learning and organization of knowledge in free recall of text passages," *Memory & Cognition*, 42(7): 1038-1048 (2014). <https://doi.org/10.3758/s13421-014-0416-z>
- [3] Kobayashi, K., "Learning by preparing-to-teach and teaching: A meta-analysis," *Japanese Psychological Research*, 61(3): 192-203 (2019). <https://doi.org/10.1111/jpr.12221>
- [4] MacIntyre, P. D., Clement, R., Dornyei, Z. and Noels, K. A., "Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation," *The Modern Language Journal*, 82(4): 545-562 (1998). <https://doi.org/10.1111/j.1540-4781.1998.tb05543.x>
- [5] Leeming, P., Vitta, J. P., Hiver, P., Hicks, D., McLean, S. and Nicklin, C., "Willingness to communicate, speaking self-efficacy, and perceived communicative competence as predictors of second language spoken task production," *Language Learning*, 74(4): 917-949 (2024). <https://doi.org/10.1111/lang.12640>
- [6] Taguchi, N., "Instructed pragmatics at a glance: Where instructional studies were, are, and should be going," *Language Teaching*, 48(1): 1-50 (2015). <https://doi.org/10.1017/S0261444814000263>
- [7] Swain, M. and Lapkin, S., "Problems in output and the cognitive processes they generate: A step towards second language learning," *Applied Linguistics*, 16(3): 371-391 (1995). <https://doi.org/10.1093/applin/16.3.371>
- [8] Storch, N., "Patterns of interaction in ESL pair work," *Language Learning*, 52(1): 119-158 (2002). <https://doi.org/10.1111/1467-9922.00179>
- [9] Sato, M. and Lyster, R., "Peer interaction and corrective feedback for accuracy and fluency development: Monitoring, practice, and proceduralization," *Studies in Second Language Acquisition*, 34(4): 591-626 (2012). <https://doi.org/10.1017/S0272263112000356>
- [10] Philp, J. and Duchesne, S., "Exploring engagement in tasks in the language classroom," *Annual Review of Applied Linguistics*, 36: 50-72 (2016). <https://doi.org/10.1017/S0267190515000094>
- [11] Pe-Pua, R. and Protacio-Marcelino, E. A., "Sikolohiyang Pilipino (Filipino psychology): A legacy of Virgilio G. Enriquez," *Asian Journal of Social Psychology*, 3(1): 49-71 (2000). <https://doi.org/10.1111/1467-839X.00054>
- [12] Creswell, J. W. and Plano Clark, V. L., *Designing and Conducting Mixed Methods Research*, 3rd ed., SAGE Publications, Thousand Oaks, CA (2018).
- [13] Fetters, M. D., Curry, L. A. and Creswell, J. W., "Achieving integration in mixed methods designs: Principles and practices," *Health Services Research*, 48(6 Pt 2): 2134-2156 (2013). <https://doi.org/10.1111/1475-6773.12117>
- [14] Koo, T. K. and Li, M. Y., "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *Journal of Chiropractic Medicine*, 15(2): 155-163 (2016). <https://doi.org/10.1016/j.jcm.2016.02.012>
- [15] Field, A., *Discovering Statistics Using IBM SPSS Statistics*, 5th ed., SAGE Publications, London (2018).
- [16] Fritz, C. O., Morris, P. E. and Richler, J. J., "Effect size estimates: Current use, calculations, and interpretation," *Journal of Experimental Psychology: General*, 141(1): 2-18 (2012). <https://doi.org/10.1037/a0024338>
- [17] Braun, V. and Clarke, V., "Using thematic analysis in psychology," *Qualitative Research in Psychology*, 3(2): 77-101 (2006). <https://doi.org/10.1191/1478088706qp063oa>
- [18] O'Connor, C. and Joffe, H., "Inter-coder reliability in qualitative research: Debates and practical guidelines," *International Journal of Qualitative Methods*, 19: 1-13 (2020). <https://doi.org/10.1177/1609406919899220>
- [19] Segumpan, S. M. I., Tac-an, T. N. F., Tidal, A.-M. R., Trasmonte, C. B., Fajardo, M. T. M. and Namoco, S. O., "Empowering Educators through Technology Integration: An Explanatory Sequential Mixed Methods Study," *International Journal for Multidisciplinary Research*, 7(4): 1-11 (2025). <https://doi.org/10.36948/ijfmr.2025.v07i04.52991>
- [20] Namoco, S. O. and Zaharudin, R., "Pedagogical Beliefs and Learning Assessment in Science: Teacher's Experiences Anchored on Theory of Reasoned Action," *Journal of Turkish Science Education*, 18(2): 304-319 (2021). <https://doi.org/10.36681/tused.2021.67>