

The Role of Debate Adjudication in Developing Metacognitive Listening Comprehension: A Case Study of Undergraduate EFL Students at University of Sultan Ageng Tirtayasa

Bayu Abdul Basit^{1*}, Syafrizal², Rosmania Rima³
Universitas Sultan Ageng Tirtayasa^{1,2,3}

*Corresponding Author's Email: bayuabdulbasit@gmail.com

Received: 06 24 2026 | Accepted: 06 29, 2026 | Published: 07 01, 2026

ABSTRACT

The majority of listening lessons for English as a foreign language (EFL) students in debate-based learning environments put students in the role of debaters, rather than the adjudicator, who might be more effective in helping students acquire metacognitive listening awareness. This qualitative case study set out to do four things: (1) explain how adjudication experience helps build metacognitive listening comprehension; (2) determine how much adjudication improves monitoring and evaluating skills; (3) find out what difficulties EFL students encounter when they actually have to be the adjudicator; and (4) compare the levels of metacognitive listening awareness between new and experienced adjudicators. Participating were four undergraduates chosen using a selective selection process from the English Debate Community (EDC UNTIRTA). Stimulated Recall Tasks (SRT), semi-structured interviews, and a modified Metacognitive Awareness Listening Questionnaire (MALQ) were used to gather data, which was then analyzed using theme analysis and triangulation. The results showed that experience with adjudication activates the planning, monitoring, and evaluating parts of metacognitive listening; the most consistently developed parts among the participants were the monitoring and evaluating parts. Time constraints and fast speech rates were obstacles that had to be overcome by using tactics like contextual inference and note-taking. When compared to their less experienced peers, accredited adjudicators exhibited superior metacognitive listening habits in terms of consolidation and multidimensionality. The results of this research show that debate adjudication is a great way to help English as a foreign language students develop their metacognitive listening skills in a real-world setting.

Keywords: debate adjudication; metacognitive listening comprehension; EFL; metacognitive strategies; case study

How to Cite:

Basit, B. A. ., Syafrizal, S., & Rima, R. (2026). The Role of Debate Adjudication in Developing Metacognitive Listening Comprehension: A Case Study of Undergraduate EFL Students at University of Sultan Ageng Tirtayasa. *Educational Journal*, 1(4), 2249-2257. <https://doi.org/10.63822/d7at4y75>

INTRODUCTION

The ability to listen is often considered to be the most important receptive skill for learning a new language. Listening comprehension in an EFL environment is more than just taking in what others say; it's a complex cognitive process that requires real-time integration of contextual information, meaning construction, and sound decoding (Vandergrift & Goh, 2012). Academic settings increase the difficulty of this work since students have a lot of time to understand logical reasoning, pick out important points, and form opinions.

A growing number of studies have shown that debate-based learning may improve students' English listening abilities by encouraging them to think critically and utilize real language (Kassem, 2021; Nasution, 2019). In contrast, the majority of current debate-based learning methods put students in the position of debaters, ignoring the debate adjudicator, a job that is both functionally rich and has distinct cognitive demands. All while listening for propositional information, monitoring comprehension, evaluating logical coherence, and producing a structured judgment—this all happening in real time. The metacognitive aspects of listening comprehension, first proposed by Flavell (1979) and later operationalized in the context of second language listening by Vandergrift and Goh (2012), are directly engaged by this constellation of demands (2012).

The ability to monitor and control one's own listening processes is known as metacognitive listening comprehension. It consists of three main parts: preparing (doing some work before listening), monitoring (keeping an eye on how well you're doing while listening), and evaluating (thinking about what you heard and making a decision based on that) (Goh, 2008; Vandergrift & Goh, 2012). While these abilities are fundamental for academic listening, they are often neglected in conventional EFL classrooms.

Chero (2023), Maftoon and Alamdari (2020), and Robillos and Bustos (2022) are only a few examples of the empirical studies that have shown how teaching EFL students metacognitive strategies improves their listening comprehension. While research has looked at how debate affects skills like critical thinking, oral production, and argumentation (Oros, 2007; Ukume & Ugah Uguma, 2017), the unique impact of the adjudicator role on the development of metacognitive listening has been mostly ignored. This research explores the possible developmental benefits of debate adjudication for undergraduate English as a foreign language (EFL) students by framing it as an innate metacognitive listening environment.

The following four questions guided the research design of this study: (1) How may students' roles as judges help them build metacognitive listening comprehension skills? (2) How much do you think the ability to monitor and evaluate is enhanced when adjudicators are involved? What are some of the obstacles that EFL students have while trying to play the role of adjudicator? (4) When comparing inexperienced and certified adjudicators, how does their metacognitive listening awareness stack up?

Both the theoretical and practical implications of this work are important. Debate adjudication is a genuine, task-embedded mechanism for metacognitive learning that does not need explicit, instructor-delivered strategy teaching; this finding expands the metacognitive listening literature. In terms of practical application, it gives debate practitioners and EFL instructors a justification based on research to include adjudicator responsibilities into listening teaching.

METHODS OF RESEARCH

Research Design

Merriam (2009) and Stake (1995) both refer to the use of a qualitative case study approach in this work. The research goal was to examine the lived experiences of a particular, limited group within a naturalistic situation, and a case study technique was used for this purpose. The chosen group was undergraduate EFL students who served as debate adjudicators. Consistent with Stake's (1995) idea of the instrumental case study, the example was used to shed light on wider pedagogical implications for English as a foreign language listening lessons.

Research Site and Participants

An English Debate Club at Indonesia's University of Sultan Ageng Tirtayasa (UNTIRTA) in Serang was the setting for the research. Using purposive sampling, we chose four undergraduate EFL students who met the following criteria: (1) they were active or recently were members of EDC UNTIRTA; (2) they had prior experience as adjudicators in formal English-language debate competitions organized under the British Parliamentary (BP) or Asian Parliamentary (AP) format; and (3) they were willing to participate in all data collection instruments and provide informed consent.

To facilitate comparison analysis, participants were divided into two groups: one group consisted of one accredited adjudicator (P4) with credentials from regional and national tournaments and more than fifteen rounds of formal adjudication experience, and the other group consisted of three novice adjudicators (P1-P3) with ten or fewer rounds of formal adjudication experience. Table 1 displays the participant demographics.

Table 1. Participant Profile

Code	Name	Faculty / Program	Category	Adj. Rounds	Accreditation	Instruments
P1	Bunga	FKIP / English Education	Novice	4–7	None	Q, I, SRT
P2	Farhan	FKIP / English Education	Novice	4–7	None	Q, I, SRT
P3	Gibran	FEB / Accounting	Novice	1–3	None	Q, I, SRT
P4	Pniel	FISIP / Communication Science	Accredited	15+	Regional/National (Certificate)	Q, I, SRT

Q = Questionnaire; I = Interview; SRT = Stimulated Recall Task. Novice participants identified by first name only to protect anonymity.

Data Collection Techniques

Data triangulation was ensured by using three supplementary instruments.

First, all participants were given a modified version of the Metacognitive Awareness Listening Questionnaire (MALQ) by Vandergrift et al. (2006) using Google Form. Planning and Evaluation (PE), Problem-Solving (PS), Mental Translation (MT, reverse-coded), Person Knowledge (PK), and Directed Attention (DA) are the five dimensions that make up the 21 Likert-scale questions that make up the MALQ.

Here, 1 = Strongly Disagree and 5 = Strongly Agree. The adjudication setting was taken into account when the instrument was modified.

The second step was to interview each participant separately using a semi-structured interview. Questions on adjudication, round listening habits, metacognitive strategy usage, difficulties faced, and perceived listening ability growth were all part of the interview guide. With participants' informed agreement, online interviews were videotaped using Zoom or Google Meet and transcribed word for word.

Third, in order to capture metacognitive thinking that was happening almost instantly, participants were given Stimulated Recall Tasks (SRT). After seeing a taped segment of the NUDC 2023 Grand Final debate (Universitas Gadjah Mada v. Universitas Udayana v. STBA PIA v. Universitas Telkom) from 09:28 to 45:34, participants verbalized their tactics for listening, managing their attention, and evaluating the argument using retrospective verbal procedures.

Data Analysis

First, the data was familiarized with. Then, initial codes were generated using both deductive (MALQ dimensions) and inductive approaches. Then, themes were searched for. After that, themes were reviewed. Finally, themes were defined and named. Finally, a report was written. This was all done using Braun and Clarke's (2006) six-phase thematic analysis model. In order to supplement the qualitative results, quantitative analysis was performed on the MALQ descriptive data at the dimension level. For descriptive statistics, IBM SPSS Statistics Version 26 was used. All three data sources were used in a systematic triangulation to ensure reliability.

RESULTS AND DISCUSSION

MALQ Descriptive Findings

Table 2 shows the mean scores of the MALQ dimensions broken down by group of adjudicators. Due to the advanced metacognitive starting position of all participants as active EDC members, the overall mean scores were similar between the novice adjudicator ($M = 3.47$) and the certified adjudication ($M = 3.52$). With a mean score of 3.67 and a standard deviation of .236 on the Directed Attention (DA) dimension, attention control emerged as the metacognitive skill that was best developed in this group. P1 (Bunga) had the highest Mental Translation (MT) score (4.00), suggesting little dependence on L1 compared to the other participants, and this resulted in the lowest mean ($M = 3.25$, $SD = .500$) and the largest within-group variation in MT.

Table 2. MALQ Dimension Mean Scores by Group

MALQ Dimension	Novice M (n=3)	Accredited M (n=1)	Δ	Overall M (SD)
Planning & Evaluation (PE)	3.58	3.50	-0.08	3.56 (.125)
Problem-Solving (PS)	3.50	3.75	+0.25	3.56 (.375)
Mental Translation (MT)*	3.33	3.00	-0.33	3.25 (.500)
Person Knowledge (PK)	3.33	3.50	+0.17	3.38 (.250)
Directed Attention (DA)	3.61	3.83	+0.22	3.67 (.236)
Overall MALQ Mean	3.47	3.52	+0.05	3.49 (.246)

Scale: 1 = Strongly Disagree, 5 = Strongly Agree. *MT scores are reverse-coded; higher scores indicate less reliance on L1 mental translation. Δ = Accredited minus Novice mean.

Theme 1: Evaluative Listening as Strategic Monitoring

The thing that all four people in the judge role had in common was that they all planned and evaluated activating strategies. All of the people who took part showed that adjudication's standards for evaluation create a fundamentally metacognitive listening setting by having to prepare before hearing, look at points as they were being made, and think about what they had heard after each round.

Participants who were new to the activity described planned movements that were clearly hard work. "Fokus aja, pikirin apa yang harus aku pikirin sekarang yaitu menilai argumen, jangan mikirin yang lain-lain" (Just concentrate, think about analyzing the arguments, don't think about other things), according to P3 (Gibran), who reported an intentional commitment of attention. An developing PE approach that aligns with the MALQ Planning and Evaluation dimension is this deliberate self-instruction. P1 (Bunga) explained how to be ready for rounds by reviewing topics in advance, which allows for anticipatory argument formulation based on schemas. P2 (Farhan) used a nuanced person-knowledge tactic known as "neutrality-as-preparation," in which he consciously set aside his previous knowledge of the subject in order to listen to the debaters' framing without prejudice.

However, the certified adjudicator (P4, Pniel) outlined automated evaluation planning using a multi-criteria framework that includes effect depth assessment, counterfactual reasoning, and spirit of the motion analysis. Instead of seeing his 2.00 (Disagree) PE1 score as evidence of a lack of preparation, it was seen as evidence of its internalization; experienced adjudicators don't 'prepare to concentrate' on purpose since the evaluative framework activates automatically. Bozorgian et al. (2022) described the move from effortful to automatized metacognitive control as the characteristic of skilled L2 listeners, and our conclusion accords with their findings.

These results are in line with the metacognitive listening development model proposed by Vandergrift and Goh (2012), whereby the methods for planning are internalized via repeated practice of challenging listening tasks. By including metacognitive responsibility into the task structure, the adjudication function seems to hasten this developmental trajectory. Adjudicators are required to express their evaluations verbally, turning their internal monitoring into communicable assessment.

Theme 2: Inference-Making Under Linguistic Uncertainty

The second topic was about how people dealt with not being able to understand because of things like foreign language, badly organized arguments, and fast speech rates. It touched on issues that relate to the Problem-Solving (PS) and Mental Translation (MT) parts of the MALQ.

There was a clear range of PS techniques that each person used. "Kecepatan berbicara dia lumayan tinggi, jadi aku terpecah fokusnya antara harus mendengarkan argumen dia yang sekarang banget sama yang udah berlalu" (his speaking speed was fast, so my attention was split)," P3 (Gibran), the judge with the least experience, admitted. Instead of actively collecting information, he used inferential placeholders to show where things were missing, which shows that he was more reactive than proactive in his PS participation. This was his way of getting better.

Farhan, P2, took a more philosophical PS method when he came across the new word "equilibrium." He didn't go back to the L1 version. He stopped the understanding gap and listened to the next speech to see how the speaker added to what they had said. This is a top-down approach that fits with Field's (2009) model of how meaning is made through discourse. During rounds, P1 (Bunga) underlined words that were hard for her and then went back to them during the test, which combined taking notes with actively reconstructing what they meant. P4 (Pniel) had the most advanced PS behavior: it made proactive evaluation decisions instead of trying to understand better after failing. It did this by using real-time evaluative reasoning, which found both the substance and logical flaws of the case.

Based on P4's MT score of 3.00, which is the same as the least experienced beginner, it seems that judging experience alone is not enough to completely get rid of L1 translation habits. In line with Vandergrift et al.'s (2006) observation that MT tendencies are partly trait-like traits, this finding suggests that even experienced judges still process some L1 information when they hear different accents or phonological diversity in debate discourse, like at national tournaments.

Theme 3: Self-Regulated Attention During Debate Rounds

All participants agreed that the third theme, steady directed attention across multi-speaker discussion rounds, was the most important problem and that it fit best with the DA dimension. The qualified judge got the best score (3.83), and the beginners had the most varied scores (P1: 3.83; P3: 3.33).

Attentional methods were found to be very different between subjects. For P3 (Gibran), the main focus was on safety measures, like getting rid of physical distractions and managing the environment, but these weren't enough to meet the needs of real-time processing. To stay focused, taking notes is the most important thing you can do. Your mind is still working while you're taking notes. P2 (Farhan) described a way to take notes that is creative. Anderson (2022) says that external job organization is a way to deal with active processing. The fact that writing can help with deep thinking is proof of this.

As Bunga, the first member, said, "I don't find myself in a stressful situation as an adjudicator because I don't have to unwind a lot of things while I'm being a debater." This result fits with Anderson's (2022) idea of cognitive load: when there is no need to make something at the same time, attentional resources can be used to check and see if knowledge is being maintained. P4 (Pniel) said that his main goal was to take detailed notes at the word level, which helped him easily follow the points of many speakers during long discussion rounds.

Theme 4: Self-Efficacy Beliefs and Listener Confidence

This section talked about how metacognitive listening is connected to self-efficacy views (the PK dimension) in the fourth part. The known judge had a slightly higher PK score (3.50 vs. 3.33 for the beginner), but the qualitative data showed that there was a big difference in how confidence profiles changed as the judge gained more experience.

P3 (Gibran) had the least trust of the three. He said his listening skills were fine for intermediate-level knowledge but that he had trouble in high-demand scenarios. In line with what Chero (2023) calls "functional-but-uncertain self-efficacy" in early metacognitive development, he accepted that he didn't understand instead of becoming anxious and acting out. Bunga, the first person, showed that she could tell the difference between her general tendency to process sounds and her academic English hearing skill. This

showed that she could tell the difference between sources of trouble at the person level and those at the language level. Bunga gave her an 8 out of 10 for her academic English hearing skills. P2 (Farhan) had the highest beginning self-efficacy based on his extensive debate experience and SRT performance, in which he followed the points of all four speakers without showing any signs of not understanding.

It's interesting that P4 (Pniel) said he had a brief doubt during a high-stakes grand final decision, but this doubt didn't affect how he was judged during the SRT. This shows that he has reached the level of metacognitive growth needed to tell the difference between listener confusion and regulatory success. This fits with what Baki (2025) says about expert metacognitive listeners.

Comparative Profile: Novice vs. Accredited Adjudicators

The main thing that set novice and certified adjudicators apart from each other was not numeric factors, but qualitative ones. These included the regularity, reliability, and automation of metacognitive methods. The overall MALQ difference in favor of certified adjudication is small. Novices' attempts at metacognition were often hard, self-aware, and not always used correctly. The most noticeable sign of competence was the certified adjudicator's integrated, automated, and reliable controlling behaviors. These behaviors showed the PE1 score inversion (2.00 vs. beginner mean 3.33), which shows that planning processes were internalized beyond the level of conscious awareness.

The starting members could have been very different from each other within the group. The qualified judge (3.52) lost to P1 (Bunga, MALQ total: 3.77) because she had better MT and DA scores. This was because she had taught English and used active strategies for thinking about cases. The lowest score was given to P3 (Gibran, total: 3.17), which could be because it doesn't have much experience with judging and comes from a field other than language (Accounting). It's clear from this difference that the growth of metacognitive listening skills through adjudication is affected by things other than the frequency of adjudication. These include exposure to language studies in the classroom, English language skills, and general debate experience.

These results broaden the scope of earlier studies in many ways. To begin with, they back up the findings of Chero (2023) and Maftoon and Alamdari (2020) regarding the development of metacognitive listening based on structured, evaluative, and reflective task engagement. However, they show that this development can also happen through authentic task participation instead of just through explicit instruction. Secondly, they demonstrate that adjudication is an evaluative and multimodal listening environment that occurs spontaneously, which advances Cheng's (2023) study on multimodal metacognitive listening training. Thirdly, they add to the research on debate as a learning tool by distinguishing the adjudicator function from the debater position as an experience with a high metacognitive yield (Kassem, 2021; Ukume & Ugah Uguma, 2017).

CONCLUSION

In this study, the metacognitive listening understanding skills of first-year English as a Foreign Language (EFL) students at UNTIRTA were checked. Four main findings can be drawn from the results.

One important role of adjudicating debates is providing a natural and effective space for learning how to listen with the intention of understanding what is being said. All five MALQ dimensions are

naturally triggered by the ruling duties of a judge, which include understanding, following, evaluating, and passing judgment on speech arguments according to Vandergrift and Goh (2012). Planning, keeping an eye on, and reviewing are some of the most important metacognitive processes.

Second, participants' experience with judging led to more structured methods for tracking comprehension in real time and evaluating reflections after each round. This made watching and evaluating aspects the most regularly developed metacognitive skills. One important part of the role of an arbitrator that changed over time was the need to share decisions based on externalized internal tracking. This is called metacognitive duty.

Third, people had a lot of trouble with things like long-term demands on their attention, different sounds, and speaking quickly. Some of the strategies used to solve these problems were active note-taking, environmental reasoning, deliberate neutrality, and schema activation. It was easier for new adjudicators to experience understanding disturbance than for more experienced ones, who relied on more automatic and reliable ways to deal with stress.

Fourth, comparing new and qualified adjudicators showed that getting better at the process over time helps your metacognitive skills grow. The changes in numeric MALQ scores were not very big, which makes sense given that EDC UNTIRTA members had already completed a lot of work. When looking at the dependability, regularity, and automaticization of metacognitive control, on the other hand, qualitative analysis showed important differences that self-report tools alone missed.

These results could help people who make lesson plans, run debate communities, or teach English as a foreign language (EFL). As part of debate-based listening training, the roles of judge and debater should be carefully added. There should also be structured post-round review and clear evaluation standards. Metacognitive processes are used in judicial interactions, even if they are short. This might not be the case with idle listening. Future research should build on these findings by using bigger ongoing designs and quantitative comparisons between before and after to get a better idea of how metacognitive listening awareness grows through judging experience.

REFERENCES

- Anderson, N. J. (2022). The role of metacognition in second language teaching and learning. ERIC Digest. <https://www.eric.ed.gov>
- Bozorgian, H., & Shamsi, E. (2025). A review of research on metacognitive instruction for listening development. *International Journal of Listening*, 39(1), 1–16. <https://doi.org/10.1080/10904018.2023.2197008>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Cheng, H. Y. (2023). Effects of metacognitive strategies on listening comprehension through multimedia supported instruction. <https://doi.org/10.21203/rs.3.rs-3233113/v1>
- Chero, C. A. C. (2023). The impact of metacognitive instruction on EFL low-level learners' listening performance and metacognitive awareness. *International Journal of Instruction*, 16(2), 291–306. <https://doi.org/10.29333/iji.2023.16217a>

- Field, J. (2009). *Listening in the language classroom*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511575945>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906–911.
- Goh, C. (2008). Metacognitive instruction for second language listening development: Theory, practice and research implications. *RELC Journal*, 39(2), 188–213. <https://doi.org/10.1177/0033688208092184>
- Kassem, H. M. (2021). Training EFL learners on debating: Effects on their oral and written performance, ideal L2 self, and communication apprehension. *MEXTESOL Journal*, 45(4).
- Krashen, S. D. (1985). *Principles and practice in second language acquisition*. Pergamon.
- Maftoon, P., & Alamdari, E. F. (2020). Exploring the effect of metacognitive strategy instruction on metacognitive awareness and listening performance through a process-based approach. *International Journal of Listening*, 34(1), 1–20. <https://doi.org/10.1080/10904018.2016.1250632>
- Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.
- Nasution, D. K. (2019). Improving the students' achievement in listening by applying debate plus method. *Indonesian Journal of Education and Mathematical Science*, 1(1). <https://doi.org/10.30596/ijems.v1i1.3912>
- Oros, A. L. (2007). Let's debate: Active learning encourages student participation and critical thinking. *Journal of Political Science Education*, 3(3), 293–311. <https://doi.org/10.1080/15512160701558273>
- Robillos, R. J., & Bustos, I. G. (2022). Learners' listening skill and metacognitive awareness through metacognitive strategy instruction with pedagogical cycle. *International Journal of Instruction*, 15(3), 393–412. <https://doi.org/10.29333/iji.2022.15322a>
- Rost, M. (2011). *Teaching and researching: Listening*. Routledge. <https://doi.org/10.4324/9781315833705>
- Seo, B. (2022). *Good arguments: How debate teaches us to listen and be heard*. Penguin Press.
- Stake, R. E. (1995). *The art of case study research*. SAGE Publications.
- Ukume, G., & Ugah Uguma, V. (2017). Debate as an effective teaching and learning strategy for developing listening comprehension in English as a second language classroom. *Language Learning*, 56(3), 431–462.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191–210. <https://doi.org/10.1017/S0261444807004338>
- Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.
- Vandergrift, L., Goh, C. C. M., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The Metacognitive Awareness Listening Questionnaire: Development and validation. *Language Learning*, 56(3), 431–462. <https://doi.org/10.1111/j.1467-9922.2006.00373.x>
- Zarrabi, F. (2020). Investigating the relationship between learning style and metacognitive listening awareness. *International Journal of Listening*, 34(1), 21–33. <https://doi.org/10.1080/10904018.2016.1276458>