



## Research Article

# AUDIO-SYNCHRONOUS ONLINE DISCUSSIONS AS SPACES FOR CRITICAL THINKING AND ORAL COMMUNICATION PRACTICE IN HIGHER EDUCATION

**Le Dinh Tung**

*Ho Chi Minh City University of Education, Vietnam*

*Corresponding author: Le Dinh Tung – Email: [tungld@hcmue.edu.vn](mailto:tungld@hcmue.edu.vn)*

*Received: May 13, 2026; Revised: August 10, 2026; Accepted: August 28, 2026*

## ABSTRACT

*This exploratory qualitative study examines how audio-synchronous online discussions provide opportunities to practice critical thinking and oral communication. Previous research has examined participation and collaboration in synchronous learning, but has paid less attention to the interactional quality of voice-based discussions. Data from seven sessions involving 20 students and three facilitators comprised discussion transcripts, 65 reflections, 18 interviews, and observation notes. These data were examined using thematic and qualitative content analyses. Audio provided a dialogic space with fewer visual cues but continued reliance on voice: students valued reduced visual exposure but experienced ambiguity in turn-taking, hesitation, and discomfort with speaking aloud. Critical thinking was evident in questioning, justification, challenge, and synthesis, whereas oral communication was understood as fluency in interaction. The findings suggest that audio can support these practices when discussions are designed as structured dialogic activities.*

**Keywords:** audio-synchronous discussion; critical thinking; dialogic pedagogy; higher education; oral communication

## 1. Introduction

Synchronous online learning has become an important part of higher education because it enables real-time interaction among instructors and students. Unlike asynchronous environments, in which learners participate at different times, synchronous environments allow students to exchange ideas immediately, respond to peers, and engage in live pedagogical interactions. Here, synchronous online learning refers to participants being present at the same time and exchanging contributions without appreciable delay. This temporal definition is useful because the challenges at issue—finding an opportunity to speak, responding to a previous contribution, and managing a turn that cannot be revised—arise from real-time interaction. Asynchronous modes allow more time for

---

**Cite this article as:** Le, D. T. (2026). Audio-synchronous online discussions as spaces for critical thinking and oral communication practice in higher education. *Ho Chi Minh City University of Education Journal of Science*, 23(8), 1964-1973. [https://doi.org/10.54607/hcmue.js.23.8.5803\(2026\)](https://doi.org/10.54607/hcmue.js.23.8.5803(2026))

reflection, whereas synchronous modes offer immediacy (Hrastinski, 2008). Audio preserves real-time speech but removes visual cues: it may reduce pressure to appear on camera while requiring students to speak and making turn-taking and interpretation more difficult.

Previous research shows that synchronous participation is shaped by the communication medium, topic sensitivity, and participant characteristics. Audio chat is not a neutral channel (Blau & Barak, 2012). Research also points to different affordances and challenges across modes: asynchronous discussion may stall at exploration without facilitation (Garrison et al., 2001), whereas synchronous modes offer immediacy rather than time to compose responses (Hrastinski, 2008). Voice chat can produce more negotiation but also greater anxiety (Satar & Özdener, 2008), and paralinguistic cues can shape turn initiation (Yamada, 2009). These findings address participation more directly than the ways students reason together. Contributions may remain parallel monologues (Hathorn & Ingram, 2002); by contrast, Mercer (2004) distinguishes cumulative from exploratory talk as different ways of constructing knowledge through interaction. This study examines the quality of interaction among participants. Three gaps motivate the research. First, critical thinking in synchronous environments has mainly been examined through text-based communication rather than spoken interaction, in which contributions cannot readily be revised. Second, studies of voice communication have largely focused on participation and learner anxiety. Third, audio has often been treated as less effective than video-based interaction. The study therefore addresses two research questions:

***RQ1. How do university students experience audio-synchronous online discussions as spaces for critical thinking and oral communication practice? RQ2. What interactional and facilitative features of audio-synchronous discussions support or constrain students' critical thinking and oral communication development?***

This study contributes to research on online dialogic pedagogy by examining audio-synchronous discussion as a voice-based communication environment that retains some features of face-to-face interaction. The educational value of this form of interaction depends on task design, facilitator practices, and the quality of peer responses. Effective audio-based learning, therefore, requires careful attention to how activities are organized and how learners respond to one another.

### ***1.1. Medium affordances and constraints in audio-synchronous discussion***

Synchronous online discussion should not be treated as a homogeneous category. Face-to-face discussion, video conferencing, audio chat, and text chat offer different communicative resources and present different challenges. Blau et al. (2017) emphasize the importance of interactions between instructors and students. Media Naturalness Theory considers the extent to which a medium resembles face-to-face communication. Audio-synchronous discussion retains real-time spoken communication but generally lacks facial expressions, gestures, and visual co-presence. Daft and Lengel (1984) classified media according to the richness of their cues and the immediacy of feedback. A medium with

fewer cues is thus constrained differently rather than necessarily deficient (Kock, 2005). Walther (1996) argues that communication can develop in reduced-cue environments in ways that media richness theory may not predict, as participants adapt to the available channel (see also Barak & Suler, 2008).

### 1.2. *Interactional practices: from participation to dialogic reasoning*

For Bakhtin (1981), each utterance responds to previous utterances and anticipates further responses. Understanding thus develops through interaction and responsiveness. Alexander (2020) argues that talk becomes educational when it is collective, reciprocal, supportive, cumulative, and purposeful. Mercer (2004; see also Mercer & Littleton, 2007) describes exploratory talk as critical yet constructive discussion in which participants make their reasoning explicit and seek agreement through argumentation, consistent with Vygotsky's (1978) perspective. Wegerif (2007) similarly argues that digital environments do not automatically produce dialogue but can be designed to create opportunities for dialogic interaction.

Accordingly, this study views critical thinking as a dialogic practice rather than solely as an individual cognitive outcome. It examines observable interactional moves, including questioning, justification, clarification, challenge, reconsideration, peer uptake, and synthesis. This approach draws on the exploratory-talk tradition, in which reasoning can be examined through conversation rather than measured solely through self-report (Mercer, 2004; Mercer & Littleton, 2007), as well as Alexander's (2020) criteria. Because coding isolated moves could emphasize form over reasoning, the moves were examined in sequence. An episode was considered evidence of critical thinking only when others took up a contribution and it changed the subsequent exchange.

Dialogue does not emerge from the communication channel alone. Anderson et al. (2001) identify three aspects of teaching presence: design and organization, facilitation of discourse, and direct instruction.

### 1.3. *Conceptual framework*

These two perspectives provide complementary analytical layers.



**Figure 1.** Conceptual framework for audio-synchronous dialogic discussion

As shown in Figure 1, audio-synchronous discussion is conceptualized as a voice-based medium shaped by its affordances and constraints, interactional practices, facilitative

conditions, and learning-oriented practices. This framework guided the analysis without assuming that audio necessarily improves learning.

## 2. Methodology

### 2.1. Research design, context, and participants

This study used an exploratory qualitative design. It aimed to examine how participants experienced and constructed audio-synchronous discussion as a learning space, rather than to test its statistical effect on achievement. A qualitative design was appropriate because the research questions concerned experience, interactional processes, and facilitative conditions.

The two research questions required different kinds of evidence. RQ1 concerned experiences that could not be observed directly: silence, hesitation, and brief contributions could reflect disengagement, processing, anxiety, or strategic listening. Reflections and interviews therefore allowed participants to explain what the transcripts alone could not reveal. RQ2 concerned what happened during the discussions and the conditions surrounding productive or unproductive episodes; it therefore required evidence from the discussions themselves. The two sources of evidence were then compared so that reported experiences could be considered alongside interactional evidence and discrepancies could be treated as analytically informative. Table 1 summarizes this alignment.

**Table 1.** Alignment of research questions, data sources, and analysis

| Research question                            | Primary data                            | Supporting data                       | Analysis                     |
|----------------------------------------------|-----------------------------------------|---------------------------------------|------------------------------|
| RQ1: Student experience                      | 65 reflections; 15 student interviews   | Facilitator interviews                | Thematic analysis            |
| RQ2: Interactional and facilitative features | 7 transcripts; 3 facilitator interviews | Observation notes; student interviews | Qualitative content analysis |

The study was conducted in a university learning context with 20 students and three facilitators. All 20 students participated in the seven sessions and were represented in the transcript and reflection corpus; 15 also took part in semi-structured interviews. The three facilitators led the sessions and contributed interviews and observation notes. Prompts addressed required participation, inclusion in audio-only settings, silence and engagement, critical thinking, assessment of oral contributions, fluency and questioning, and the design of productive audio discussions. Participants were recruited through purposive sampling with a maximum-variation strategy that encompassed habitual initiators, students who contributed after listening, and students who rarely spoke. Variation in year level and discipline was also sought. Participation was voluntary and independent of assessment.

The student group included 10 female and 10 male students across four year levels: Year 1 (n = 4), Year 2 (n = 8), Year 3 (n = 5), and Year 4 (n = 3). Their disciplinary backgrounds were English Studies (n = 5), Business English (n = 4), TESOL (n = 4), Communication (n = 3), Translation (n = 2), and Applied Linguistics (n = 2). For analysis, the students were grouped into three descriptive participation profiles: active speakers (S01–S06), moderate participants (S07–S14), and quiet or less confident participants (S15–S20).

The profiles supported variation in the analysis and were not treated as fixed psychological categories. Table 2 presents the profiles by participation orientation because the study focuses on how students experienced the same audio-synchronous environment in different ways.

Active speakers were comfortable initiating turns and extending discussions. Moderate participants contributed after listening to several peers and often served as a bridge between contributions. Quiet or less confident participants were less visible in spontaneous talk; their accounts offered insight into anxiety, wait time, ambiguity in turn-taking, and conditions for equitable participation.

**Table 2.** *Participant profile and data contribution*

| Participant group                 | IDs and number                 | Background profile and analytic relevance                                                                                                                                                                                                                   |
|-----------------------------------|--------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Active speakers                   | S01-S06; n=6; 3 female, 3 male | Year 2-4 students from English Studies, Business English, TESOL, Applied Linguistics, and Communication. They generally reported high or medium-high confidence and often initiated claims, challenged ideas, gave examples, or synthesized peer responses. |
| Moderate participants             | S07-S14; n=8; 4 female, 4 male | Year 1-4 students from TESOL, Business English, English Studies, Translation, and Communication. They usually entered after listening to several viewpoints and often served as connectors between active and quieter students.                             |
| Quiet/low-confidence participants | S15-S20; n=6; 3 female, 3 male | Year 1-4 students from English Studies, TESOL, Translation, Business English, Communication, and Applied Linguistics. They were central to analysis of silence, hesitation, wait time, invitation, pronunciation anxiety, and missing visual cues.          |
| Facilitators                      | F01-F03; n=3                   | The facilitator roles were course instructor, teaching assistant, and language-skills lecturer. Their contributions were used to interpret task design, turn-taking management, facilitative probing, oral communication support, and observation notes.    |

## **2.2. Data sources and procedures**

Data came from five sources: seven discussion transcripts, 65 student reflections, semi-structured interviews with 15 of the 20 students, three facilitator interviews, and facilitator observation notes. These sources enabled comparison of interactional practices with participants' retrospective accounts. Session prompts asked students to compare alternatives, justify choices, and respond to peers. Topics included whether live participation should be required, whether audio was more inclusive than video, and how silence should be interpreted. The prompts were chosen to elicit disagreement rather than statements of preference. Reflections captured experiences that transcripts could not reveal, including anxiety, delayed processing, and self-evaluation. Observation notes provided context for silence and overlapping speech.

## **2.3. Data analysis, trustworthiness, and ethics**

The reflections and interviews were examined using thematic analysis, whereas the transcripts were coded for dialogic moves using qualitative content analysis. The analyses

were then compared to assess whether interactional evidence aligned with participants' reported experiences.

### 3. Results and discussion

The analysis identified four interconnected themes: (1) audio as a dialogic space with fewer visual cues but continued reliance on voice, (2) critical thinking as peer-responsive reasoning, (3) oral communication as fluency in interaction, and (4) facilitation as dialogic design. Figure 2 shows the relationships among these themes and their main components.

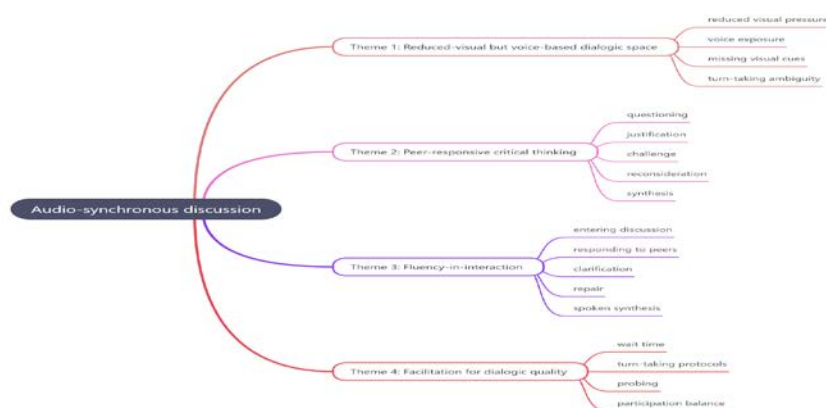


**Figure 2.** Thematic map of audio-synchronous discussion findings

#### 3.1. Audio as a reduced-visual but voice-based dialogic space

Students consistently described audio discussion as less visually demanding than video or face-to-face discussion. Some felt more comfortable because they did not have to manage their camera presence, appearance, or background. However, the absence of gaze, gesture, and facial expression also made it harder to judge when to speak.

S08: Audio is inclusive for me because I do not worry about my camera or room.

S17: But I cannot see when someone wants to speak, so I sometimes stop myself.

S02: Maybe audio removes visual pressure but creates turn-taking pressure.

Silence did not necessarily indicate disengagement; it could also reflect thinking, anxiety, or difficulty finding a suitable opportunity to enter the discussion. A cultural interpretation warrants caution: East Asian reticence has been widely discussed, but explanations of it remain contested (Cheng, 2000; Liu & Jackson, 2008).

#### 3.2. Critical thinking through peer-responsive reasoning

Critical thinking was evident when students moved beyond sharing opinions to reason in response to peers: they questioned assumptions, requested evidence, justified claims, reconsidered positions, and synthesized competing views. The clearest episodes occurred when one contribution prompted further reasoning by others.

S01: Some requirements are necessary because optional discussion usually benefits already confident students.

S15: I agree partly, but if it is required, I feel nervous because I need time to prepare my sentence.

S04: Would it be better if the requirement was one prepared point, not spontaneous speaking?

S03: So maybe the issue is not required versus optional, but what kind of participation counts.

This sequence moved the discussion beyond a binary choice between required and optional participation toward questions of fairness, preparation, and what counts as participation. Reflections supported this reading: active speakers reported revising their views after listening to less confident peers, whereas moderate participants described how questions and peer responses encouraged analytical thinking. Similar shifts appeared elsewhere in the discussion, including when participants distinguished silent listening from complete absence from the interaction.

### **3.3. Oral communication as fluency-in-interaction**

Students and facilitators did not define oral communication solely in terms of speaking quickly, accurately, or frequently. Rather, they understood it as the ability to sustain interaction despite hesitation. This could involve asking a brief question, responding to peers, clarifying an idea, repairing a misunderstanding, or summarizing a line of reasoning.

S11: Good questions are most important because they open the discussion.

S06: Responding critically is important because, without response, questions do not develop.

S15: For me, asking a short question is easier than speaking fluently.

S03: Maybe fluency is not speed, but the ability to continue interaction even with hesitation.

F03: Excellent distinction: fluency-in-interaction, not perfect grammar.

This finding links oral communication and critical thinking: the same moves that supported reasoning also provided opportunities for oral practice. Audio encouraged students to speak in response to others through a range of interactional moves, including questions, clarifications, summaries of emerging agreement, and self-repair. For quieter participants in particular, a brief but well-timed contribution represented communicative progress.

### **3.4. Facilitation as dialogic design**

Facilitation shaped the quality of discussion because audio alone did not ensure balanced participation or critical engagement. Reasoning-based prompts, sufficient wait time, turn-taking protocols, specific invitations to contribute, probing questions, and reflection activities supported productive discussion.

S07: The rubric could include one question, one response, and one reflection instead of focusing only on speaking time. F02: That is a design solution. What makes it fair? S07: It provides different ways to demonstrate thinking, rather than only measuring fluent speech.

Facilitation operated at two related levels. At the interactional level, facilitators managed turn-taking, invited less confident students to participate, and allowed wait time

so that faster speakers were less likely to dominate. At the cognitive level, they prompted participants to give reasons, make comparisons, and justify proposed solutions.

### 3.5. Discussion

Audio discussion should be evaluated by the quality of interaction as well as the amount of participation. In terms of Media Naturalness Theory, audio retains speech and synchronicity while reducing visual cues. This reduction can make interaction difficult, although it may also ease anxiety for students who prefer less visual exposure. Its value therefore depends in part on how discussion is designed and facilitated. Students may experience the same medium differently, so a single participation policy may not meet all their needs. Critical thinking was evident when students built on one another's contributions by asking questions, challenging ideas, and synthesizing perspectives.

### 3.6. Limitations and future research

This study involved 20 students and three facilitators in a single university context. The findings provide detailed insight into their experiences, but their transferability to other settings requires further examination. Interactional moves identified here may be more common in disciplines where speaking is central to learning. Future research could examine discussions across disciplines, proficiency levels, group sizes, and institutions.

## 4. Conclusion

This study examined audio-synchronous discussion as a setting for practicing critical thinking and oral communication. Its educational value lay in peer-responsive reasoning and interactional oral practice supported by purposeful design. Dialogue was limited when prompts elicited personal opinions without comparison, turn-taking was unclear, and assessment emphasized speaking time. Under these conditions, quieter students tended to withdraw as more confident speakers occupied more of the discussion. Facilitation shaped how the medium affected interactional practice. Audio did not itself create or prevent dialogue; its characteristics shaped the work facilitators needed to do to support productive interaction.

❖ **Conflict of Interest:** Author have no conflict of interest to declare.

## REFERENCES

- Alexander, R. (2020). *A dialogic teaching companion*. Routledge.  
<https://doi.org/10.4324/9781351040143>
- Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks*, 5(2), 1-17.  
<https://doi.org/10.24059/olj.v5i2.1875>
- Bakhtin, M. M. (1981). *The dialogic imagination: Four essays* (C. Emerson & M. Holquist, Trans.). University of Texas Press.

- Barak, A., & Suler, J. (2008). Reflections on the psychology and social science of cyberspace. In A. Barak (Ed.), *Psychological aspects of cyberspace: Theory, research, applications* (pp. 1-12). Cambridge University Press. <https://doi.org/10.1017/CBO9780511813740.002>
- Blau, I., & Barak, A. (2012). How do personality, synchronous media, and discussion topic affect participation? *Educational Technology & Society*, 15(2), 12–24.
- Blau, I., Weiser, O., & Eshet-Alkalai, Y. (2017). How do medium naturalness and personality traits shape academic achievement and perceived learning? An experimental study of face-to-face and synchronous e-learning. *Research in Learning Technology*, 25, 1974. <https://doi.org/10.25304/rlt.v25.1974>
- Cheng, X. (2000). Asian students' reticence revisited. *System*, 28(3), 435-446. [https://doi.org/10.1016/S0346-251X\(00\)00015-4](https://doi.org/10.1016/S0346-251X(00)00015-4)
- Daft, R. L., & Lengel, R. H. (1984). Information richness: A new approach to managerial behavior and organizational design. *Research in Organizational Behavior*, 6, 191-233.
- Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. *American Journal of Distance Education*, 15(1), 7-23. <https://doi.org/10.1080/08923640109527071>
- Hathorn, L. G., & Ingram, A. L. (2002). Cooperation and collaboration using computer-mediated communication. *Journal of Educational Computing Research*, 26(3), 325-347. <https://doi.org/10.2190/7MKH-QVVN-G4CQ-XRDU>
- Hrastinski, S. (2008). Asynchronous and synchronous e-learning. *Educause Quarterly*, 31(4), 51-55.
- Kock, N. (2005). Media richness or media naturalness? The evolution of our biological communication apparatus and its influence on our behavior toward e-communication tools. *IEEE Transactions on Professional Communication*, 48(2), 117-130. <https://doi.org/10.1109/TPC.2005.849649>
- Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal*, 92(1), 71-86. <https://doi.org/10.1111/j.1540-4781.2008.00687.x>
- Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, 26(13), 1753-1760. <https://doi.org/10.1177/1049732315617444>
- Mercer, N. (2004). Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, 1(2), 137–168. <https://doi.org/10.1558/japl.v1i2.137>
- Mercer, N., & Littleton, K. (2007). *Dialogue and the development of children's thinking: A sociocultural approach*. Routledge. <https://doi.org/10.4324/9780203946657>
- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 1-13. <https://doi.org/10.1177/1609406919899220>
- Satar, H. M., & Özdenler, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal*, 92(4), 595-613. <https://doi.org/10.1111/j.1540-4781.2008.00789.x>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

- Walther, J. B. (1996). Computer-mediated communication: Impersonal, interpersonal, and hyperpersonal interaction. *Communication Research*, 23(1), 3-43. <https://doi.org/10.1177/009365096023001001>
- Wegerif, R. (2007). *Dialogic education and technology: Expanding the space of learning*. Springer. <https://doi.org/10.1007/978-0-387-71142-3>
- Yamada, M. (2009). The role of social presence in learner-centered communicative language learning using synchronous computer-mediated communication: Experimental study. *Computers & Education*, 52(4), 820-833. <https://doi.org/10.1016/j.compedu.2008.12.007>
- Young, R. F. (2011). Interactional competence in language learning, teaching, and testing. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning*, 2, 426-443. Routledge.

---

**THẢO LUẬN TRỰC TUYẾN ĐỒNG BỘ BẰNG ÂM THANH NHƯ KHÔNG GIAN RÈN LUYỆN TƯ DUY PHẢN BIỆN VÀ GIAO TIẾP NÓI TRONG GIÁO DỤC ĐẠI HỌC**

**Lê Đình Tùng**

*Trường Đại học Sư phạm Thành phố Hồ Chí Minh, Việt Nam*

*Tác giả liên hệ: Lê Đình Tùng – Email: tungld@hcmue.edu.vn*

*Ngày nhận bài: 13-5-2026; Ngày nhận bài sửa: 10-8-2026; Ngày duyệt đăng: 28-8-2026*

**TÓM TẮT**

*Nghiên cứu định tính khám phá này xem xét cách thức các cuộc thảo luận trực tuyến đồng bộ bằng âm thanh có thể hoạt động như những không gian rèn luyện tư duy phản biện và giao tiếp nói. Các nghiên cứu trước đây đã xem xét sự tham gia và hợp tác trong học tập trực tuyến đồng bộ, nhưng ít chú ý đến chất lượng tương tác của thảo luận dựa trên giọng nói. Dữ liệu từ bảy phiên thảo luận với 20 sinh viên và ba người điều phối bao gồm các bản ghi thảo luận, 65 bài phản tư, 18 cuộc phỏng vấn và ghi chú quan sát, được phân tích theo phân tích chủ đề và phân tích nội dung định tính. Thảo luận bằng âm thanh tạo ra một không gian đối thoại giảm yếu tố thị giác nhưng dựa trên giọng nói: sinh viên đánh giá cao việc giảm mức độ hiện diện thị giác, song cũng trải nghiệm sự mơ hồ trong phân lượt, sự ngập ngừng và cảm giác bị phơi bày qua giọng nói. Tư duy phản biện được thể hiện qua việc đặt câu hỏi, biện minh, phản biện và tổng hợp; trong khi giao tiếp nói được hiểu phù hợp nhất như sự lưu loát trong tương tác. Thảo luận bằng âm thanh hỗ trợ các hoạt động này khi được thiết kế như một hoạt động đối thoại có cấu trúc.*

**Từ khóa:** thảo luận đồng bộ bằng âm thanh; tư duy phản biện; sự phạm đối thoại; giáo dục đại học; giao tiếp nói