Collaborative listening using multimedia through metacognitive instruction: a case study with less-skilled and more-skilled EFL learners

Esmat Shamsi[1*] and Hossein Bozorgian[1*]

*Correspondence:
shamsiesmat@gmail.com;
h.bozorgian@umz.ac.ir

[1] Department of English
Language and Literature,
University of Mazandaran,
Babolsar, Mazandaran, Iran

Abstract

Multimedia is vastly used as authentic and available input in second and foreign-language contexts. To deal with the difficulties and complexities of comprehending these authentic materials, collaborative learning through metacognitive instruction helps develop learners' listening comprehension. The present study has mainly explored the role of using multimedia input through metacognitive instruction of a group of 20 upper-intermediate EFL learners divided into five groups (each including two less-skilled and two more-skilled learners). They received eight weeks of metacognitive instruction in listening at an English language learning institute. Evidence gathered from observation, learners' logs, and semi-structured interviews shows that collaboration, interaction, and the use of metacognitive strategies help learners improve overall listening comprehension. Moreover, the learners had positive attitudes towards utilizing multimedia input, which provided them with some kinesics and prosodic hints to understand the listening content. The findings and pedagogical implications for language learners and teachers are discussed further.

Keywords: Collaboration, Less/more-skilled listeners, Listening comprehension, Metacognition, Multimedia

Introduction

Listening comprehension is essential in the second and foreign language (L2/FL) learning domain. Multiple factors such as class time limitation (Cross, 2014; Goh & Vandergrift, 2022), lack of authentic input, material, and context (Renandya & Hu, 2018), vocabulary knowledge deficiencies (Vandergrift & Baker, 2015, 2018; Wallace, 2020; Wang & Treffers-Daller, 2017), speed of speech (Siegel, 2013), and prosody and connected speech (Cross, 2009b) contribute to the EFL learners' listening deficiency. Therefore, to compensate for these problems, L2/FL learners need more opportunities to listen to authentic classroom input. One available listening input is multimedia (multiple content forms such as text, audio, and video into a single presentation), which improves L2/FL listening comprehension, motivation, and listening fluency (Cross, 2011a).

 Macaro et al. (2007) maintain that in listening learning, using some strategies such as cognitive (e.g., thinking aloud, guessing, and verifying guesses), socio-affective (e.g.,

awareness of motivational and affective states of the learners), and metacognitive strategies (e.g., planning, monitoring, problem-solving, and evaluation) can help learners to cope with difficulties and problems during listening learning. Research demonstrated that metacognitive strategies are practical and helpful in listening learning (Maftoon & Fakhri Alamdari, 2020; Zeng & Goh, 2018; Zhang et al., 2022). The focus of these strategies is on the listening process rather than the product, reaching a desirable amount of understanding collaboratively and having control over learning activities and affective factors (e.g., anxiety, self-confidence, and motivation) (Goh, 2018). These strategies may specifically help learners recall information more quickly when they listen and systematically plan for further learning development (Goh & Hu, 2013; Ngo, 2019).

Metacognitive instruction (MI) as a pedagogical procedure can help learners develop their knowledge about themselves, listening tasks, and appropriate strategies for fulfilling the tasks. In this approach, learners learn how to plan, monitor, and evaluate their listening comprehension and self-regulate their learning (Goh & Vandergrift, 2022). Some studies so far have investigated the effects of MI on EFL learners' listening development with different proficiency levels (Fung & Macaro, 2019; Kök, 2017). The results revealed that metacognitive instruction considerably improved learners' metacognitive awareness and listening comprehension. However, how learners practically deal with metacognitive strategies is still under-researched and needs careful and close examination. As such, this study aims to scrutinize how learners with different levels of proficiency who receive MI use and share metacognitive strategies and whether multimedia input facilitates their listening comprehension. Unlike previous studies that mainly used tests and questionnaires (e.g., Ahmadi Safa & Motaghi, 2021; Ghorbani Nejad & Farvardin, 2022; Mahdavi & Miri, 2017), the present study utilized three different data sources (observation, logs, and interviews) to provide additional and significant insights into the issue.

Review of literature

Multimedia in collaborative listening learning

The rationale behind applying multimedia as a learning tool is rooted in Mayer's (2009) theory of multimedia learning and Paivio's (1990) Dual Coding Theory. According to Clark and Mayer (2016), learners understand the materials better when they engage in active learning by connecting the words and pictorial materials. Learners need to organize these materials into their cognitive representation and combine them with their prior knowledge. This yields meaningful learning in which understanding is supported by learners' active participation in the learning process. Furthermore, Paivio (1990) states that the concepts that are simultaneously coded by verbal and visual channels may readily be retrieved and recalled.

Vygotsky's (1978, 1996) sociocultural theory (SCT) has been employed by researchers who study the effectiveness of multimedia players in language learning (Warschauer, 2005). They asserted that a physical tool (such as a multimedia player) and a symbolic tool (such as language) could be referred to as mediation that aids humans with their learning skills (Lantolf & Poehner, 2014). Through social interaction of this mediation, learners were able to reach a higher mental ability. This theory emphasizes the individual cognitive enhancement of meaning construction via social interaction with competent peers and teachers. By utilizing learning tools, language acquisition is promoted by the

Shamsi and Bozorgian *Asian. J. Second. Foreign. Lang. Educ.* (2024) 9:25

Page 3 of 18

aid of receiving efficient input, mental processing of linguistic input, and negotiating the meaning with others (Chapelle, 2009).

Cross (2018, p.1) defined collaborative listening as "the verbal interaction between pairs or small groups of language learners to facilitate and support one another's (a) comprehension of L2 listening texts and (b) L2 listening development". The purpose of collaborative listening is to raise learners' awareness, develop control over mental processes, find solutions for listening problems, and plan the listening process by discussing, selecting materials, and identifying text structure and genre. Tanewong (2019) believed that metacognitive instruction accented a collaborative approach to promote effective interaction among peers and teachers. It assisted less-skilled listeners in compensating for their deficiencies by obtaining enough information to share with their peers and understanding as much as more-skilled listeners did. Moreover, interaction inspired learners to guide, support, and correct each other.

In a systematic review research, Zhang and Zou (2021) found that multimedia input has been utilized to improve vocabulary learning, listening comprehension, and grammar learning, respectively. In the listening area, multimedia could add liveliness to activities to reach a better understanding of the content. These sources also aided the learners with constructing both auditory and visual representations of the content. Moreover, captions/subtitles helped learners lower their cognitive role in fulfilling listening tasks.

Regarding practical research, many inquiries have examined the effects of using multimedia input on EFL learners' listening comprehension (Azmee, 2022; Babaei & Izadpanah, 2019; Hsieh, 2019). As a case in point, Hesieh (2019) investigated the effects of watching videos with captions and the type of captions on Chinese EFL learners. The results showed that receiving two simultaneous modalities (i.e., text and audio) was more effective than receiving only one modality (either text or audio). A combination of images, captions, and audio provided learners with an easy and accessible multimodal material source through different channels. The results confirmed the positive effects of multimedia input on EFL learners' listening and vocabulary learning. However, it did not take into account learners' individual differences, learning styles, or potential effects of their language proficiency and background knowledge.

Similarly, Babaei and Izadpanah (2019) examined the impact of vocabulary, previewing comprehension questions, and multimedia annotations on EFL learners' listening comprehension. The result revealed that multimedia annotations had the most positive effect on learners' listening comprehension. The study investigated the issue quantitatively, and evidence was gathered through listening tests. Although this inquiry opened new horizons on employing visual aids to improve listening comprehension, learners' feelings, emotions, difficulties, or challenges were not examined thoroughly.

In EFL instruction, researchers have paid growing attention to the role of multimedia through MI in learning English listening interactively and collaboratively. For example, Jones (2006) scrutinized the effects of multimedia listening and collaborative learning on learners' listening comprehension. The results revealed that students who worked collaboratively with the presence of captions outperformed their counterparts who worked individually without the presence of captions. Interaction and negotiation allowed the learners to deepen their understanding, exchange missing words and information, answer questions, connect keywords, process the aural input, and enhance their

Shamsi and Bozorgian *Asian. J. Second. Foreign. Lang. Educ.* (2024) 9:25

Page 4 of 18

motivation. Her study did not examine the impact of using the integrative power of multimedia and collaborative listening learning to improve comprehensible output. Moreover, the effectiveness of interactive strategies in solving problems or sharing ideas was not investigated.

In a small-scale study, Cross (2010) investigated learners' construction and co-construction of metacognitive awareness through peer-to-peer dialogues. The results showed that encountering a listening comprehension problem could trigger discussion to find a strategic solution and self-regulatory strategic behaviors. In another study, Cross (2011b) investigated how interaction and contradiction between two members of a pair resulted in their metacognitive awareness and listening comprehension. Findings affirmed that the contradictions could support and constrain their learning by (1) encouraging them to redress the balance in their pair work and (2) revealing some sociocultural norms and aging differences that inhibit them from openly expressing their ideas. In this study, negotiation and discussion between a pair with different ages, background knowledge, and level of proficiency were observed and reported. The research aimed at scrutinizing the impacts of social and cultural oppositions on listening learners in Japan as a strictly culture-bounded society that avoids open disagreement and expression of feelings. Some kinds of imbalance in the amount of cooperation and partnership were obvious due to age differences and sociocultural norms in Japan. Therefore, according to the results, not all opportunities that metacognitive instruction provides for listening learning necessarily lead to better ways of listening learning. The contextual conditions and individual differences should be considered when prescribing such instructions.

The notion of dialogic interaction was further examined by Bozorgian and Fakhri Alamdari (2018) and Fakhri Alamdari and Bozorgian (2022). They conducted these two studies to compare metacognitive instruction (MI) and metacognitive instruction through dialogic interaction (MIDI) in a large-scale population. Findings demonstrated that learners in MI and MIDI groups outperformed their peers in the control group. Moreover, MIDI members had a better performance than the MI group due to the chances for more comprehension and facilities that visual material and interaction provided for them. These two studies reported positive evidence of listening comprehension development and clearly revealed the effectiveness of MI and MIDI on Iranian EFL learners with different genders, ages, and proficiency levels. However, many practical aspects of the metacognitive approach remained uncovered. These aspects could be revealed by conducting in-person interviews, directly and closely observing learners' learning behaviors, and collecting data from their personal journals, logs, or diaries

Less-skilled vs. more-skilled listeners and metacognitive instruction

L2 Listening ability is defined by learners' success in listening comprehension that is a function of their vocabulary size, topical and grammar knowledge, memory, and attentional control (Vandergrift, 2006; Wallace, 2020). Research has shown that learners with different levels of listening ability benefited from MI (Cross, 2011c; Fung & Macaro, 2019). Research also revealed that by applying MI, both less- and more-skilled listeners enhanced their listening comprehension, and in many cases, less-skilled ones showed a higher degree of improvement (Roussel et al., 2017). For instance, Vandergrift and Tafaghodtari (2010)

found that less-skilled listeners performed better than more-skilled ones since they did not transfer natural approaches from their first language to their second language. Teachers and peers guided less-skilled listeners to regulate their listening learning better. On the other hand, more-skilled listeners did not translate the text mentally and could focus on self-monitoring and problem-solving strategies. They had a broader lexical knowledge and a greater ability to employ their topical and background knowledge whenever needed. This experimental study was conducted on French as a second language learners at the beginner level in which they were not presented with a transcription of the text for the final verification stage. Therefore, the learners could not compare the articulated speech with its written form. Different and more robust results would be obtained if the researchers have provided such transcriptions for the learners.

In Cross's (2011c) study, less-skilled listeners arranged both cognitive and metacognitive strategies simultaneously to reach a stronger memory of information from listening to texts. However, more-skilled listeners had a better realization and could appropriately and simultaneously apply many strategies. Moreover, in another study, Tanewong (2019) claimed that in problem-solving strategies such as guessing unknown words, less-skilled listeners performed better than more-skilled. Their improvement in some abilities, namely predicting, lexical segmentation, and avoiding mental translation, was evident. Furthermore, they always tried to get prepared for doing the listening task and gather more information to share with more-skilled listeners. Gradually, their awareness of their weaknesses increased due to interacting with more-skilled listeners. The researcher used both audio and authentic video texts as the input of the study. The learners were interviewed and their attitudes towards the impact of MI on their person, task, and strategy knowledge were elicited. The focus was on the affordances of MI and its stages, as well as learners' prioritized, favorite, or unpreferred strategies. In this study, the advantages or disadvantages of utilizing audio vs. video texts for less- and more-skilled learners were not considered.

Additionally, Read and Barcena Madera (2016) maintained that less-skilled listeners had more control over the process of fulfilling listening tasks, which was the result of structuring and scaffolding. On the other hand, more-skilled listeners utilized metacognitive strategies automatically and unconsciously because they had internalized these strategies before. The research was conducted as mobile-assisted distance learning without the presence of a teacher. The main purpose of the study was to investigate MI's effects on learners' self-regulation, scaffolding, self-awareness, and self-evaluation skills resulting from applying a specified social app for listening learning. Thus, noticing the nature of distance-learning, the learners did not enjoy collaboration, interaction, and negotiation, nor did they receive teachers' explicit instruction and immediate feedback.

Noticing existing literature and with respect to the purpose of this study, the following research question was addressed: 'What are the roles of collaborative learning and using multimedia through metacognitive instruction in less- and more-skilled female EFL learners' listening comprehension?'.

Method

Design
As scrutinizing EFL learners' behaviors and reporting their attitudes, concerns, and feelings require close observation and self-reflection data sources, a qualitative approach was

deemed appropriate for this study (Dörnyei, 2007). This approach could address the topic, meet the need for investigating learners' behaviors while making decisions on selecting and using a specific strategy, and affordances that multimedia listening through MI may provide for them.

Participants

With a purposeful sampling method, 20 participants out of 63 volunteers were selected. These participants were non-English-major university students, relatively homogenous concerning their willingness to take part in the study, their age (18–20), gender (all female), level of language proficiency (upper intermediate based on Oxford Placement Test (OPT) results conducted before training sessions), and their common goal and interest to enhance listening skills. All the participants declared their agreement in consent forms.

   The following instruments and materials were utilized to collect data and respond to the research question:

Observation field notes

As the main data source for documenting participants' activities and behaviors, the first researcher, as a participant in this study, briefed learners on placing cameras in the classroom and recorded all sessions in a semi-structured form (Creswell & Creswell, 2017). After each session, the researchers labeled each recording file with the time, date, and setting of sessions and uploaded them to their laptops for further access and data analysis. They watched the recordings twice, transcribed them verbatim, and prepared the data for coding and content analysis.

Learners' log

In order to investigate the learners' difficulties, internal monologues, and preferred strategies during training sessions, they were asked to write down their reflections in personal journals (Goh, 2010) at the end of each session (see Appendix A). They were required to record all details, even those that did not seem significant to them, and there was no limitation to respondents' amount of statements. The last 20 min were allocated to filling learners' logs in all sessions. They were asked to write their journals immediately after each session ended since their memories were fresh and they could remember more details. The teacher asked them to write about their feelings, difficulties, thoughts, and further plans and suggestions. Totally, 20*8 = 160 logs were collected.

Semi-structured interviews

A semi-structured interview was conducted in person one week after the sessions ended to verify log data as well as to collect more evidence for learners' attitudes and feelings towards the program that may not be gathered by observation. Interview questions were adapted from Cross (2009a) and contained six open-ended questions (see Appendix B).

Material

The material used in this study were entertaining and informative short videos (5–10 min) about culture, lifestyle, and social issues (Cross, 2011a) selected from TED Talks and YouTube with exciting topics related to the learners' course books and

comprehensible for language learners with upper-intermediate level. The teacher indirectly monitored the process of selecting suitable videos by providing some hints about the length, content, accent, and speed of speech. She asked the learners to choose short videos with a topic relevant to the lessons of their course book in an American accent with a moderate speed of speech.

Procedure

The study lasted eight weeks, and the class met once a week for ninety minutes. In the first session, the teacher (the first researcher of this study) divided 20 learners into five foursome groups, which contained two less-skilled and two more-skilled learners. Their listening proficiency level was determined based on their listening scores by administrating the listening section of the IELTS Practice Test. The learners scoring above the mean were classified as more-skilled, and those scoring below the mean were classified as less-skilled listeners. Then, the teacher distributed Guide Sheets (Goh & Vandergrift, 2022) for listening among them. All sessions were held according to Goh and Vandergrift's (2022) outline for a video listening lesson supported by metacognitive activities.

Throughout the pre-listening discussion, by reading the video title, looking at a picture related to the content, and noticing the teacher's explanations, the learners predicted what they would hear and see in the video and orally shared what they knew about that topic with their groupmates. Then, in their guide sheets, they wrote down five main ideas they thought would be mentioned in the video. After that, they discussed their predictions with their groupmates and wrote down more ideas that their groupmates included in their list of predictions, and they considered logical possibilities.

Groups watched the video without the caption, listened carefully, and took notes of content words. After first listening, they put check marks beside the ideas that were mentioned in the video and they had predicted correctly and wrote down other ideas that they had not predicted but were mentioned in the video. Using keywords and phrases, they together could produce a short text close to the original videotext. They shared difficulties and challenges that they experienced trying to understand the content and planned how they would listen out for the main ideas and details throughout the second listening.

After verifying their predictions and discussing listening results with groupmates, the learners watched the video without caption for the second time and wrote down the remaining keywords and phrases as well as unheard and misheard parts. Then, they compared their notes with each other and added more words and phrases to their own script. Further, they checked the results to resolve any discrepancies in comprehension among group members, made a list of words that they did not understand at all, and planned to figure them out throughout the third listening. In this phase, they identified exactly how many words should be focused on and where are their approximate location in the video. They also determined whether they were new and unfamiliar words or a combination of two familiar ones that were misheard in the process of mixing or removing final sounds. Finally, they added further points and important details that they did not recognize at all through the first listening stage.

The learners watched the video for the third time with the caption 'on' to verify comprehension after a group discussion of the content of the text or a reading of the

Shamsi and Bozorgian *Asian. J. Second. Foreign. Lang. Educ.* (2024) 9:25

Page 8 of 18

text transcript. They checked their notes, corrected wrong words and phrases, and added new ones. They identified words in the captions they did not understand, copied them, and listened to their pronunciation in the connected speech. Then, they put their notes aside and watched the video without the captions again. They were not allowed to add anything to their notes anymore.

Finally, the learners read their script and decided which was closest in meaning and detail to the original text. Afterward, they shared their notes in the group, explained their feelings when they saw the caption, and discussed whether captions helped or limited their listening. For the evaluation stage, they discussed what difficulties/successes they had with trying to understand the text, how successful the strategies they employed were, whether any other strategy could be tried, and what they planned to do for the next session. These activities are summarized in Table 1.

Data analysis

Thematic analysis was done to respond to the research question. First, all interviews were transcribed verbatim, and responses were reassembled and disassembled multiple times (Yin, 2011). Logs and field notes were also analyzed at the end of each session. The second researcher and an external coder, a Ph.D. holder of TEFL (teaching English as a foreign language) familiar with qualitative data analysis, separately and autonomously coded data, and their codings were compared. They read field notes, logs, and interview transcription several times to find repetitive and significant patterns. Then, initial codes were written down based on selecting and utilizing strategies by less- and more-skilled learners in different stages. The coders categorized the codes after several recoding and refining. Regular meetings were held, and steps taken across the data analysis process were shared. To assess the credibility of findings through member checking, the researchers sent the final report containing all

Table 1 Multimedia listening instruction stages based on Metacognitive processes (Goh & Vandergrift, 2022)

Stages	Activities
Stage one: Planning/prediction	The learners read the title and saw a picture/preview related to the video, and discussed the probable information that they would receive and the words and phrases they would hear. Then, they wrote down their ideas, negotiated them, and added more ideas that their groupmate shared
Stage two: First verification	They watched the video without the caption, listened carefully, confirmed correct predictions, and wrote down unknown words and phrases. Using keywords, they could produce a short text close to the original text, identify unheard parts, and plan for the second verification stage
Stage three: Second verification	In this stage, they watched the video again without the caption, tried to complete their texts, compared their notes with each other, and fulfilled incomplete sections. They also prepared a list of remaining unknown words to find out further
Stage four: Final verification	They watched the video with turned-on captions, corrected the wrong words, and wrote down the rest of the unknown words; after that, they put aside their notes and watched the video again
Stage five: Reflection and evaluation	They read their texts to each other and shared their feelings, difficulties, and attitudes toward the effectiveness of strategies that they used and would use in upcoming sessions

descriptions and categories to the participants and asked them to confirm information accuracy and add any missed details that they thought were necessary.

Moreover, the researchers asked an auditor to read the final report and examine it regarding the relationship between collected data and the research question, the accuracy of transcription, and the stages of data analysis from raw data to interpretation (Creswell & Creswell, 2017). These two strategies did not change categories or final report extremely and only yielded a minor modification to descriptions. By assessing the reliability of findings through a cross-checking process, an acceptable inter-coder agreement (reliability = 0.82) was obtained (Gibbs, 2007).

Findings

We categorized the obtained patterns into four broad themes: planning in collaborative multimedia listening, monitoring in collaborative multimedia listening, problem-solving in collaborative multimedia listening, and evaluation in collaborative multimedia listening.

Planning in collaborative multimedia listening

The goal of this phase was to preview task demands and prepare for the first verification listening. The learners were supposed to be prepared and think deeply about the goal of listening tasks and what they needed to take from the video for producing further output. In this phase, they brought together all they knew about the topic. More skilled learners warned others about the initial words of videotext, which contain important information and are usually missed by many listeners.

In this phase, groups decided on what kind of text they would listen to, the probable information they would receive, and particular names, dates, or specific numbers they would hear. The video title, the teacher's explanations, the introduction, and the images that were shown before playing the video provided important hints for them in this stage. Some of them could write five predictions they were supposed to, and others could not; however, they all attempted to complete their notes by and large and wrote down as much as possible. Gradually, the number of ideas increased and became more relevant to the topic.

Regarding this issue, the following excerpt was selected from one of the less-skilled participants' logs:

Initially, I did not know what to write and whether my predictions were correct and related to the topic. I could hardly write one or two ideas. When I shared my thoughts with my groupmates, I realized that in this phase, we should freely write whatever comes to our minds. The more predictions and anticipation we make, the more precise and vivid scenarios we create. We could further modify or share our scenarios with our groupmates to refine them (Lily, 19, Session 1).

Utilizing cultural, scientific, social, or educational hints of topics to anticipate the content of the video and brainstorming among group members is evident in a more-skilled learner's log in this way:

The topic of today's vodcast was "How is your city tackling the climate crisis?" We quickly read the explanation and tried to write five ideas. Considering its brief

introduction, we could guess that this video would be about the role of humans in damaging the planet, environmental pollution, sustainable energies, and the role of governments and environmentalists in protecting the planet against the harm of climate change. After watching the video, we realized that most of our predictions were correct, which was a great success for us (Tara, 20, Session (2).

Monitoring in collaborative multimedia listening

The learners checked their hypotheses and predicted ideas throughout the three verification stages. They put a checkmark next to correct ideas and removed incorrect and irrelevant ones. They wrote down all key terms, phrasal verbs, idioms, and other words they did not know their meaning, had not heard before, or misheard. Many of these unknown items were detected by intergroup discussion. More-skilled members helped others find and guess the meaning and replace misheard words with correct ones. They agreed on items that remained unknown, located them, and determined to listen more carefully in the second and third verification stages.

The most problematic mishearing words and phrases for learners were those with some kinds of elision, assimilation, and linking of two consonants. During close observation, it was perceived that in some cases, learners knew the written form of a word in a sentence; however, they had severe problems with verbal forms. They could not comprehend the speech of native speakers but could completely recognize subtitles in written form. This can be attributed to a lack of familiarity with phonemic aspects of language. A participant commented: "When we saw subtitles, we realized that we knew most of the words but could not realize their verbal form, which is a problem that most of us encountered" (Betti, 18, Session 2).

All in all, this phase focused on checking the accuracy of predictions and watching performance consciously. Furthermore, attention was paid to raising awareness about the listening process and the probable breakdown that occurred during this process. Priority in this stage was given to active involvement, using strategies consciously, and evaluating the amount of success in predicting and achieving the determined goals of the group.

Problem-solving in collaborative multimedia listening

Since the nature of listening is fleeting and listeners usually hear only once, information can be lost for many reasons: unfamiliar accent (silent 'r' in a British accent), speed of speech, distraction, and mind wandering. Observations confirmed that utilizing some metacognitive strategies helped cope with these issues. The aim here was to present a comprehensive text, and the learners were supposed to pay attention to the information they received in detail. Every member tried to bring as much as possible, and at last, they arranged these segments next to gather like the pieces of a jigsaw puzzle to form a complete text. If they failed to find all the pieces, the teacher introduced an appropriate strategy they had not employed before. For example, a less-skilled learner referred to this strategy proposed by the teacher about coping with high speed of speech: "We would pause the videos at the end of every chunk to reflect on the text and repeat it in our mind. We would replay it to find out the message completely" (Hannah, 19, Session 5).

In their interviews, learners asserted that some hints in videos aided them in better comprehension. For instance, to guess the name of an unknown object, the learners could indirectly guess the name of that object when the speaker said its name and pointed to it. Furthermore, speakers' gestures and body language were also helpful to understand their intentions. Some other strategies, such as exploiting known words to guess the meaning of unknown words, were also noticeable. One of the less-skilled learners stated:

> In some cases, we could guess the meaning of unknown words. Some words are very similar to words that we regularly use. For example, I always used the word 'option' and did not know the meaning of 'opt,' which is a verb. I guessed there should be a relationship between these two, and I was right (Tina, 19, Session 5).

When less-skilled learners encountered comprehension problems, more skilled could propose efficient strategies since they had often experienced them. The teacher encouraged more-skilled learners to explain suitable strategies to less-skilled ones so that their own knowledge would be solidified and deepened. Gradually, there was a sense of competition and rivalry among the groups, and all the members engaged in activity to achieve their common goal and receive teachers' positive feedback. All the learners were close friends, relatively the same age, and already knew each other. Thus, there was no excessive respect or shyness among them resulting from cultural or social differences. They easily challenged each other and disagreed with inefficient strategies.

An interesting strategy that was mainly used by more-skilled participants and transferred to less-skilled ones was ignoring unknown words. It was evident that when less-skilled learners encountered an unknown word, they felt anxious and highly focused on it and subsequently lost further sections. Regarding this issue, this excerpt was selected from a less-skilled learner's log:

> I realized that I used to overemphasize unknown words and phrases. Today, our teacher told us that even a native speaker of English may not know the meaning of some terms or may lose track of their listening. The first strategy is ignoring unknown words. Furthermore, their meaning may be revealed in the following sentences, or you may guess the meaning by paying attention to adjacent words or sentences. In a text consisting of 300 words, you may not know the meaning of two or three words, and it is ignorable. Do not look at a text word by word; consider it a whole text trying to convey a message. Do not translate it in your mind because it would slow your listening down and cause distraction (Sarah, 19, Session 6).

Some learners referred to the speed of speech. They paused or replayed the videos at a slower speed to adjust themselves to the high speed of native speakers' speech. Learners gradually increased the pace of the video to an average or even high speed. Moreover, paralinguistic multimedia aspects in listening comprehension, such as kinesics and prosody, provided the learners with helpful cues to understand the text and grasp the meaning. It was aptly put in a learner's interview:

> Facial expressions and body movements contain significant messages for understanding unknown words. They may carry either positive or negative feelings and can be replaced with verbal statements. The tone and stress of the spokesperson's

Shamsi and Bozorgian *Asian. J. Second. Foreign. Lang. Educ.* (2024) 9:25

Page 12 of 18

speech are also beneficial for understanding the content. Some noteworthy informa-
tion and prominent parts of speech are articulated with a higher pitch and more
stress. Yes, it (visual information) helped me understand better (Ella, 20, Session 5).

Evaluation in collaborative multimedia listening

The last phase of the instruction was to evaluate learners' performance and decide on fulfilling further listening tasks. The learners decided what strategies were helpful and what others were not. They discussed why some strategies did not meet their purposes and the main reasons for this failure; whether they were satisfied with their performance; what they could do to fulfill their tasks and achieve their in-common goals; and what emotions and feelings (positive or negative) they experienced during their collaborative listening learning. A learner evaluates her performance as follows:

I think others were better than me. They could guess more ideas in a shorter time.
Maryam (a more-skilled member) was more active than the other students. She
told us how to focus on unheard or misheard parts in the second and third listen-
ing rounds. She inspired us to be more active and write our ideas even if we thought
they were wrong. She asked us to divide the text into chunks that comprise similar
ideas and information. It helped us to adjust ourselves to lengthy and complex texts
(Rosie, 18, Session 8)

Two other learners referred to their negative emotions regarding failure in accommodating to the speed of speech and guessing the meanings of unknown words in this way: "Losing the track or mind wandering due to intrusive thoughts was disappointing for me, but the point is you must immediately refocus on the video and be aware of your listening process. Try harder and don't give up" (Anna, 19, Session 7). The other one stated: "Sometimes there is no possible way to guess the meaning of a word, and it really makes me dispirited because grasping the message of a video closely depends on its keywords; consequently, many parts are lost and remain unknown" (Mina, 20, Session 6).

Discussion

The following research question directed and led this study: "What are the roles of collaborative learning and using multimedia through metacognitive instruction in less- and more-skilled female EFL learners' listening comprehension?" Evidence from direct observation, learners' logs, and interviews illustrated the importance of interaction, collaboration, and using and sharing metacognitive strategies in fulfilling listening tasks. Strategies such as predicting ideas, problem-solving, evaluating, and monitoring activities were used more than other strategies. In addition, utilizing kinesics and phonological cues such as speakers' body gestures and visual hints embedded in multimedia input could help learners reach a better comprehension. In this study, collaboration among more-skilled and less-skilled learners is also found to help construct new ways of thinking and meaning making. On the one hand, a sense of competition among members of one group as well as between one group and others, worked as an impetus for more effort and activity. On the other hand, there was a sense of sharing and interaction among more-skilled learners in order to receive their teacher's approval and consolidate their own knowledge. These advantages led the

more-skilled learners to aid less-skilled ones in elevating their level of listening proficiency and collaborating in group work. Learners in this study gradually enhanced their self-efficacy and self-confidence and managed their approach to listening with the help of metacognitive strategies. Almost all of the participants in our study stated that receiving audio and visual input simultaneously eased their listening comprehension and helped them with more effective listening. For these learners, speakers' facial expressions, hand gestures, and body language were helpful for better understanding the videos. These features could convey important messages about content and aid listeners with guessing the meaning of unknown words.

Regarding the impact of using multimedia on listening comprehension, the finding of the present study was consistent with Azmee (2022), Babaei and Izadpanah (2019), and Hsieh (2019) in which the positive effect of using multimodal input in a multimedia listening environment was confirmed. The studies mentioned above attributed positive results to the accessibility, comprehensibility, and multimodality of multimedia resources. Some options, such as pausing, rewinding, or slowing down the videos, were also mentioned as listening comprehension facilitators. When a failure occurred in either verbal or visual channels, receiving information from two sensory channels simultaneously helped the learners recover it successfully. Therefore, learners could utilize two kinds of information in working memory without any competition between them.

The finding also aligned with Cross's (2011a) notions that multimedia supported learners' listening comprehension. He referred to captions and subtitles as other helpful factors; however, in some cases, the multimodality of contiguous information provided by computer-generated animations was considered detrimental to learners' understanding. Additionally, our finding was in line with Jones (2006), who claimed that sharing perceived information and keywords in collaborative listening improved learners' performance and enhanced their motivation in examining listening tasks more closely and connecting words and relevant ideas. They could compare their understanding of the material with their peers, aid their partners with supplying missing and unknown words, and answer the comprehension questions. Furthermore, the results of our study confirmed Jones's claim that multiple retrieval routes were helpful in recalling vocabulary and propositions of listening passages and deepened understanding through negotiation and interaction.

The effectiveness of interaction in the co-construction of meaning (brainstorming, sharing ideas, and negotiating) obtained in this study was congruent with Cross (2010), demonstrating that interaction raises discussion, which in turn leads to discovering a proper solution for listening comprehension problems and increasing metacognitive awareness. However, Cross' study was not a classroom-based one. He grouped learners in dyads; each pair completed five lessons separately at different times, and they did not receive multimedia input from the researcher. Striving to detect an effective and suitable solution was observed among participants of the present study in different stages, specifically in problem-solving. When groups encountered a common listening issue that needed the contribution of all members, it would turn into a common goal and a stimulus to an active interaction and find a reasonable solution; subsequently, they used appropriate strategies, and listening development took place in different levels for almost all learners in the group.

The finding of this study also affirms Cross (2011b), who reported that some factors such as interaction, common goal, motivation, and collaboration shaped the participants' listening development. The difference was that in Cross' research, participants' social, cultural, and historical conditions were different, and sometimes, these factors caused some inhibiting contradictions in learners' listening development. However, in the present study, selected participants were homogeneous in terms of their gender, age, and cultural conditions so that they could share ideas and strategies freely and easily without any restriction caused by respect or shame. The purpose here was to observe collaborative behaviors that led to successful or unsuccessful listening comprehension.

Like this study, collaboration and interaction among learners through MI and receiving multimedia input were declared as major determinants of successful listening comprehension in two studies conducted by Bozorgian and Fakhri Alamdari (2018), and Fakhri Alamdari and Bozorgian (2022). Both of them demonstrated that interaction through MI assisted EFL learners in improving listening performance and metacognitive awareness. This improvement was assigned to the systematic and collective nature of MI that actively involved learners in listening learning and increased their motivation. Furthermore, since multimedia input was easily available, comprehensible, and ubiquitous, teachers could integrate it with listening pedagogy and enhance learners' motivation and willingness to listen.

With regard to less- and more-skilled learners' use of metacognitive strategies through MI, the result of this study can empirically support the different use of these strategies by learners with different levels of listening ability. Less-skilled listeners mainly relied on their prior knowledge about the topic, cultural information, and guessing the meaning of unknown words with the help of known words. More-skilled listeners, instead, were able to direct their attention and avoid mind wandering and mental translation. They seemed less anxious, more motivated and could use metacognitive strategies unconsciously. These findings are aligned with relevant previous literature (Cross, 2011c; Read & Barcena, 2016; Tanewong, 2019; Vandergrift & Tafaghodtari, 2010) in which the effectiveness of MI on learners with different levels of listening ability was confirmed. The mentioned studies reported that less-skilled listeners may also benefit from MI even more than more-skilled ones, and MI was applicable to learners with different levels of proficiency. In our study, both less- and more-skilled learners used metacognitive strategies throughout the listening process; however, the kinds of strategies and frequency of using them differed. All in all, when learners with the same age, gender, and level of proficiency are grouped in small groups and share metacognitive strategies, they may construct and co-construct new knowledge, compensate for their listening deficiency, and overcome emotional barriers such as listening anxiety, disappointment, and lack of self-confidence. Moreover, multimedia can act as a supportive factor and compensate for poor vocabulary knowledge, mishearing, and mind-wandering.

Conclusion

This study attempted to explore the role of collaborative multimedia listening in learners' listening comprehension through metacognitive instruction. Findings demonstrated promising and advantageous evidence for the effectiveness of this approach, which can be helpful for teachers and learners in EFL classrooms. The key contributing factors

were negotiation, collaboration, and close discussion, accompanied by the affordances that multimedia inputs provided for EFL learners. The implication of this study is essential since it contributes to our understanding of new listening teaching approaches and using varied and more attractive material to help learners enhance their motivation, collaboration, and willingness to learn. For language teachers, the metacognitive instruction approach is assumed beneficial since it is likely to develop listening skill by taking more control over the learning process and meeting its challenges. Grouping learners with different listening ability levels in small groups might make them eager to elevate their listening competence, help other groupmates take more practical strategies, and compensate for their listening deficiencies.

However, the findings of this study should be interpreted cautiously and in light of some limitations. First, the number of participants was limited; therefore, the findings are not easily generalizable to a larger population. Further longitudinal research is needed to replicate the process with more participants. Second, the study was conducted with female upper-intermediate learners. Future studies may consider learners with different levels of proficiency and gender. Finally, adopting a mixed-methods research design may shed more light on this path and increase the degree of certainty about the inquiry's findings.

Appendix A
Variations in Focus and Prompts of Learners Listening Logs (Adapted from Goh, 2010).

Please answer the following question in short statements

1. What was the topic of the listening lesson?

2. How much of it did you understand at the first time listening?

3. What did you do to understand as much of it as possible?

4. How did you feel about what you did?

5. Are you pleased with the result?

6. What did you plan to do to fulfill your listening task in this session?

Complete the table regarding the degree of listening skill using the symbols

√ = Yes, I feel I have learned the skill

x = No, I have not learned the skill very well

? = Not sure; I am not sure I have learned the skill

Skills I learned to improve my performance

Asking for repetition and explanation/clarification

Listening for descriptions of past events

Recognizing keywords

Distinguishing different styles

Listening for positive and negative opinions

Inferring speakers' attitude from their tone of voice and gestures

1. Describe your previous experiences in learning to listen

2. Do you think the strategies you've applied so far are helpful? Compare them with your previous experiences

You will be given 20 min at the end of the class to write about:

1. The listening comprehension activities done in each class

2. Your feelings towards the class today and what makes you feel that way

3. The problems and difficulties encountered

4. Your plans to overcome similar problems in the future

Appendix B

Interview Questions (Adapted from Cross, 2009a).

1. How did you try to understand unknown words in the videos?
2. Are visual cues helpful in understanding the videos? If yes, how?
3. What difficulties did you encounter in comprehending videos?
4. Are using metacognitive strategies helpful in comprehending the videos? If yes, how?
5. Are you content with your overall performance? Why/why not?
6. What are your suggestions for similar studies that will be conducted in the future?

Abbreviations
EFL   English as a foreign language
MI   Metacognitive instruction
MIDI   Metacognitive instruction through dialogic interaction
IELTS   International English language testing system
TEFL   Teaching English as a foreign language

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
All the participants consented to take part in this study.

Competing interests
We have no known conflict of interest to disclose.

References
Ahmadi Safa, M., & Motaghi, F. (2021). Cognitive vs metacognitive scaffolding strategies and EFL learners' listening comprehension development. *Language Teaching Research*. https://doi.org/10.1177/13621688211021821

Azmee, A. (2022). The impact of the podcast (BBC Sounds) to raise metacognitive awareness in developing L2 listening comprehension among Bangladeshi undergraduate learners. *Computer Assisted Language Learning Electronic Journal (CALL-EJ), 23*(4), 103–119.

Babaei, S., & Izadpanah, S. (2019). Comparing the effects of different advance organizers on EFL learners' listening comprehension: Key vocabularies, previewing comprehension questions and multimedia annotations. *Cogent Education, 6*(1), 1–25. https://doi.org/10.1080/2331186X.2019.1705666

Bozorgian, H., & Fakhri Alamdari, E. (2018). Multimedia listening comprehension: Metacognitive instruction or metacognitive instruction through dialogic interaction. *ReCALL, 30*(1),131–152. https://doi.org/10.1017/S0958344016000240

Chapelle, C. A. (2009). The relationship between second language acquisition theory and computer-assisted language learning. *The Modern Language Journal, 93*, 741–753. https://doi.org/10.1111/j.1540-4781.2009.00970.x

Clark, C. R., & Mayer, R. E. (2016). *e-learning and the science of instruction*. John Wiley & Sons Inc.

Creswell, J., & Creswell, D. (2017). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Beverley Hill: Sage Publication.

Cross, J. (2009a). Effects of listening strategy instruction on news videotext comprehension. *Language Teaching Research, 13*(2), 151–176. https://doi.org/10.1177/1362168809103446

Cross, J. (2009b). Diagnosing the process, text, and instruction problems responsible for L2 listeners' decoding errors. *The Asian EFL Journal, 11*(2), 31–53.

Shamsi and Bozorgian *Asian. J. Second. Foreign. Lang. Educ.* (2024) 9:25

Page 17 of 18

Cross, J. (2010). Raising L2 listeners' metacognitive awareness: A sociocultural theory perspective. *Language Awareness, 19*(4), 281–297. https://doi.org/10.1080/09658416.2010.519033

Cross, J. (2011a). Comprehending news videotexts: The influence of the visual content. *Language Learning and Technology, 15*(2), 44–68.

Cross, J. (2011b). Social-cultural-historical contradictions in an L2 listening lesson: A joint activity system analysis. *Language Learning, 61*(3), 820–867. https://doi.org/10.1111/j.1467-9922.2011.00658.x

Cross, J. (2011c). Metacognitive instruction for helping less-skilled listeners. *ELT, 65*(4), 408–416. https://doi.org/10.1093/elt/ccq073

Cross, J. (2014). Promoting autonomous listening to podcasts: A case study. *Language Teaching Research, 18*(1), 8–32. https://doi.org/10.1177/1362168813505394

Cross, J. (2018). Collaborative listening: Teaching listening pedagogical perspectives on listening. *The TESOL Encyclopedia of English Language Teaching (First Edition)*. https://doi.org/10.1002/9781118784235.eelt0598

Fakhri Alamdari, E., & Bozorgian, H. (2022). Gender, metacognitive intervention, and dialogic interaction: EFL multimedia listening. *System, 104*, 1–13. https://doi.org/10.1016/j.system.2021.102709

Fung, D., & Macaro, E. (2019). Exploring the relationship between linguistic knowledge and strategy use in listening comprehension. *Language Teaching Research*. https://doi.org/10.1177/1362168819868879

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press

Ghorbani Nejad, S., & Farvardin, M. T. (2022). Roles of general language proficiency, aural vocabulary knowledge, and metacognitive awareness in L2learners' listening comprehension. *International Journal of Listening, 36*(3), 187–206. https://doi.org/10.1080/10904018.2019.1572510

Gibbs, G. R. (2007). Analyzing qualitative data. In U. Flick (Ed.), *The Sage qualitative research kit.* Sage publications.

Goh, C. (2010). Listening as process: Learning activities for self-appraisal and self-regulation. In N. Harwood (Ed.), *English language teaching materials: Theory and practice* (pp. 179–206). Cambridge University Press.

Goh, C. (2018). Metacognition in second language listening. *The TESOL Encyclopedia of English Language Teaching (first Edition)*. https://doi.org/10.1002/9781118784235.eelt0572

Goh, C., & Hu, G. (2013). Exploring the relationship between metacognitive awareness and listening performance with questionnaire data. *Language Awareness, 23*(3), 255–274. https://doi.org/10.1080/09658416.2013.769558

Goh, C., & Vandergrift, L. (2022). *Teaching and learning second language listening metacognition in action* (2nd ed.). Tylor & Francis Group: Routledge.

Hsieh, Y. (2019). Effects of video captioning on EFL vocabulary learning and listening comprehension. *Computer Assisted Language Learning, 33*(5–6), 567–589. https://doi.org/10.1080/09588221.2019.1577898

Jones, L. C. (2006). Effects of collaboration and multimedia annotations on vocabulary learning and listening comprehension. *CALICO Journal, 24*(1), 33–58.

Kök, I. (2017). Relationship between listening comprehension strategy use and listening comprehension proficiency. *International Journal of Listening, 32*(3), 163–179. https://doi.org/10.1080/10904018.2016.1276457

Lantolf, P. L., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. London: Routledge.

Macaro, E., Graham, S., & Vanderplank, R. (2007). A review of listening strategies: Focus on sources of knowledge and on success. In E. Macaro & A. Cohen (Eds.), *Language learner strategies: 30 years of research and practice* (pp. 165–185). Oxford UniversityPress.

Maftoon, P., & Fakhri Alamdari, E. (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

Mahdavi, N., & Miri, M. (2017). Co-shaping metacognitive awareness and developing listening comprehension through process-based instruction. *International Journal of Listening*. https://doi.org/10.1080/10904018.2016.1260454

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge: Cambridge University Press.

Ngo,. (2019). Understanding the impact of listening strategy instruction on listening strategy use from a sociocultural perspective. *System, 81*, 63–77. https://doi.org/10.1016/j.system.2019.01.002

Paivio, A. (1990). *Mental representations: A dual coding approach*. Oxford: Oxford University Press.

Read, T. M., & Bárcena Madera, M. E. (2016). Metacognition as scaffolding for the development of listening comprehension in a social MALL app RIED. *Revista Iberoamericana De Educación a Distancia, 19*(1), 103–120.

Renandya, W. A., & Hu, G. (2018). L2 listening in China: An examination of current practice. In A. Burns & J. Siegel (Eds.), *International Perspectives on Teaching the Four Skills in ELT* (pp. 37–50). Palgrave Macmillan.

Roussel, S., Gruson, B., & Galan, J. (2017). What types of training improve learners' performances in second language listening comprehension? *International Journal of Listening, 33*(1), 39–52. https://doi.org/10.1080/10904018.2017.1331133

Siegel, J. (2013). Second language learners' perceptions of listening strategy instruction. *Innovation in Language Learning and Teaching, 7*(1), 1–18. https://doi.org/10.1080/17501229.2011.653110

Tanewong, S. (2019). Metacognitive pedagogical sequence for less-proficient Thai EFL listeners: A comparative investigation. *RELC Journal, 50*(1), 86–103. https://doi.org/10.1177/0033688218754942

Vandergrift, L. (2006). Second language listening: Listening ability or language proficiency? *The Modern Language Journal, 90*(1), 6–18. https://doi.org/10.1111/j.1540-4781.2006.00381.x

Vandergrift, L., & Baker, S. (2015). Learner variables in second language listening comprehension: An exploratory path analysis. *Language Learning, 65*(2), 390–416. https://doi.org/10.1111/lang.12105

Vandergrift, L., & Baker, S. (2018). Learner variables important for success in L2 listening comprehension in French immersion classrooms. *The Canadian Modern Language Review, 74*(1), 79–100. https://doi.org/10.3138/cmlr.3906

Vandergrift, L., & Tafaghodtari, M. (2010). Teaching L2 Learners How to listen does make a difference: An empirical study. *Language Learning, 60*(2), 470–497. https://doi.org/10.1111/j.1467-9922.2009.00559.x

Vygotsky, L. (1978). Mind in society: The development of higher psychological processes. In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Harvard University Press.* London.

Vygotsky, L. (1996). *Thought and language* (9th ed.). MIT Press.

Wallace, M. P. (2020). Individual differences in second language listening: Examining the role of knowledge, metacognitive awareness, memory, and attention. *Language Learning, 72*(1), 5–44. https://doi.org/10.1111/lang.12424

Wang, Y., & Treffers-Daller, J. (2017). Explaining listening comprehension among L2 learners of English: The contribution of general language proficiency, vocabulary knowledge and metacognitive awareness. *System, 65*, 139–150. https://doi.org/10.1016/j.system.2016.12.013

Warschauer, M. (2005). Sociocultural perspectives on CALL. In J. Egbert & G. M. Petrie (Eds.), *CALL research perspectives* (pp. 41–51). Lawrence Earlbaum.

Yin, R. K. (2011). *Qualitative research from start to finish*. Guilford Press.

Zeng, Y., & Goh, C. (2018). A self-regulated learning approach to extensive listening and its impact on listening achievement and metacognitive awareness. *Studies in Second Language Learning and Teaching, 8*(2), 193–218. https://doi.org/10.14746/ssllt.2018.8.2.2

Zhang, R., & Zou, D. (2021). A state-of-the-art review of the modes and effectiveness of multimedia input for second and foreign language learning. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2021.1896555

Zhang, L., Yu, X., & Goh, C. C. (2022). Improving performance and metacognitive awareness in listening through metacognitive instruction for Chinese language learners. *RELC Journal*. https://doi.org/10.1177/00336882221129645

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.