

RESEARCH

Open Access



The effects of interactionist versus interventionist dynamic assessment models on Iranian EFL learners' speaking sub-skills: a mixed-method study

Saeid Sarabi Asl¹, Mojgan Rashtchi^{2*} and Ghafour Rezaie³

*Correspondence:
mojgan.rashtchi@gmail.com

¹ Islamic Azad University, Science and Research Branch, Tehran, Iran

² Islamic Azad University, North Tehran Branch, Tehran, Iran

³ Islamic Azad University, Garmsar Branch, Garmsar, Iran

Abstract

Dynamic assessment has been proven to effectively promote EFL learners' speaking proficiency, but its implementation in teaching speaking skills has been limited. One of the main reasons is that, thus far, very few studies have scrutinized the impacts of its two main models, interactionist and interventionist, on the speaking sub-skills of EFL learners. Instead of examining speaking as a general skill, this research focused on four speaking sub-skills, including grammatical range and accuracy, vocabulary, fluency, and pronunciation, with a concurrent experimental mixed-method design. To this aim, 30 undergraduate students from the Islamic Azad University, North Tehran Branch, were recruited through convenience sampling and participated in both phases. The quantitative data were collected from participants' pre-and post-test scores in the IELTS speaking module, and qualitative data consisted of the transcribed recordings of the intervention sessions. The integration of the quantitative and qualitative results revealed that while both models positively affected EFL learners' speaking proficiency, they differed in the extent and mechanisms of their impacts. The interactionist DA effectively improved grammatical range and accuracy, pronunciation, and vocabulary depth, and the interventionist DA enhanced fluency and vocabulary breadth more efficiently. The findings dissect the differential effects of DA models in one-to-one speaking lessons, offering practical insights for ELT practitioners and stakeholders.

Keywords: Dynamic assessment, Speaking accuracy, Speaking fluency, Speaking vocabulary, Pronunciation, Mixed-method

Introduction

As one of the significant areas in education, assessment has always been at the center of the attention of ELT practitioners. Syllabus designers and material developers try to make their products more attractive by adjusting them to standardized tests. Teachers find themselves involved in learning about assessment theories to become vigilant evaluators of their learners and help them succeed in their tests. Therefore, there has been an increased focus on formal and standardized assessments among English language

instructors (Volante, 2004). While such assessments might be essential in some contexts, mere focus on the content and requirements of formal standardized tests has given birth to undesirable practices such as teaching to the test. These practices are criticized because they only aim to increase the test scores, not develop the learners' performances and realize their potential (Bennett, 2011). Cheng (2003) asserts that when teaching to the test replaces teaching the learners, the scores do not represent the learners' competencies but show to what extent test training instructions have been helpful. Poehner (2008) claims that the social values gained from higher standardized test scores can hinder language learning. Lidz and Gindis (2003) also note that standardized assessment forms fail to help educators recognize an individual's abilities and potential, so students rarely receive proper guidance from their instructors. One of the attempts to address the shortcomings of formal assessments has been the promotion of "assessment for learning instead of assessment of learning" (William, 2017, p. 984). According to William, assessment for learning focuses on providing feedback and support to students during the learning process rather than solely evaluating their final performance. Harlen (2002) argues that this approach enhances students' understanding of their learning progress and enables them to take ownership of their learning. However, formative assessments have been unable to eliminate teaching-to-the-test issues and have limitations in providing detailed diagnosis and targeted support.

Cheng and Curtis (2004) argue that instead of being a separate or post-teaching activity, assessment should be merged with teaching to form a unified action, providing various forms of support for learners to make real progress. To address such considerations, scholars in assessment have proposed alternative assessment frameworks. One of the most prominent ones, Dynamic Assessment (DA), was inspired by Vygotsky's (1978) Sociocultural Theory (SCT). SCT proponents argue that responses to learners' needs are best addressed when the critical role of social interactions and cultural constructs in the learning process is thoughtfully considered (Lantolf & Thorne, 2006). This form of assessment, embedded in instruction, attempts to gauge and improve the learning capacities of learners at the same time. Feuerstein (2000) believes that bonding assessment and instruction helps learners perform beyond their current ability level, meaning that implementing DA can move learners to higher levels of development. Poehner (2009) further argues that the core emphasis of DA, which is the co-creation of dialogical interaction between teacher/examiner and students, merging instruction and assessment, enables a mediator to help learners more purposefully.

While some models and categorizations of DA have emerged in the last two decades, Lantolf and Poehner's (2004) interventionist and interactionist labels for categorizing the DA approaches seem to be the most proper, as other distinctions and versions fall under one of these models. Generally, what differentiates the two approaches is their way of providing mediation. According to Lantolf and Poehner (2007), interactionist DA has close links with what Vygotsky regarded as dialoguing cooperatively, and mediations materialize due to interaction between the teacher/assessor and the learner. The main focus of the interactionist DA approach is enhancing learners' skills without considering such issues as the amount of effort and time. Moreover, the interactionist model does not define a clear point of achievement.

On the other hand, Lantolf and Poehner (2011) point out that interventionist DA is closer to the psychometric origins of NDA. It utilizes systematic assistance practices to develop quantifiable results for between and within-group comparisons. Unlike interactionist DA, this model regards the pre-planned mediations as an indicator of the pace of learning to estimate the needed amount of effort and time to reach a certain final point. In interventionist DA instruction, mediation moves from the most implicit to the most explicit for accurate responses; if learners fail to accomplish a task, the instructor provides them with needed prompts.

DA can be a viable alternative to the current teaching and assessing frameworks with its solid theoretical basis and the advantages of its main models in ELT. However, according to Haywood and Lidz (2007), DA has not drawn considerable attention among English language teachers, especially in speaking lessons. One of the main reasons for teachers' reluctance to implement DA might be because they do not clearly understand the distinct differences between the DA models in speaking sub-skills. Therefore, they might not be sure which model to employ based on their lesson objectives.

Conducting mixed-method research studies on the impacts of DA models on specific speaking subskills seems necessary for two main reasons. First, speaking sub-skills differ in rate and improvement route, and DA studies considering speaking a single-dimensional construct do not provide a precise picture (Dincer, 2017). Second, according to Riazi and Farsani (2023) while MMR designs have been receiving more attention from ELT scholars in the past decade, they have not been utilized in strands proportionally equally. That is, the body of mixed-method research studies on such areas as testing and assessment has remained limited. Therefore, focusing on speaking sub-skills and the dominance of quantitative studies in this area stimulated the incentives for the present study. The insights provided by the present study provide ELT researcher and practitioners with more profound familiarity with the DA models differences in developing EFL learners' speaking sub-skills, helping them choose and implement the DA model that better aligns with their lesson outcomes.

Literature review

DA has its roots in Vygotsky's Sociocultural theory. SCT states that "higher forms of consciousness, such as voluntary control of memory, perception, and attention, occur through a process of internalization whereby these functions initially occur as an interaction between human beings but are then transformed into cognitive abilities" (Poehner et al., 2018, p. 5). The most well-known component of Vygotsky's (1978) theory of mind that has influenced psychology and education is the Zone of Proximal Development (ZPD), which refers to the gap between a person's level of ability to function with and without mediation. Vygotsky asserts that the ZPD serves as the basis for development-oriented pedagogies. His argument that emergent abilities are most amenable to instructional intervention encouraged scholars in education and assessment fields to focus on practical applications of the ZPD later, such as DA (Poehner, 2008; Tzuriel, 2012).

One of the distinguishing elements of DA is the provision of mediation during assessment practices, which is also referred to as scaffolding. Scaffolding is deemed a tenet of DA, one of the means of actualizing learners' ZPD. Through scaffolding, the contribution of the more expert partner is lowered. At the same time, the learner becomes

increasingly skilled, meaning that scaffolding becomes less and finally removed until the learner can perform alone (Lantolf, 2006). Through scaffolding, teachers ease students into comprehension and mastery of new skills by incrementally reducing the teacher's guidance. Rashtchi (2019) states that scaffolding is a crucial factor in developing the productive skills of ESL learners as it can provide students with meta-cognitive strategies and thought-provoking opportunities. As students grapple with tasks outside their current competency—often called the ZPD, they are driven to analyze, synthesize, and evaluate information. As such, scaffolding helps expand the learners' ZPD by allowing them to test and develop their language skills in supportive and stimulating environments. Meta-cognitive strategies may involve learners planning how they will approach a task, monitoring their comprehension as they carry out the task, and evaluating the effectiveness of their approach afterward. The scaffolding mediations in DA assist L2 learners in performing better on tests and going above and beyond their existing abilities (Lantolf & Poehner, 2011; Lantolf & Thorne, 2006; Poehner, 2008). According to Lantolf (2006), DA is based on obtaining a more accurate view of L2 learners' abilities, meaning that assessments should be supplied with incremental, consistent mediations. Thus, DA is an interactive assessment in which L2 learners are given gradational, congruent mediations to help them progress beyond their current capacities (Haywood & Tzuriel, 2002; Lantolf & Poehner, 2004).

Dynamic Assessment has proven to impact EFL learners' affective and linguistic development effectively. Studies suggest that implementing DA can help with second language learners' affective factors, such as promoting their intrinsic motivation (Azizi & Khafaga, 2023; Malmeeer & Zoghi, 2014), lowering debilitating anxiety (Pishghadam et al., 2011; Sherkuzyeva et al., 2023) and developing facilitative anxiety. As for language skills, studies have found DA advantageous in developing language skills such as writing (Ebadi & Rahimi, 2019; Farrokh & Rahmani, 2017), reading (Naeini & Duvall, 2012; Yang & Qian, 2017), listening (Ebadi et al., 2023; Hidri, 2014), and speaking skills (Riswanto et al., 2023; Ritonga et al., 2022).

DA main models

Different interpretations of the ZPD have given birth to various DA practices and manifestations. Those DA researchers who consider the ZPD a cognitive development tool tend to adjust DA within the larger framework of standardized assessment types through standardized procedures. Such interpretation in the discourse of DA is known as the interventionist model. According to Poehner (2008) in the interventionist model, "mediators are not free to respond to learners' needs as these become apparent during the procedure but must instead follow a highly scripted approach to mediation in which all prompts, hints, and leading questions have been arranged hierarchically" (pp. 44–45). The motivation behind interventionist approaches is addressing the test's objectivity, making them comparable to mainstream assessments.

Through the interventionist DA model, learners receive tasks; if needed, standardized mediations are offered to finish a given task successfully. When learners gain the competence to solve problems without mediation, examiners seek to determine to what extent the test takers can transfer their skills to a new problem independently. To this end, examiners give learners some "novel examples" of the initial problem, including

"near transfer" problems, which require the learners to apply the same principles to the original task in an alternative combination. Next, assessors present "far transfer" problems to the learners, which demand the use of more advanced but related principles, and finally, "very far transfer" problems, which are more complicated than the original problem (Campione et al., 1984, p. 81). In the Graduated Prompt Approach, the researcher can create learners' profiles based on their performance throughout the procedure. The learners' profiles measure how quickly learners internalize new principles and how far they can extend the new knowledge to new problems (Ferrara et al., 1986). The types of mediation and the amount of interaction based on this interpretation of the ZPD are pre-planned. This interpretation of the ZPD, while bringing about the comfort of better generalizability and large-scale administration, deprives learners and instructors of co-creating a ZPD, just as Vygotsky had described.

On the other hand, some DA researchers maintain that DA is a critical tool for intellectual development, which can be achieved through flexible interaction between the mediator and the learner. This interpretation of SCT is regarded as interactionist DA. According to Feuerstein et al. (2006), interactionist DA highlights "how stimuli experienced in the environment are transformed by a mediating agent, usually a parent, teacher, sibling, or another intentioned person in the life of the learners" (p. 25). Interactionist DA centers on cooperative dialogue as recommended by Vygotsky and is sensitive to the ZPD of the learners. Lantolf and Poehner (2004) maintain that interactionist DA approaches prioritize instruction and learning to the speed and efficiency of development. That is why interactionist DA does not regard time and effort as measurement tools and, therefore, has no predetermined endpoint. This feature makes interactionist DA models more in line with the Vygotskian assessment perspective. Additionally, interactionist DA does not require mediators to follow specific rules and hierarchies to respond to the learners' needs during DA procedures. The DA sessions are mainly shaped by tasks or problems that mediators and learners cooperatively try to complete and are highly individualized and interactive, and interventions emerge as needed.

Related studies

Several studies have investigated the impact of these dynamic assessment approaches on EFL learners' speaking subskills. As for pronunciation, Yang and Qian (2017) conducted an experimental mixed-method investigation on the effects of DA on Chinese English learners' pronunciation proficiency. The study included 36 college students majoring in English who had difficulty improving their pronunciation. In addition, the study utilized two questionnaires to gauge the participants' attitudes, motivation, and anxiety in a pronunciation class. Yang adopted an interventionist DA model and provided oral medication to the participants in the experimental group in the form of hints, suggestions, explanations, or demonstrations. The findings showed that the implementation of DA significantly impacted the pronunciation proficiency of the experimental group since it gave learners a more positive attitude, higher motivation, and lower anxiety. Shafiee et al. (2018) studied the effect of interactionist DA teaching English rhythm to 30 Iranian EFL learners in a mixed-experimental design. The results revealed that the experimental group outperformed the control group in the post-test, mainly due to their more

positive attitude in the presence of DA implementation. The results of the experimenters' questionnaires provided quantifiable verification of the claim.

Regarding fluency, Kao (2020) reported the results of her exploration of the influence of the interactionist DA on Chinese EFL learners' oral fluency in two levels of proficiency: elementary and advanced. She researched at a Taiwanese university with 119 first-year students, and her finding suggests that interactionist DA does not benefit learners' pronunciation at the lower level of proficiency since the teacher's constant intervention stops the progress of learners' content development. Safdari and Fathi (2020) investigated the effect of DA on the speaking fluency of Iranian EFL learners in an experimental study. They found that implementing DA did not significantly impact the participants' fluency.

The effectiveness of interactionist and interventionist DA in improving the grammatical knowledge of EFL learners has been studied by several researchers. Jafary et al. (2012) found that applying interactionist DA improved the grammatical knowledge of Iranian male pre-university students. Estaji and Ameri (2020) conducted a study on pre-intermediate and upper-intermediate EFL learners. They found that interventionist DA was significantly more effective in developing the participants of the experimental groups, with lower-level participants' post-test scores higher than high-intermediate ones. The learners in the experimental groups had a more positive attitude toward grammar learning and preferred DA techniques over conventional ones. The findings of the study are in line with similar studies conducted by Alavi et al. (2012), Farangi and Kheradmand Saadi (2017), and Shabani (2012), which indicated that the success of DA could be attributed to teachers' accurate diagnostic feedback, learners' involvement in DA interactions, and the possibility of discussing grammar issues with their teachers.

Researchers have explored the use of DA in vocabulary learning and retention of English learners in various settings, especially with direct links to speaking skills. Hessamy and Ghaderi's (2014) quasi-experimental study on the role of DA in the vocabulary learning of 50 male adult learners in Iran revealed that the learners who received interventions outperformed the control group members. The study conducted by Mirzaei et al. (2017) investigated the effect of interactionist Cumulative Group Assessment (CGA) on the depth of vocabulary knowledge in 60 female Iranian junior high school students. The results showed that the GDA group outperformed the control group in vocabulary depth of speaking in immediate and delayed post-tests. The study suggests that implementing the interactionist cumulative GDA can help learners gain confidence, a positive attitude towards vocabulary learning, retention, and use, and a higher motivation to keep developing their vocabulary knowledge.

Despite all positive support from studies for the effectiveness of DA on speaking skills, two significant issues appear to be under-explored and, thus, call for further investigations. First, more empirical studies must compare the effectiveness of interactionist and interventionist DA approaches in terms of specific speaking sub-skills. The existing studies, such as Ebrahimi (2015) and Safdari and Fathi (2020), face generalizability challenges and do not include pronunciation and vocabulary sub-skills. Second, most investigations comparing DA models' impact are predominantly quantitative; hence, they could not capture how ZPD-sensitive feedback exchanges aid in developing speaking subskills

in learners. To provide a more in-depth comparison of DA models, the present study aimed to investigate the following mixed-method research question:

In what ways and to what extent do the interactionist and interventionist DA differ in their impact on Iranian EFL learners' speaking skills components, namely accuracy, fluency, pronunciation, and vocabulary?

The mixed-method research question was broken into quantitative and qualitative questions to facilitate finding the answers.

1. Do interactionist and interventionist DA have different effects on the speaking components, including grammatical accuracy and complexity, fluency, pronunciation, and vocabulary of intermediate Iranian EFL learners?
2. How do interactionist and interventionist DA differ concerning their effects on the intermediate Iranian EFL learners' speaking accuracy, fluency, pronunciation, and vocabulary?

Method

This study utilized a concurrent mixed-method experimental design to answer the main research question. According to Riazi (2016), this design is a procedure where researchers collect both quantitative and qualitative data simultaneously and then integrate the two forms of data to interpret the overall results to provide a more comprehensive picture of the research problem than a standalone method. The quantitative phase of the study's design was a two-group-pre-test-post in which the main participants were randomly assigned to two experimental groups through the block randomization method, which divides participants into similar-sized groups (Suresh, 2011). In this study, the two experimental groups, namely the interactionist and interventionist groups, each had 15 participants. The quantitative data were the participants' scores in the IELTS speaking test's pre-and post-test administration.

Regarding the qualitative phase, the data was collected from all participants during the treatment stage of the quantitative phase. All the instructional sessions for all participants were recorded, transcribed, and served as the raw data for qualitative analysis. Inductive thematic analysis was employed to analyze the dataset. In line with Braun and Clarke's (2006) framework, the researchers familiarized themselves with the data by reading and re-reading the transcripts to search for meaning, potential patterns, and initial ideas. The researchers separated the data into consistent units that could stand alone and independently within the text. Then, the transcripts were uploaded to Delve software and coded line-by-line. In this process, such instances as learners' linguistic errors, misunderstandings, self-corrections, and indications of learning strategies were examples of meaningful units and subject to investigation. The transcriptions of the interactions during the sessions were broken into excerpts. The researchers grouped the excerpts into 82 codes, which were thoroughly analyzed to identify the potential concepts or indicators of the differences between the impacts of the two treatments. Next, the researchers constantly compared and contrasted the initial codes and regrouped them to establish more focused codes. The result included 33 focused codes, which were regrouped around the most predominant codes. The codes were selected based on the researchers' interpretation and helped them to reach theoretical conceptualizations. It

is worth mentioning that the focused codes were of two types; some were only present in one of the experimental groups, and some were common but manifested themselves differently. The researchers grouped the codes into categories and themes. The themes were reviewed and refined repeatedly, which included merging the overlapping themes and breaking down the overgeneralized ones until the refinements did not lead to any substantial changes.

A connection was drawn between the two datasets after completing the separate analyses of the quantitative results (derived from pre-test and post-test scores) and qualitative results (gathered from thematic analysis of learners and instructor interactions). The emergent themes from the qualitative data and the trends from the quantitative data were juxtaposed, which allowed the researchers to link the observations with statistical results for a comprehensive understanding. Specifically, the qualitative findings provided an in-depth explanation and context to the quantitatively measured effects of the different treatments on the participants' speaking subskills.

Participants

The researchers drew the sample from a population of adult Iranian EFL learners. Convenience sampling was used to recruit the participants from the existing network of students in the English department of Islamic Azad University, North Tehran Branch. Fifty-one Iranian undergraduate students majoring in Teaching English as a Foreign Language (TEFL), aged between 21 and 35, with intermediate English proficiency, participated in the study. All of them were residing in Tehran during the experiment period, were native speakers of Persian, and had already experienced learning English via the Google Meet platform. To ensure the homogeneity of the learners in terms of general English language proficiency, the participants took the B1 Preliminary test. Thirty learners (including 19 female and 11 male students) whose scores were between 135 and 152 ($M=142$, $SD=8.18$) in the B1 Preliminary test were selected as the legitimate participants of the study.

In line with the ethical research protocol, all participants were informed of the purpose and the detailed procedure of the study. In addition, they were reminded that they could leave the study at any point if they did not feel comfortable. The participants provided their written informed consent to the researchers at the beginning of the study and indicated their voluntary participation. They were ensured that their personal identities would not be revealed to any other individuals, and their responses would be handled carefully to ensure confidentiality. Moreover, the intervention sessions were scheduled according to the participants' availability and preferences to ensure their comfort.

Instruments

B1 preliminary test

A sample of past papers of B1 Preliminary (paper version) was used to select the most homogeneous group of participants regarding language proficiency. The B1 Preliminary test is an intermediate-level assessment of English language proficiency for adult learners, consisting of three papers: Reading and Writing (90 min), Listening (30 min), and Speaking (12 min). The B1 Preliminary test is a valid instrument of general language

proficiency for L2 learners and is scored from 120 to 170, with scores between 153 and 159 corresponding to CEFR's B1 level of proficiency. As for reliability, the B1 Preliminary has been shown to enjoy indices of 0.92 (Cambridge Assessment English, n.d.) and 0.86 in the current study.

IELTS speaking test

As four speaking sub-skills, namely speaking accuracy and complexity, fluency, pronunciation, and lexical resources, were the dependent variables of the current research, the IELTS Speaking test was used as a pre-and post-test to measure the variables objectively. Using the IELTS speaking test for pre-and post-test stages required the test to have the face and construct validities. O'Loughlin (2001) maintains that the speaking component of the IELTS test is administered directly, and a direct speaking assessment provides the most salient evidence for face validity. The IELTS speaking construct is generally defined as oral competency; however, from an examiner's perspective, this speaking construct is further defined by the rating criteria, which is an operationalized format of the primary characteristics described in detail. Thus, the IELTS speaking test reduces oral proficiency to four overarching variables: fluency and coherence, lexical resource, grammatical range and accuracy, and pronunciation. While it is true that the speaking sub-skills are not limited to the present study's variables, the IELTS speaking test's rubrics and components are the best match for objectively measuring them in the quantitative phase of the study.

The IELTS speaking module consists of three sections of a recorded oral interview between a candidate and an examiner, and the entire test takes about 15 min. The speaking module consists of three sections of a recorded oral interview between a candidate and an examiner. Candidates answer broad questions about themselves and their lives in part one, and then in section two, the examiner gives them a card about a specific topic about which they must speak for one to two minutes. In this part, candidates have one minute to prepare and take notes if they want. Part two ends with the examiner asking one or two rounding-off questions. In the last section, the examiner prompts candidates to discuss more abstract topics, which are thematically related to the topic in part two. Candidates receive a band score between 1 and 9 for each criterion (speaking sub-skills).

The overall speaking band score is the average of the scores obtained in these four criteria. In this study, the members of both groups took the IELTS speaking as the pre-test to ensure the equality of the participants in subskills of speaking proficiency such as fluency, accuracy, pronunciation, and lexical resources. After the treatment stage, the participants were given the IELTS speaking test as a post-test for statistical analysis purposes. One of the researchers administered the pre-and post-tests through one-to-one online interviews. The test was recorded so that another research team member, as the second rater (with a Ph.D. in TEFL and 35 years of experience), could do the second round of scoring. The final score of the participants was determined by averaging the scores given by each rater. Both raters independently assessed the candidate's performance using the IELTS public version speaking rubrics. The individual scores were added and divided by two to calculate the average score. This process ensured fairness and consistency in scoring. The inter-rater reliability between the scorers was calculated,

resulting in a coefficient of 0.82 for the pre-test and 0.79 for the post-test, which indicated an appropriate level of agreement in their assessments.

Instructional materials

The instructional instruments used in this study were authentic materials and aimed at eliciting oral responses from participants. The type of material and speaking tasks were the same for all participants in both groups. However, the topics varied based on the negotiation between the researchers and each participant before the treatment phase. That is, the topics were not selected in advance but were selected based on each participant's interests and background information. Therefore, the topics varied between the experimental groups and within the groups. Each instrument was used two times in two separate sessions so the learners could transfer and use the mediations they received in the first session. The instruments' list and corresponding tasks were as follows: News stories: Each participant could select a recent news story based on their choice of topic (such as a discovery) from the BBC News website. They narrated the news in their own words and explained why they learned from them. Ted-Ed videos: Learners could choose one of Ted-Ed's educational videos aligned with their choice of topic. After watching the clip, they summarized it and shared their thoughts. Book/Movie Reviews: The participants could choose a movie or a novel they liked, do some online research about it (e.g., IMDB, Wikipedia), and discuss the plot, the characters, the moral of the story, and their opinions. All materials and task instructions were given to the students in advance so that they had sufficient preparation time.

Procedure

The treatment stage of the present experiment took ten sessions for each individual, meaning there were 300 hundred sessions. The sessions were held one-on-one via Google Meet, and the instructor (the first author) was the same in both groups. Each instructional session lasted for 60 min and was conducted one-on-one. All sessions were recorded with the participants' written consent for the qualitative phase of the experiment. Each session began with small talks and the news task. Then, participants had to deliver the planned task of the session. As mentioned earlier, the materials were given to the learners in advance. Depending on the task and mediation type, the participants had to repeat the task twice or three times until no mediation was required. According to Poehner and Lantolf (2005), meaningful progress toward the learner's ZPD should be evident in tasks other than the original tasks; transfer tasks should be incorporated into the DA programs to determine how far learners' knowledge and abilities can be extended beyond the assessment environment and how much assistance they need. In the present study, the second assignment of the tasks (e.g., story narration) served as the transfer tasks. Although the task types were the same in both groups, the mediations offered by the instructor differed. The following sub-sections provide explanations and samples of mediation.

Mediations in the interventionist group

The mediation provided in the interventionist group was carried out in the most implicit to the most explicit order and concluded with an accurate response. To this end, Lantolf and Poehner's (2011) mediation scale was employed based on each student's responses. The instructor's mediations focused on all four components of speaking: grammatical range and accuracy, vocabulary, fluency, and pronunciation. If the student's response was correct, mediation was unnecessary. However, when the student's response was incorrect, the instructor chose one of the eight possible responses from the scale in the following forms:

1. The instructor paused.
2. The instructor questioned the syllabus, the entire word, or the phrase.
3. The instructor repeated only the incorrect part of the utterance.
4. The instructor posed a question: What is wrong with this sentence?
5. The instructor pointed out the inaccuracy.
6. The instructor posed either/or questions.
7. The instructor clarified the correct answer.
8. The instructor explained why the participant's response was inaccurate.

The list progressed from the most implicit to the most straightforward way of delivering mediation to the interventionist group participants. The instructor maintained his mediation systematically, thanks to the mediation inventory. The following excerpt, which is part of a conversation between the instructor (I) and Participant F, is an example of how the mediation helped the participant pronounce a word correctly.

F: "I disagree with the presenter because luck [lɒk] is not the most important thing in success."

I: (Paused, signaling something was wrong in the utterance)

F did not notice anything wrong in their utterance and thought the instructor was surprised by their opinion. Therefore, the instructor moved to the second level of mediation; he repeated the part of the sentence, which entailed the pronunciation error with rising intonation.

I: "lock is not the most important thing"?

F noticed the error in the second half of their response but failed to identify its type and repeated the word incorrectly.

Moving to the next level of mediation, the instructor explicitly mentioned the wrong pronunciation of the word:

I: "Lock."

F: "Oh..... luck [lɒk] is not the most important thing in success."

Mediations in the interactionist group

The interactionist mediations were not constructed in advance but were developed due to a rigorous analysis of the mediator's cooperative discussions with each learner during each session. The mediations were constructed like Aljaafreh and Lantolf's (1994) Regulatory Scale. The regulatory scale's key principle is contingency, which means that

assistance is offered contingent on learners' needs and is progressively adjusted depending on their responses. The main goal is to gradually reduce teacher intervention as learners become more capable of independent function. However, since the interactionist DA's distinguishing characteristics establish flexible grounds for effective learner development, the instructor did not consider the Regulatory Scale as specific and prescriptive rules or standards. The interactionist interventions served two purposes: first, to address the learners' difficulties, and second, to determine whether the participants found the mediations beneficial. The following excerpt, part of the conversation between the instructor and Participant M, demonstrates how interactionist mediation was offered for a participant in the interactionist group. This participant was particularly weak at using appropriate English tenses in speaking. Their responses indicated that there were many instances of L1 transfer.

I: "Oh, that's a great one! I've seen that movie three times. Can you tell me about a movie you've seen more than once?"

M: "I saw the Gladiator two times!"

Instructor: "Well done in communicating your idea, but your sentence had a grammatical error, do you know what it was?"

M: "[pauses], I guess I shouldn't use 'the' before the name of movies."

I: That's true in some cases, but for this movie, you used 'the' correctly. Try to find the error based on our previous session's discussion.

M: "You mean simple past?"

I: "You're getting there!"

M: "I suppose the correct sentence is: I have seen the movie two times!"

I: "Yes, you got it! Now can you remember why we use the present perfect instead of the simple past?"

M: "I'm sorry, I just don't remember the rule!"

I: "It's ok! In English, we usually use the present perfect tense to talk about a non-specific time in the past. So, you should say 'I've seen the Gladiator two times.' Can you give me other examples?"

M: "Like, I've been to Europe two times, ... or I've visited the Golestan Palace three times."

I: Perfect! Do you think you'll remember the rule?

M: "I think I have a better understanding now."

I: "What do you think you can do to avoid such errors?"

M: "I took notes about what you say, but I guess I mustn't think in Farsi [Persian] words."

Data analysis

Quantitative results

As Table 1 shows, the mean values for all variables were similar between the groups, and there was no notable difference in pronunciation, fluency, coherence, grammatical range, accuracy, or lexical resources. To examine the difference between the variables in the two experimental groups at the pre-test stage, a one-way MANOVA was employed. Before running the test, the assumptions of multivariate normality,

Table 1 Descriptive statistics of pre-test scores

Group	N	Mean	Std. deviation	Skewness		
				Statistics	Std. Error	Ratio
<i>Interactionist</i>						
Pronunciation	15	4.40	.94	-.32	.58	-.55
Fluency and coherence	15	4.16	.95	.36	.58	.62
Grammatical range	15	4.40	.89	-.00	.58	-.00
Lexical resources	15	4.40	1.02	-.15	.58	-.25
<i>Interventionist</i>						
Pronunciation	15	4.56	1.04	-.04	.58	-.06
Fluency and coherence	15	4.70	.79	.33	.58	.56
Grammatical range	15	4.40	.89	-.00	.58	-.00
Lexical resources	15	4.16	.93	.33	.58	.56

Table 2 Results of one-way MANOVA for pre-test scores

Effect	Value	F	Hypothesis df	Error df	Sig	Partial Eta Squared	Observed power ^b
Group							
Pillai's trace	.226	1.401 ^a	5.000	24.000	.259	.226	.407
Wilks' lambda	.774	1.401 ^a	5.000	24.000	.259	.226	.407
Hotelling's trace	.292	1.401 ^a	5.000	24.000	.259	.226	.407
Roy's largest root	.292	1.401 ^a	5.000	24.000	.259	.226	.407

^a Exact statistic

^b Computed using alpha = .05

outliers, linearity, multicollinearity, and homogeneity of covariance matrices were examined in both groups of participants. In the first stage, univariate outliers and normality were checked using boxplots and the Shapiro–Wilk test. Results indicated that all variables were normally distributed in interactionist and interventionist groups, with no significant departure from normality for all dependent variables in groups (p values > 0.05). Mahalanobis's distances were calculated. All values were less than the critical value of 13.68, supporting the multivariate normality of the distributions and indicating no multivariate outliers.

In the second stage, linearity, multicollinearity, and homogeneity of covariance matrices were examined. Scatterplot matrices indicated that both groups' dependent variables were approximately linearly related. The multicollinearity assumption was investigated using the tolerance and Variance Inflation Factor (VIF) indices. Results showed no multicollinearity issues, with Tolerance indices not lower than 0.1 and VIF values not exceeding five. Finally, the homogeneity of covariance matrices was examined using Box's M test, which showed a value of 10.86 associated with a non-significant p -value of 0.919, supporting the equality of covariance matrices between the interactionist and interventionist groups.

Table 3 Descriptive statistics of post-test scores

Group	N	Mean	Std. deviation	Skewness		
				Statistics	Std. error	Ratio
<i>Interactionist</i>						
Pronunciation	15	5.70	.84	-.21	.58	-.36
Fluency and coherence	15	4.19	.75	.78	.58	1.34
Grammatical range	15	5.80	.92	-.546	.58	-.96
Lexical resources	15	5.03	.91	.61	.58	1.05
<i>Interventionist</i>						
Pronunciation	15	4.43	1.09	.00	.58	.00
Fluency and coherence	15	6.38	.74	-.72	.58	-1.24
Grammatical Range	15	5.00	.96	-.20	.58	-.34
Lexical resources	15	4.96	.61	.14	.58	.24

Table 4 Results of on-way MONVA for post-test results

	Value	F	Hypothesis df	Error df	Sig	Partial eta squared	Noncent. parameter	Observed power ^b
Pillai's trace	.857	28.742 ^a	5.000	24.000	<.001	.857	143.712	1.000
Wilks' lambda	.143	28.742 ^a	5.000	24.000	<.001	.857	143.712	1.000
Hotelling's trace	5.988	28.742 ^a	5.000	24.000	<.001	.857	143.712	1.000
Roy's largest root	5.988	28.742 ^a	5.000	24.000	<.001	.857	143.712	1.000

^a Exact statistic

^b Computed using alpha = .05

After verifying the assumptions, a one-way MANOVA was conducted to compare the pre-test results. As presented in Table 2, the outcomes revealed no significant MANOVA effect, $F(5, 24) = 1.401$, $p = 0.259$, Wilk's $\lambda = 0.774$. The findings indicated that there was no significant difference in speaking skills, including accuracy, fluency, vocabulary, and pronunciation, between the participants of the interactionist and interventionist groups prior to the treatments.

Post-test results

Table 3 shows that the interactionist group had a higher mean score for pronunciation, while the interventionist group had higher mean scores for fluency and coherence. In terms of grammatical range and accuracy, and lexical resources, the interactionist group had slightly higher mean scores than the interventionist group.

A one-way MANOVA was used to test the research hypothesis, and prior to running the MANOVA, the test assumptions, similar to the pre-test stage, were checked. The univariate outliers and normality were checked using boxplots and the Shapiro–Wilk test, indicating that all variables were normally distributed in interactionist and interventionist groups (p values > 0.5). The Mahalanobis's distances were calculated, and the results did not suggest any multivariate outliers. Scatterplot matrices showed that each group's variables were roughly linearly related. The Tolerance and VIF values were calculated, and the results indicated no multicollinearity issues in the distributions. The Box's M value of 29.027, associated with a p -value of 0.078, supported the equality of

covariances matrices between the groups. Then, a one-way MANOVA was run to find out if there were any significant MANOVA factors in all variables. Table 4 shows the results.

As Table 4 demonstrates, the Multivariate test results indicated a significant MANOVA effect, $F(5, 24) = 1.401, p < 0.001$, Wilk's $\lambda = 28.742, \eta_p^2 = 0.857$, suggesting difference(s) between the interactionist and interventionist groups in the mean values. The partial eta squared (η_p^2) value of 0.857 suggests a very high effect size, indicating that the dynamic assessment models (interactionist and interventionist) account for approximately 85.7% of the variance in our dependent variable, the speaking subskills of EFL learners.

Leven's Test was run before investigating the Between-Subject Effects and testing each of the study's hypotheses. Based on a series of the F test results, the homogeneity assumption of variances was satisfied and assumed equal. Table 5 demonstrates the results of the Test of Between-Subject Effects.

Results for speaking accuracy and complexity

The researchers used a Bonferroni-adjusted alpha of 0.025 to reduce the chances of false-positive results. The results showed a significant effect on the grammatical range and accuracy scores, $F(1, 28) = 4.88, p = 0.024, \eta_p^2 = 0.310$, with scores higher in the interactionist group ($M = 5.8, SD = 0.92$) compared to the interventionist group ($M = 5.00, SD = 0.96$). The result suggest that interactionist DA is moderately more effective in developing Iranian EFL learners' grammatical range and accuracy in speaking.

Results for speaking fluency

The results indicated a significant effect on the Fluency and Coherence scores, $F(1,28) = 35.81, p < 0.001, \eta_p^2 = 0.692$, with scores higher in the interventionist group ($M = 6.38, SD = 0.74$) in comparison to the interactionist group ($M = 4.59, SD = 0.75$). The results suggest that the interventionist DA is significantly more effective in developing Iranian EFL learners' speaking fluency than the interactionist DA.

Table 5 Results of the *between-subjects effects*

Source	Dependent variable	Sum of squares	df	Mean square	F	Sig	Partial eta squared	Noncent. parameter	Observed power ^a
Grouping	Pronunciation	12.033	1	12.033	12.557	.001	.310	12.557	.873
	Fluency and Coherence	35.861	1	35.861	63.338	<.001	.693	63.338	1.000
	Grammatical Range and Accuracy	4.880	1	4.880	5.682	.024	.169	5.682	.514
	Lexical Resources	.033	1	.033	.055	.816	.002	.055	.029

^a Computed using alpha = .025

Results for pronunciation

There was a significant effect on the pronunciation variable, $F(1,28) = 12.033$, $p = 0.001$, $\eta_p^2 = 0.31$, with the interactionist group's score ($M = 5.70$, $SD = 0.84$) higher than the interventionist group's ($M = 4.43$, $SD = 1.09$). The evidence suggests that the interactionist DA is moderately more effective in improving the EFL learners' pronunciation skills.

Results for vocabulary

For the fourth hypothesis of the investigation, the results did not indicate a significant effect on the lexical resources variable, $F(1,28) = 0.033$, $p = 0.816$, between the interactionist group ($M = 5.03$, $SD = 0.91$) and the interventionist group ($M = 4.43$, $SD = 1.09$). The evidence supported the lack of significant differences between the impacts of interactionist and interventionist DA models on the Iranian EFL learners' speaking vocabulary.

Qualitative findings

The analysis of the qualitative data revealed that the DA models differed in their impacts on Iranian EFL learners' speaking subskills due to the following items: (a) the degree of negotiated mediations, (b) enhancing different communicative strategies, (c) developing different learning strategies, and (d) promoting different preparation orientations.

The degree of negotiated mediations

The degree and quality of collaborative interactions between the instructor and the individual learners were among the main factors affecting the impact of the interactionist and interventionist treatments. Since the researcher did not have to follow standardized feedback and prompts in the interactionist approach, there were more meaningful collaborative interactions than in the interventionist approach. Therefore, the instructor and the learners could negotiate mediation, providing ample opportunities for deeper learning.

Thanks to the negotiated mediation, the length and depth of mediation went beyond the task level enabling the instructor to identify the source(s) of the learners' difficulties and address them effectively. For instance, when the instructor realized a learner's utterance contained fossilized errors in grammar and pronunciation, he did not stop when the learner produced an error-free utterance but introduced the fossilization concept and some standard techniques to tackle the fossilized errors. Then the teacher and the learner engaged in a collaborative conversation to choose a technique that fitted the student's style and preferences. As a result of the learner's active contribution to the learning process, they were more willing to employ the techniques to spot their fossilized errors and rectify them autonomously.

In the case of interlingual errors, especially in pronunciation and grammar, in the interventionist DA, the most explicit mediation went far only to explain why the learner's response was erroneous. Nevertheless, in the interactionist approach, the mediator not only helped the learner with the task at hand but had the opportunity to provide a simple contrastive analysis and draw the learners' attention to the linguistic or cultural differences between the L1 and L2 and elicit more instances from the learner. Thus,

familiarizing the learners with the negative transfer concept resulted in fewer interlingual errors in the subsequent sessions.

Another benefit of the negotiated mediation in the interactionist group was the opportunity the learners had to internalize some critical yet often neglected aspects of English pronunciation, such as the rhythm. While the suprasegmental aspects of English pronunciation were also addressed in the interventionist group, the mediation's length and depth did not allow the instructor to familiarize all the participants with the underlying concepts and the governing rules. However, in the interactionist approach, the instructor's mediation included guided discovery techniques or detailed explanations to help learners find out the rules and patterns concerning the rhythm of pronunciation when speaking in English.

Enhancing different communicative strategies

The second emerging theme regarding the differences between the interventionist and interactionist DA models was their impact on promoting different communicative strategies. The data suggests that the scope and flexibility of the interactionist mediations were more effective in helping the learners develop achievement strategies. However, the structured interventionist mediations contributed to enhancing the reduction strategies. Færch and Kasper (1983) assert that accomplishment communication techniques incorporate hypothesis and the communicator's practical assertion and can facilitate language learning. Conversely, when reduction communication tactics are utilized, the original objective is altered and may result in less language learning.

Such features of the interventionist DA as standardized and systematic evaluation and feedback provision required explicit endpoints where the learners could achieve the task requirements without the mediator's help. The existence of predetermined endpoints in this approach reinforced the participants to use reduction strategies such as uncomplicated (or simplified versions of) words, expressions, and structures to avoid errors, minimize the frequency and explicitness of the mediations, and increase their fluency. Nonetheless, implementing the reduction strategies impeded the learners from expanding their range of grammatical structures and lexical resources.

On the other hand, the participants in the interactionist group utilized achievement strategies more often. For instance, the learners employed appealing strategies at the onset of their responses to ensure the appropriateness of their choice of vocabulary, grammatical structures, or content organization. In addition, they resorted to this strategy, accompanied by code-switching, to ensure their understanding of newly learned grammatical, lexical, and phonological items.

Developing different learning strategies

The third difference regarding the impacts of the interventionist and interactionist DA models on speaking subskills was rooted in developing different learning strategies. According to Oxford (2002), the learning strategies can be categorized based on the mental processing of the target language or providing support for language learning. Those strategies that directly include the target language, such as memory, cognitive, and compensation strategies, are the direct ones. Indirect strategies, on the other hand,

help learners through focusing, planning, and self-evaluation and include metacognitive, affective, and social strategies.

Based on the data analyses, the interventionist mediations provided more opportunities for learners to boost direct learning strategies, but the interactionist model contributed more to the indirect strategies. As stated before, since the interventionist group participants tried to reduce the explicitness of the instructor's interventions, they were more exposed to implicit mediations that demanded more elicitations and, consequently, more adoption of guessing strategies.

The promotion of guessing as a subset of compensation strategies helped the participants in the interventionist group bypass their fluency obstacles by guessing less complicated linguistic items in terms of morphology and pronunciation when responding to follow-up questions for which they had to improvise. Furthermore, the interactions and the researchers' memos analyses indicated that most learners in the interventionist group successfully minimized the repetition of some phonological or lexical errors for which they had received explicit mediations. Thus, the evidence suggests that the learners adopted memory strategies for storing and retrieving the correct forms of linguistic items and cognitive strategies for practicing them before each session.

The interactionist approach promoted metacognitive and affective strategies as subsets of indirect learning strategies more than the interventionist model. In the interactionist group, the instructor had more opportunities to encourage the learners to reflect on their speaking performance after finishing the tasks. The learners were prompted to explain how they could have done the task better. As a result, the learners benefitted from discussing their problematic areas, even those which did not necessarily stem from linguistic incompetency, especially regarding cohesion and coherence, grammar, and lexical resources.

For instance, some learners mentioned that their anxiety negatively affected their performance. The instructor and the learner focused on the possible roots of the learner's anxiety. During the discussions, some participants mentioned that they had perfectionism attitudes toward English learning and were obsessed with delivering error-free performance. The instructor provided why committing errors is necessary for learning a language as a skill and discussed the solutions to alter their perfectionist mindset, which promoted their affective learning strategies.

Some other participants believed that their performances were not satisfactory since they felt vulnerable when doing a speaking task in front of the instructor, a competent language user. The subsequent conversation between the instructor and the participant focused on self-esteem and how to boost it by any means they found more aligned with their style. The analyses indicate that these learners were more willing to take risks and provide more complex grammatical structure and vocabulary responses as the sessions progressed.

Moreover, regarding metacognitive strategies, the instructor's mediations based on the learners' ZPD included the introduction of learning resources such as the collocation dictionary and thesaurus for lexical resources and text-to-speech tools for pronunciation, to name a few. The participants of the interactionist group had to find out how to take advantage of the resources on their own and discuss the outcome with the instructor. As a result of the personalization of the efficient use of the out-of-class learning

tools, the interactionist group participants managed to navigate their learning process in certain areas and conduct a more accurate evaluation of their progress. An excellent example of this was evident in the development of lexical resources of the participants. The personalized use of resources helped the participants expand and deepen their vocabulary knowledge by learning the words' connotations, collocations, and some morphological features through self-study and attempting to use them in class. For another, the external resources provided the learners with opportunities to evaluate their progress, evident in their satisfaction with their improved collocations and segmental and suprasegmental pronunciation elements.

Promoting different preparation orientations

The fourth emerging difference between the two models of the DA sheds light on the planning and preparation of the participants' orientations. The analyses of the participants' responses and interactions with the instructor showed that the interventionist group participants appeared more prepared with in-advance planning for each session. The frequent use of notes, sentence stems, and prefabricated expressions by the learners supported the idea that the learners' preparations were more focused and organized for successful task achievement purposes.

Unlike the interventionist group, the interactionist group had a different orientation toward planning and preparation, resulting from integrating teaching and assessment. The apparent frequent attempts to use uncommon vocabulary and more advanced grammatical structures indicated that the learners' preparations were more targeted at making progress than satisfactory task accomplishment. Thus, according to the data, the interventionist group participants had a product-oriented preparation, which helped them be more confident and fluent during the post-test compared to the interactionist group participants, who had developed a process-oriented mindset. While the process-oriented mindset contributed to the learner's intrinsic motivation to learn more items, it negatively affected their fluency. Their responses in the post-test entailed more hesitations, delays, and clarification requests.

Integration of quantitative and qualitative phases

Combining quantitative and qualitative findings revealed that interactionist and interventionist DA models favor different aspects of speaking subskills. The quantitative results did not demonstrate the interactionist DA's significant superiority regarding grammatical range and accuracy. However, the qualitative findings revealed that the degree of negotiated mediations allowed for more profound learning and error correction. The interactionist DA encouraged in-depth negotiations, which assisted students in identifying and correcting fossilized and interlingual errors. As evidenced by the high scores for this aspect, such interactions resulted in effective grammar implementation improvement. The quantitative data, on the other hand, supports the interventionist DA model for speaking fluency. When combined with qualitative data, it is clear that systematic feedback and standardization in the interventionist method promote fluency by encouraging learners to use reduction strategies, which might limit their lexical and grammatical range. As for pronunciation, the interactionist DA was practically more beneficial because of its flexible discourse, allowing a reflective focus on fossilized errors and an understanding of phonological rules and

patterns. Finally, while there was no statistically significant difference in vocabulary development across both models, qualitative findings revealed more. The interventionist model emphasizes procedural preparation, promoting pre-planned lesson strategies based on existing words and patterns. However, the interactionist model encourages a discovery-oriented approach, expanding learners' lexical resources by exploring unfamiliar vocabulary and their collocations. The findings show that both assessment models have strengths that target different aspects of EFL speaking skills. The choice of interactionist or interventionist DA models is determined by the specific learning outcomes desired for students' language development.

Discussion

The findings of the present investigation revealed that the two DA models differed in their impacts on speaking sub-skills. As for the grammatical range and accuracy, the integrated results of the present investigation were in line with those of previous studies comparing the effects of the two DA models like Ebrahimi (2015), who conducted a study to explore the influence of the interactionist and interventionist DA on the students' complexity, accuracy, and fluency (CAF) of oral production. Her results indicated that the interactionist group outperformed the interventionist group regarding complexity and accuracy components. Khoshsima and Farokhipours (2016) reported similar results, which revealed that the grammatical accuracy and complexity scores of the interactionist group were noticeably higher than those of the interventionist group. In addition, the results of the present study accorded with Safdari and Fathi (2020) and Malmeer and Zoghi (2014) concerning the positive impact of interactionist DA on EFL learners' grammatical competence. The flexibility of the interactionist model allows the participants to be more actively involved in the learning process and engage in mediation negotiation. The learners' involvement helps the instructor identify and address the learners' problematic areas in grammar, such as fossilized errors, negative L1 transfers, and inappropriate attitudes toward language learning. Thus, the instructor had ample opportunities to provide the learners with tailor-made tackling strategies (Lantolf & Poehner, 2011).

In terms of fluency, the findings of the present investigation revealed that the interactionist DA was more effective than the interventionist model. The interventionist DA participants gained higher scores in fluency by avoiding complicated linguistic items and utilizing the simplified versions of words and structures in their answers. This finding aligned with Nakatani's (2005) conclusion that strategies that align with the task objectives help learners maintain their fluency. In addition, the learners' orientation to their preparations and exercises before the class was more product-oriented, aiming at successful task achievement, with minimum errors and mediations, rather than improving their speaking skills. Such orientation not only boosted the student's speaking confidence but lowered their anxiety by receiving less corrective feedback (especially explicit ones) from the instructor. Since anxiety is a critical factor that negatively affects learners' speaking fluency (Dörnyei & Ryan, 2015), it could be argued that the interventionist DA positively promoted learners' fluency by reducing their anxiety. This explanation finds support from Zhang and Rahimi (2014) and Estaji and Farahanynia (2019), who reported that interventionist DA effectively lowered the learners' anxiety. However, the results

were inconsistent with Ebrahimi's (2015) and Safdari and Fathi's (2020), who reported that DA models did not significantly impact the speaking fluency of Iranian EFL learners. However, the results of this study were in line with Ghahderijani et al. (2021). They compared the effect of group-dynamic assessment (G-DA) and computerized-dynamic assessment (C-DA) on upper-intermediate Iranian EFL learners' speaking performance.

For pronunciation, the present study revealed that the interactionist DA mediations enabled the instructor not only to assist the learners with the correct pronunciation of the words but also to help them internalize the essential features of English pronunciation through personalized and explicit instructions, which has shown to be practically effective in improving the EFL learners' pronunciation skill (Elliott, 1997; Lee & Lyster, 2015; Saito, 2011). Such evidence supports Shafiee et al. (2018), who indicated that implementing interactionist DA was significantly effective in acquiring rhythm among Iranian EFL learners. Thus, the participants' internalized awareness of the English pronunciation governing rules and features was one reason for the interactionist DA's higher efficiency than the interventionist model.

Concerning lexical resources, while there were no significant differences in the effects of DA models, qualitatively speaking, the findings revealed that the interactionist and interventionist models impact the vocabulary component of speaking skills differently. The vocabulary knowledge categorization, according to Read (2000), consists of vocabulary breadth (the number of words in learners' repertoire) and vocabulary depth (the degree to which learners know about the words). The present study findings suggested that interactionist DA contributed to promoting the vocabulary depth of the learners, which is in line with Mirzaei et al. (2017) findings. As for the effect of interventionist DA on the learners' vocabulary breadth, the present study showed that the participants in the interventionist group tended to use memory strategies, which helped them not to repeat their recurring errors in the vocabulary component. Additionally, adopting the memory strategy facilitated the retrieval of the new and advanced vocabulary items they had prepared to use in their task responses. This finding was in accord with Pérez and Alvira (2017), who investigated vocabulary acquisition through memory strategies; their results indicated that all participants benefitted considerably from memory strategy, allowing them to acquire, memorize, and retrieve the meaning of new words. The results are comparable to those of Hessamy and Ghaderi (2014) and Marzban and Nafarzadehnafari (2018), who investigated the effectiveness of the interventionist DA on EFL learners' vocabulary learning. They reported that the experimental group participants who received DA-based mediations outperformed the control group who received non-DA corrections. For the interactionist DA, Gibbons' (2003) case study demonstrated that applying ZPD-based mediations improved the learners' understanding and contextualization of new vocabulary items. Ebadi et al. (2018) conducted a study to shed light on the vocabulary acquisition process and the promotion of vocabulary knowledge for novice EFL learners. In addition to the abovementioned investigations, the current study accorded with Hamavandi et al. (2017). Their studies supported the positive impact of DA on enhancing the vocabulary knowledge of Iranian EFL learners.

Conclusion and implications

The study found that both DA models positively impacted speaking proficiency. However, the interactionist DA model was more effective in improving speaking grammar and pronunciation, while the interventionist DA model was more effective in promoting fluency. The study also found that the learners' prioritization choices and task features influenced their fluency. The interventionist DA model led learners to use simplification and reduction strategies to achieve good fluency at the expense of grammar and pronunciation complexity. The study further revealed that vocabulary use in speaking did not significantly differ between the two models. However, the learners' orientations toward course preparations influenced their vocabulary knowledge progress differently.

The comparative findings of this study could assist language educators in choosing the DA model for their online classes that suits their course objectives. Therefore, depending on the speaking lesson's objectives, teachers can implement the according DA model. If the lesson aims to provide learners with opportunities to develop the complexity of their grammatical, phonological, and lexical knowledge, they can adopt the interactionist model. However, the interventionist DA model can be more advantageous for courses such as free discussion and topic-based conversations, where speaking fluency is a higher priority. In addition, interventionist DA can be used in speaking exam preparation courses to increase the test-taker's confidence and help them develop appropriate preparation approaches.

Limitations and suggestions for further research

Concerning the limitations and delimitations of the study, some suggestions are recommended for further research studies. First, to enhance the generalizability of the findings, in addition to having a bigger sample size, future investigations can benefit from a control group, which enhances the accuracy and generalizability of the quantitative results. The experiment was conducted on the Google Meet platform and in a one-to-one teaching context. Similar studies can be conducted in face-to-face, private, or group classes. Fourth, future investigations can enrich the literature by diversifying the participants. The main participants of the current study were all at the intermediate level of proficiency; further studies need to be carried out with participants in lower or upper levels of proficiency. In addition, the present study benefitted from the IELTS speaking test and its rubrics. Future studies can utilize other measurement instruments to provide comparison opportunities or instruments that include the components or subcomponents not included in the IELTS rubrics. Finally, the quantitative data analyses revealed no significant differences between interactionist and interventionist DA. Further research studies can be conducted with longer and more treatment sessions to examine DA models' impacts more accurately on vocabulary resources.

Abbreviations

DA	Dynamic assessment
EFL	English as a foreign language
ELT	English language teaching
ZPD	Zone of proximal development
TEFL	Teaching English as a foreign language
L2	Second language
CEFR	Common European framework of references for languages
MMR	Mixed-methods research
AL	Applied linguistics

IELTS International English language testing system
 TED-Ed Technology, entertainment, design, education
 VIF Variance inflation factor

Author contributions

The authors had equal contribution in the process of the paper. The authors read and approved the final manuscript.

Funding

This study received no funding.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethical approval and consent to participate

The study adhered to ethical principles, ensuring participant confidentiality and privacy throughout the research process, and informed consent was obtained from all participants.

Competing interests

The authors declare that they have no competing interests.

Received: 21 May 2023 Accepted: 15 November 2023

Published online: 08 April 2024

References

- Alavi, S., Kaivanpanah, S., & Shabani, K. (2012). Group dynamic assessment: An inventory of mediational strategies for teaching listening. *Teaching English as a Second Language Quarterly*, 3(4), 27–58. <https://doi.org/10.22099/jtls.2011.370>
- Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, 78(4), 465–483. <https://doi.org/10.2307/328585>
- Azizi, Z., & Khafaga, A. (2023). Scaffolding via group-dynamic assessment to positively affect motivation, learning anxiety, and willingness to communicate: A case study of high school students. *Journal of Psycholinguistic Research*. <https://doi.org/10.1007/s10936-023-09935-6>
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Campione, J. C., Brown, A. L., Ferrara, R. A., & Bryant, N. R. (1984). The zone of proximal development: Implications for individual differences and learning. *New Directions for Child and Adolescent Development*, 1984(23), 77–91. <https://doi.org/10.1002/cd.23219842308>
- Cheng, L. (2003). Looking at the impact of a public examination change on secondary classroom teaching: A Hong Kong case study. *The Journal of Classroom Interaction*, 38(1), 1–10.
- Cheng, L., & Curtis, A. (2004). Washback in language testing. In L. Cheng, Y. Watanabe, & A. Curtis (Eds.), *Washback in language testing: Research contexts and methods* (pp. 3–19). Taylor & Francis.
- Dincer, A. (2017). EFL learners' beliefs about speaking English and being a good speaker: A metaphor analysis. *Universal Journal of Educational Research*, 5(1), 104–112.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited* (1st ed.). Routledge. <https://doi.org/10.4324/9781315779553>
- Ebadi, S., Karimi, E., & Vakili, S. (2023). An exploration into EFL learners' perspectives on online computerized listening comprehension dynamic assessment. *Language Testing in Asia*. <https://doi.org/10.1186/s40468-023-00221-9>
- Ebadi, S., & Rahimi, M. (2019). Mediating EFL learners' academic writing skills in online dynamic assessment using Google Docs. *Computer Assisted Language Learning*, 32(5–6), 527–555. <https://doi.org/10.1080/09588221.2018.1527362>
- Ebadi, S., Weisi, H., Monkareisi, H., & Bahramlou, K. (2018). Exploring lexical inferencing as a vocabulary acquisition strategy through computerized dynamic assessment and static assessment. *Computer Assisted Language Learning*, 31(7), 790–817. <https://doi.org/10.1080/09588221.2018.1451344>
- Ebrahimi, E. (2015). The effect of dynamic assessment on complexity, accuracy, and fluency in EFL learners' oral production. *International Journal of Research Studies in Language Learning*, 4(3), 107–123. <https://doi.org/10.5861/ijrsl.2015.982>
- Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania*, 80(1), 95–108. <https://doi.org/10.2307/345983>
- Estaji, M., & Ameri, A. F. (2020). Dynamic assessment and its impact on pre-intermediate and high-intermediate EFL learners' grammar achievement. *Cogent Education*. <https://doi.org/10.1080/2331186x.2020.1740040>
- Estaji, M., & Farahanynia, M. (2019). The immediate and delayed effect of dynamic assessment approaches on EFL learners' oral narrative performance and anxiety. *Educational Assessment*, 24(2), 135–154. <https://doi.org/10.1080/10627197.2019.1578169>
- Færch, C., & Kasper, G. (1983). Plans and strategies in foreign language communication. In C. Færch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp. 20–60). Longman.

- Farangi, M. R., & Kheradmand Saadi, Z. (2017). Dynamic assessment or schema theory: The case of listening comprehension. *Cogent Education*, 4(1), 1–13. <https://doi.org/10.1080/2331186x.2017.1312078>
- Farrokh, P., & Rahmani, A. (2017). Dynamic assessment of writing ability in transcendence tasks based on Vygotskian perspective. *Asian-Pacific Journal of Second and Foreign Language Education*. <https://doi.org/10.1186/s40862-017-0033-z>
- Ferrara, R. A., Brown, A. L., & Campione, J. C. (1986). Children's learning and transfer of inductive reasoning rules: Studies of proximal development. *Child Development*, 57(5), 1087–1099. <https://doi.org/10.2307/113043>
- Feuerstein, R. S. (2000). Dynamic cognitive assessment and the instrumental enrichment program: Origins and development. In A. Kozulin & Y. Rand (Eds.), *Experience of mediated learning: An impact of Feuerstein's theory in education and psychology* (pp. 147–165). Pergamon Press.
- Feuerstein, R., Feuerstein, R. S., Falik, L., & Rand, Y. (2006). *Creating and enhancing cognitive modifiability: The Feuerstein Instrumental Enrichment Program, Part 1 Theoretical and conceptual foundations, Part 2, Practical applications of the Feuerstein Instrumental Enrichment Program*. ICELP Publications.
- Ghahderijani, B. H., Namaziandost, E., Tavakoli, M., Kumar, T., & Magizov, R. (2021). The comparative effect of group dynamic assessment (GDA) and computerized dynamic assessment (C-DA) on Iranian upper-intermediate EFL learners' speaking complexity, accuracy, and fluency (CAF). *Language Testing in Asia*. <https://doi.org/10.1186/s40468-021-00144-3>
- Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly*, 37(2), 247–273. <https://doi.org/10.2307/3588504>
- Hamavandi, M., Rezaei, M. J., & Mazdayasna, G. (2017). Dynamic assessment of morphological awareness in the EFL context. *Cogent Education*, 4(1), 1–14. <https://doi.org/10.1080/2331186x.2017.1324254>
- Harlen, W. (2002). *A Systematic review of the impact of summative assessment and tests on students' motivation for learning*. <https://dspace.stir.ac.uk/handle/1893/19607>
- Haywood, H. C., & Lidz, C. S. (2007). Dynamic assessment in practice: Clinical and educational applications. *Cambridge University Press*. <https://doi.org/10.1017/CBO9780511607516>
- Haywood, H. C., & Tzuril, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education*, 77(2), 40–63. https://doi.org/10.1207/s15327930pje7702_5
- Hessamy, G., & Ghaderi, E. (2014). The role of dynamic assessment in the vocabulary learning of Iranian EFL learners. *Procedia - Social and Behavioral Sciences*, 98, 645–652. <https://doi.org/10.1016/j.sbspro.2014.03.463>
- Hidri, S. (2014). Developing and evaluating a dynamic assessment of listening comprehension in an EFL context. *Language Testing in Asia*, 4, 4. <https://doi.org/10.1186/2229-0443-4-4>
- Jafary, M. R., Nordin, N., & Mohajeri, R. (2012). The effect of dynamic versus static assessment on syntactic development of Iranian college preparatory EFL learners. *English Language Teaching*. <https://doi.org/10.5539/elt.v5n7p149>
- Kao, Y. T. (2020). A comparison study of dynamic assessment and nondynamic assessment on EFL Chinese learners' speaking performance: Transfer of learning. *English Teaching & Learning*, 44(3), 255–275. <https://doi.org/10.1007/s42321-019-00042-1>
- Khoshsima, H., & Farokhipours, S. (2016). On the role of different models of dynamic assessment on promoting speaking. *International Journal of Humanities and Cultural Studies*, 2(4), 586–600.
- Lantolf, J. P. (2006). Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition*, 28(01), 67–109. <https://doi.org/10.1017/s0272263106060037>
- Lantolf, J. P., & Poehner, M. E. (2004). Dynamic assessment of L2 development: Bringing the past into the future. *Journal of Applied Linguistics*, 1(1), 49–72. <https://doi.org/10.1558/japl.1.1.49.55872>
- Lantolf, J. P., & Poehner, M. E. (2007). *Language proficiency or symbolic capability: A dialectical perspective*. The Pennsylvania State University.
- Lantolf, J. P., & Poehner, M. E. (2011). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research*, 15(1), 11–33. <https://doi.org/10.1177/1362168810383328>
- Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.
- Lee, A. H., & Lyster, R. (2015). The effects of corrective feedback on instructed L2 speech perception. *Studies in Second Language Acquisition*, 38(1), 35–64. <https://doi.org/10.1017/s0272263115000194>
- Leong, L. M., & Ahmadi, S. M. (2017). An analysis of factors influencing learners' English speaking skill. *International Journal of Research in English Education*, 2, 34–41.
- Lidz, C. S., & Gindis, B. (2003). Dynamic assessment of the evolving cognitive functions in children. In B. Gindis, S. M. Miller, A. Kozulin, & V. S. Ageyev (Eds.), *Vygotsky's educational theory in cultural context* (pp. 99–116). Cambridge University Press. <https://doi.org/10.1017/CBO9780511840975.007>
- Malmeer, E., & Zoghi, M. (2014). Dynamic assessment of grammar with different age groups. *Theory and Practice in Language Studies*, 4(8), 1707–1713. <https://doi.org/10.4304/tpls.4.8.1707-1713>
- Marzban, A., & Nafarzadehnafari, F. (2018). The effect of interventionist classroom vs MALL dynamic assessment on EFL learners' vocabulary learning. *Applied Linguistics Research Journal*, 2(3), 58–66. <https://doi.org/10.14744/alrj.2018.36844>
- Mirzaei, A., Shakibei, L., & Jafarpour, A. A. (2017). ZPD-based dynamic assessment and collaborative L2 vocabulary learning. *The Journal of Asia TEFL*, 14(1), 114–129. <https://doi.org/10.18823/asiatefl.2017.14.1.8.114>
- Naeini, J., & Duvall, E. (2012). Dynamic assessment and the impact on English language learners' reading comprehension performance. *Language Testing in Asia*. <https://doi.org/10.1186/2229-0443-2-2-22>
- Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *The Modern Language Journal*, 89(1), 76–91. <https://doi.org/10.1111/j.0026-7902.2005.00266.x>
- O'Loughlin, K. J. (2001). *The equivalence of direct and Semi-Direct speaking tests (studies in language testing)* (1st ed.). Cambridge University Press.
- Oxford, R. L. (2002). Language learning strategies in a nutshell: Update and ESL suggestions. In J. Richards & W. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 124–132). Cambridge University Press. <https://doi.org/10.1017/CBO9780511667190.018>

- Pérez, L. M., & Alvira, R. (2017). The acquisition of vocabulary through three memory strategies. *Colombian Applied Linguistics Journal*, 19(1), 103–116. <https://doi.org/10.14483/calj.v19n1.10032>
- Pishghadam, R., Barabadi, E., & Kamrood, A. M. (2011). The differing effect of computerized dynamic assessment of L2 reading comprehension on high and low achievers. *Journal of Language Teaching and Research*, 2(6), 1353. <https://doi.org/10.4304/jltr.2.6.1353-1358>
- Poehner, M. E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development (educational linguistics book 9)* (2008th ed.). Springer. <https://doi.org/10.1007/978-0-387-75775-9>
- Poehner, M. E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly*, 43(3), 471–491. <https://doi.org/10.1002/j.1545-7249.2009.tb00245.x>
- Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research*, 9(3), 233–265. <https://doi.org/10.1191/1362168805lr1660a>
- Poehner, M. E., van Compernelle, R. A., Esteve, O., & Lantolf, J. P. (2018). A Vygotskian developmental approach to second language education. *Journal of Cognitive Education and Psychology*, 17(3), 238–259. <https://doi.org/10.1891/1945-8959.17.3.238>
- Rashtchi, M. (2019). Scaffolding argumentative essay writing via reader-response approach: a case study. *Asian-Pacific Journal of Second and Foreign Language Education*. <https://doi.org/10.1186/s40862-019-0078-2>
- Read, J. (2000). Research on vocabulary acquisition and use. In: *Assessing Vocabulary* (Cambridge Language Assessment, pp. 38–73). Cambridge University Press. <https://doi.org/10.1017/CBO9780511732942.004>
- Riazi, A. M. (2016). The Routledge encyclopedia of research methods in applied linguistics. In *Routledge eBooks*. <https://doi.org/10.4324/9781315656762>
- Riazi, A. M., & Farsani, M. A. (2023). Mixed-methods research in applied linguistics: Charting the progress through the second decade of the twenty-first century. *Language Teaching*. <https://doi.org/10.1017/s0261444823000332>
- Riswanto, N., Teferi, H., & Ibrahim, K. A. (2023). Cultivating EFL learners' productive skills by employing dynamic and non-dynamic assessments: attitude in focus. *Language Testing in Asia*. <https://doi.org/10.1186/s40468-023-00228-2>
- Ritonga, M., Farhangi, F., & Ajanil, B. (2022). Interventionist vs interactionist models of dynamic assessment (DA) in the EFL classroom: impacts on speaking accuracy and fluency (SAF), foreign language classroom anxiety (FLCA), and foreign language learning motivation (FLLM). *Lang Testing in Asia*, 12, 43. <https://doi.org/10.1186/s40468-022-00195-0>
- Safdari, M., & Fathi, J. (2020). Investigating the role of dynamic assessment on speaking accuracy and fluency of pre-intermediate EFL learners. *Cogent Education*, 7(1), 1818924. <https://doi.org/10.1080/2331186x.2020.1818924>
- Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: An instructed SLA approach to L2 phonology. *Language Awareness*, 20(1), 45–59. <https://doi.org/10.1080/09658416.2010.540326>
- Shabani, K. (2012). Dynamic assessment of L2 learners' reading comprehension processes: A Vygotskian perspective. *Procedia - Social and Behavioral Sciences*, 32, 321–328. <https://doi.org/10.1016/j.sbspro.2012.01.047>
- Shafiee, S., Talakoob, F., & Fatahi, M. (2018). Effects of dynamic assessment on the acquisition of the rhythm of English: The case of EFL learners' attitudes. *International Journal of English Linguistics*, 8(5), 181–191. <https://doi.org/10.5539/ijel.v8n5p181>
- Sherkuziyeva, N., Gabidullina, F. I., Ibrahim, K. A. A., & Bayat, S. (2023). The comparative effect of computerized dynamic assessment and rater mediated assessment on EFL learners' oral proficiency, writing performance, and test anxiety. *Language Testing in Asia*, 13, 1. <https://doi.org/10.1186/s40468-023-00227-3>
- Suresh, K. (2011). An overview of randomization techniques: An unbiased assessment of outcome in clinical research. *Journal of Human Reproductive Sciences*, 4(1), 8–11. <https://doi.org/10.4103/0974-1208.82352>
- Tzurriel, D. (2012). Dynamic Assessment of Learning Potential. In *Springer eBooks* (pp. 235–255). https://doi.org/10.1007/978-94-007-4507-0_13
- Volante, L. (2004, September 25). *Teaching to the Test: What every Educator and Policy-Maker should know*. <https://eric.ed.gov/?id=EJ848235>
- Vygotsky, L. S. (1978). Interaction between learning and development. In M. Cole, V. Jolm-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: Development of higher psychological processes* (pp. 79–91). Harvard University Press. <https://doi.org/10.2307/j.ctvj9vz4.11>
- Williams, P. (2017). Assessing collaborative learning: Big data, analytics and university futures. *Assessment & Evaluation in Higher Education*, 42(6), 978–989. <https://doi.org/10.1080/02602938.2016.1216084>
- Yang, Y., & Qian, D. D. (2017). Assessing English reading comprehension by Chinese EFL learners in computerized dynamic assessment. *Language Testing in Asia*. <https://doi.org/10.1186/s40468-017-0042-3>
- Zhang, L. J., & Rahimi, M. (2014). EFL learners' anxiety level and their beliefs about corrective feedback in oral communication classes. *System*, 42, 429–439. <https://doi.org/10.1016/j.system.2014.01.012>

Publisher's Note

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