AJAL

The impact of dynamic assessment on tertiary EFL students' speaking skills

Prathana Siwathaworn
Graduate School, Chulalongkorn University, Thailand

Jirada Wudthayagorn
Language Institute, Chulalongkorn University, Thailand

This study discusses the use of dynamic assessment for pedagogical purposes. Conducting dynamic assessment (DA), grounded in Vygotsky's zone of proximal development (ZPD), we studied a group of Thai university students who had difficulties in speaking English. We adopted DA as an alternative assessment in the context of teaching English as a foreign language (EFL). This investigation is aimed at the students' potential to improve their speaking skills in the test task called elicited imitation (EI). EI, which prompted the students to repeat sentences, was used to target specific features of the students' English speaking in this study namely meaning, syntax, vocabulary, pronunciation, and fluency. Through the test-train-retest design, DA improved the students' English speaking. The findings of both qualitative and quantitative data indicated positive impacts of DA on the students. Data from retrospective interviews and diaries showed that DA brought about meaningful learning experiences. The students also exhibited positive attitudes toward DA. The findings suggest that DA can be used in a classroom practice to promote development in the students' English speaking ability.

Keywords: dynamic assessment; traditional assessment; speaking skills; elicited imitation; EFL students; Thailand

Introduction

In Thailand, over decades, a series of educational policies have pushed schoolteachers to implement a communicative learner-centred approach to teaching English in the classroom (Teng & Sinwongsuwat, 2015). Thai students' ability to speak English fluently is prioritized by policymakers across all levels of education. Thus, the progress of Thai students' speaking skills has become the main concern of the teachers who teach English in the classroom. However, there has been limited success, despite a great deal of effort and considerable resources put into English instruction (James, 2015). Khamkhien (2010) suggests that the EFL context where Thai students grow up is insufficiently supportive to allow them to speak English in everyday life and this may be one of the factors stifling their motivation for speaking English. The teachers' knowledge and skill in assessing speaking are also regarded as a challenge for administering constructive instruction and assessment (Khamkhien, 2010; Ratanapinyowong, Poonpon, & Honsa, 2007). This combination of factors indicates a need to investigate and improve the instruction and assessment of Thai students' speaking skills in the classroom context.

Standardized traditional assessment is the dominant method in most English language classrooms in Thailand (Jaturapitakkul, 2013). It mainly reports what and how much the students gain from prior teaching. However, teachers, especially those who teach remedial courses, may consider this information inadequate. Alternative assessments are needed to allow those teachers to gain diagnostic information about students who encounter learning difficulties, and to predict their students' future courses of action (Grigorenko & Sternberg, 1998; Tzuriel, 2000).

Sternberg et al. (2007) note that students who perform poorly on traditional assessments tend to be less competent in dealing with learning difficulties on their own. As a result, those who are unable to pass the test in an English course and have to repeat the course without knowing how to improve their poor performance may eventually become trapped in that course. These students need assistance in finding their way out of the trap; otherwise, they may lose their self-confidence or even internalize negative perceptions toward their own English abilities.

Abbott, Reed, Abbott, and Berninger (1997) claimed that dynamic assessment (DA) lends itself to understanding the poor performance of these kinds of students. In using DA, students who would potentially get stuck in the test can be equipped with direct instruction, guided practice, and effective strategies to cope with their difficulties while taking the test. With this cooperative and responsive orientation, DA has been proved to be an effective tool for working with such students (Kozulin, 2001; Tzuriel, 2000).

The students who are the focus of the study reported here were EFL students with low English proficiency who performed poorly in traditional assessments. The aim of the study was to evaluate the potential of DA to help these students improve their English speaking skills. It is important to note that the use of DA is not common in Thailand.

Literature review

Lev Vygotsky's sociocultural theory is appealing to numerous scholars and practitioners of language assessment who are interested in students' learning potentials. He proposed that teachers should take their students forward through a social process of co-construction in the Zone of Proximal Development (ZPD) which refers to the area just beyond what students can achieve independently. Within this area, it is possible for students to accomplish a task when their level of potential development is stretched to the extent to which they can move away from their existing performance level toward the possible higher level in a social setting with more proficient people (Bavali, Yamini, & Sadighi, 2011; Lantolf & Poehner, 2004). Vygotsky's concept of ZPD has been extensively adopted by language testing researchers to probe an individual learner's ongoing learning and cognitive development. It is also a theoretical foundation of DA, which is the main focus of this study.

Dynamic assessment (DA)

DA is unique in its ability to look not only backward, but also forward (Sternberg, 2000) in a process-oriented approach of assessment which is sometimes called an assessment of learning potential (Dörfler, Golke, & Artelt, 2009). Its purpose is to bring about improved performance through the provision of the examiner's assistance (or mediation) during the test. It is claimed by DA proponents (see, among others, Kozulin & Garb, 2004; Leung, 2007; Lidz, 1991; Poehner & Lantolf, 2010) that the students' responsiveness to the mediation is an indication of what they will be able to do alone in

the future. The mediation includes hints, reminders, leading questions and other guidance to draw an individual student's conscious attention to the test task at hand. The student is encouraged to actively collaborate and interact with the examiner who is also the mediator during the course of the test task.

DA is also used to cause a positive change in the student by redirecting and reorienting his/her ability to learn while implementing the task. As an integration of instruction and assessment, it provides instructionally relevant information about the students' underlying ability to enable teachers to understand their students' poor performance and make a plan to tackle those aspects within the students' reach (Lantolf & Poehner, 2011; Roehr & Ganem-Gutiérrez, 2013).

DA versus traditional assessment

DA has been widely used in educational psychology and second language pedagogy to compensate for what traditional assessment lacks. Thus, it complements rather than replaces traditional tests (Heritage, Walqui, & Linquanti, 2015; Nazari, 2012), as summarised in Table 1. DA is more concerned with individual empowerment to promote students' engagement in the assessment process. Traditional assessment prioritises consistency of test scores because its goal is a summative product used to compare individuals for the purpose of classification. DA serves the purpose of this study because it facilitates the investigation of the students' problems and their learning process on an individual basis.

Table 1. The comparison of traditional assessment and dynamic assessment (based on Baek & Kyoung, 2003; Haywood & Lidz, 2007)

	Traditional assessment	Dynamic assessment
Focus	On product	On process
Examining process	Decontextualized, objective, and standardized	Contextualized, interactive, and individualized
Role of examiner	Neutral observer	Participant
Orientation	Retrospective approach	Prospective approach
Interpretation of results	Limit on performance	Learning potential (obstacles and way to overcome them)

DA and classroom practice

According to van der Veen, Dobber, and van Oers (2016), although DA has great potential for classroom practices, it has not been carried out in many classrooms yet. There are a number of reasons for this. Firstly, DA originated in the fields of psychology and psycholinguistics, so there are difficulties in transferring the clinical context of DA to the classroom context. Secondly, there is a difficulty with the scale of DA practice. DA practitioners typically concentrate on an individual's modifiability through problem-solving activities, while in the classroom context teachers have to be responsible for teaching to a large class. Finally, there is a worry that the procedure of DA may challenge the fairness and reliability of the test (Haywood & Lidz, 2007).

However, it has been argued that these potential problems relate only to the question of where to place the priority (Mehri & Amerian, 2015; Poehner, 2005; Weir, 2005). In particular, supporters of DA emphasise its concern with "the whole individual rather than simply the individual's scores on a test" (Murphy, 2011, p. 194) and also point out that, despite their reliability, the scores of a standardized test may not disclose valuable information about the examinees' real potential (Lauchlan, 2012) DA is found to be accurate when the objective of the examiner is primarily on enhancing the examinees' learning ability rather than making a decision based on the reported scores.

DA and speaking tests

Underhill (1987) pointed out that speaking tests should offer examinees an opportunity to behave as individuals. O'Sullivan's (2000) has identified three features that influence individuals' language ability in their test performance: physical/physiological characteristics; psychological characteristics; and experiential characteristics. The first set of characteristics includes special needs that require special measurements, accommodations, or modification for examinees' physical illness or disabilities. The second set includes examinees' interests, emotional stage, motivation, learning strategies, and learning styles. The third set consists of external influences like former education, examination preparedness, examination experience, and communication experience and other similar external influences that can have significant effects on examinees' ability to speak. DA is equipped to support all three sets of characteristics and, thus, can be integrated into speaking tests in order to optimize the speaking skills of the examinees.

Elicited Imitation as a speaking test task and its measurement

Elicited imitation (EI) is a language test task that has been extensively used to examine second language oral proficiency and development. The test method of EI requires that in order to remember and successfully imitate a sentence, the examinees should organize it in some manner (Hamayan, Saegert, & Larudee, 1977). The EI task, which takes the form of sentence repetition, does not target pragmatic knowledge. Instead, it focuses on how the examinees understand the meaning and use their internalized oral grammar, word knowledge, and pronunciation abilities to reconstruct the sentence (Burger & Chrétien, 2001). Brown (2004) added that EI elicits not only phonological ability but also discourse and overall oral production ability. Thus it is possible to employ EI in order to tap into the targeted components of the participants' speaking skills, which are meaning, syntax, vocabulary, pronunciation, and fluency. Since the EI task of this study was classroom-based, it was mandatory to include these five components to be relevant to the course content. In terms of parroting, to ensure that participants' responses were not a result of rote memorisation, Erlam's (2006) means of inserting a 5-second interval between listening and repeating in each sentence was used. Due to limited space in this paper, the explanation of how each component was measured is shown in Appendix 1.

Research questions

The current study is designed to answer the following research questions:
1. To what extent does DA assist EFL undergraduate students to improve their speaking skills?
2. What are the students' attitudes toward DA?

Methodology

Participants

Ten participants (both male and female) were selected through purposive sampling. They are referred to in this paper by pseudonyms to protect their confidentiality. None of the participants had been abroad and had studied English exclusively in an EFL classroom. The participants were first-year university students all members of the same class of a course called *Foundation English I* in a public university in Thailand. The range of their TOEIC scores was 285-315, indicating a level of basic users (or waystage users). Their levels of speaking abilities were measured through individual oral interviews. The interviewer was a native speaker of English who was also a university instructor. He has had more than ten years of English teaching experience in EFL contexts. The results of the interview showed that eight participants were at A1 level, and two participants were at A2. Before training them for the speaking task in DA sessions, they were provided with information about the objectives, processes and time frame of the research, and their right to withdraw from the study. They all signed consent forms.

Instruments

Both qualitative and quantitative research instruments were adopted. Qualitative instruments consisted of stimulated recall, retrospective interview and participants' diaries. The latter consisted of a rating of test scores. Participants' speaking performance in six DA sessions was video recorded (with their permission) speaking performance. Individual participants watched the video of themselves dealing with the task immediately after finishing it so that they could follow their thinking process and verbalize it. Immediately after each participant's self-report a retrospective interview was conducted to gather further information based on the self-report. Most of the questions were open-ended and flexible. Example questions are: "what are you doing/ thinking at this point?", "how do you feel when you say that word?" and "why are you moving your finger?" This verbal report also gave the participants a chance to self-evaluate. The diaries were cross-referenced with other data to enable triangulation.

The quantitative instruments consisted of 3 tests containing parallel sets of 15 EI sentences that covered the language focus taught in the language course. The tests were administered as a pretest (at the beginning of the course), a posttest two weeks before the end of the course), and a delayed posttest (at the end of the course). The EI sentences were recorded prompts which participants repeated after a five-second pause. All sentences were spoken by a native speaker at a natural speed, and they were in an increasing order of difficulty. They ranged in length from 6 words to 15 words (see Table 2 for examples of the EI sentences). It must be acknowledged that with such a small number of participants statistical analyses of quantitative data will not be conclusive but may be indicative of trends.

Table 2. Examples of EI sentences

- These new teachers aren't from Russia.
- That's seven euros and twenty-five cents altogether.
- Could I have the sandwich but no apple juice, please?
- My brother doesn't play all kinds of games on the Internet.
- There isn't a living room but there is a bathroom in this apartment.
- I haven't got my own office and my manager isn't friendly.
- I have breakfast with my roommate four or five times a month.
- The post office is opposite my school and on the left of the museum.

Data collection

Data collection was integrated into all the stages of the course (see Figure 1) which consisted of an initial training in which participants learned about the goals of the activity, the provision of mediation, the examiner's role and their role as an examinee in DA sessions. This was followed by a mock DA session and then a verbal report and diary training based on the video recording of that session. As the course progressed each participant accomplished six weekly DA sessions individually (each lasting about 45 minutes) which were video recorded and followed by a verbal report and a retrospective interview. Each participant also wrote a diary entry on the same day.
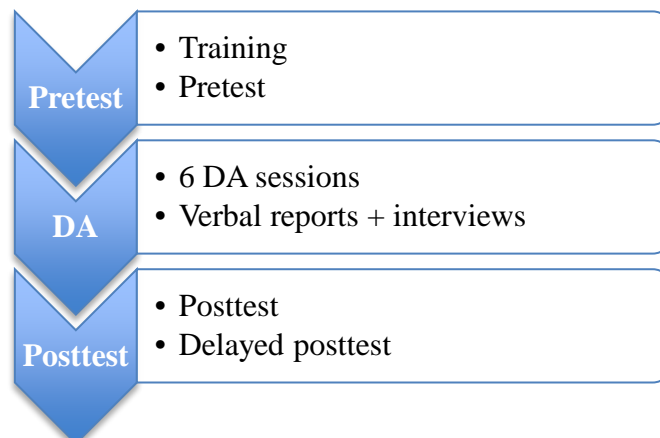

Pretest
- Training
- Pretest

DA
- 6 DA sessions
- Verbal reports + interviews

Posttest
- Posttest
- Delayed posttest

Figure 1. The stages of data collection

The mediation

Being aware of the drawbacks of subjectivity in DA research, the mediation guidelines of previous research were adopted. These provided a systematic approach in a form of scripted prompts with increasing explicitness (see Table 3). While following the prompts, the mediator retained some flexibility to vary the mediation according to different needs, responses, or speaking errors of the participants. The participants also had the opportunity in the first prompt to think, notice, and identify the problem by themselves in order to correct the error before receiving explicit explanation.

Table 3. Mediation prompts (adapted from van Compernolle & Zhang, 2014)

Sequence	Mediation prompt for the EI task
1	Shaking head to show rejection, saying "try again," Replaying the item
2	Giving the first hint (by naming the source of problem e.g. sentence structure, pronunciation, vocabulary, meaning), Replaying the item
3	Giving the second hint (more explicit than the second prompt), Replaying the item
4	Correct response and explanation provided

Scoring

The participants were informed that the scores of all tests in this study were not for evaluative purposes. They were rather used to gain insights into the participants' speaking skills. Two raters conducted the rating of test scores. One of the two raters was the researcher and the other was an instructor of the course. The findings of an earlier pilot study were used to check the practicality of the rating scales and to moderate inter-rater agreement. The analytic scoring method adapted from Gaillard (2014) was employed. The scoring rubric consisted of five criteria: meaning, syntax, vocabulary, pronunciation, and fluency. It is a 4-point rubric. Each scale contained a specific descriptor.

Data analysis

The qualitative data were analysed using a thematic analysis. The data of participants' self-reports and diaries, as well as retrospective interviews were coded and categorized to find changes in the participants' speaking skills, and to diagnose their strengths and weaknesses. Additionally, descriptive statistics was calculated for the test scores.

Results and discussion

Research question 1: To what extent does DA assist EFL undergraduate students to improve their speaking skill?

Analysis of the quantitative data shows improvement in speaking across the pretest, posttest, and delayed posttest (Figure 2). The mean score rose from 2.64 (SD = 1.15) in the pretest to a mean score of 4.83 (SD = 1.96) in the posttest and then to a mean score of 7.57 (SD = 2.37) in the delayed posttest. Regarding the individual scoring, each participant achieved higher scores in every test. While statistical measurement is not reliable with such a small number of participants, the results indicate a general ongoing improvement of the participants' independent performance before and after the DA sessions which suggests the intervention of DA caused positive changes in participants' independent performance. The continued improvement within the delayed posttest scores reflects potential sustainability of the improvement. These findings are in line with previous studies using DA to improve English students in other skills.
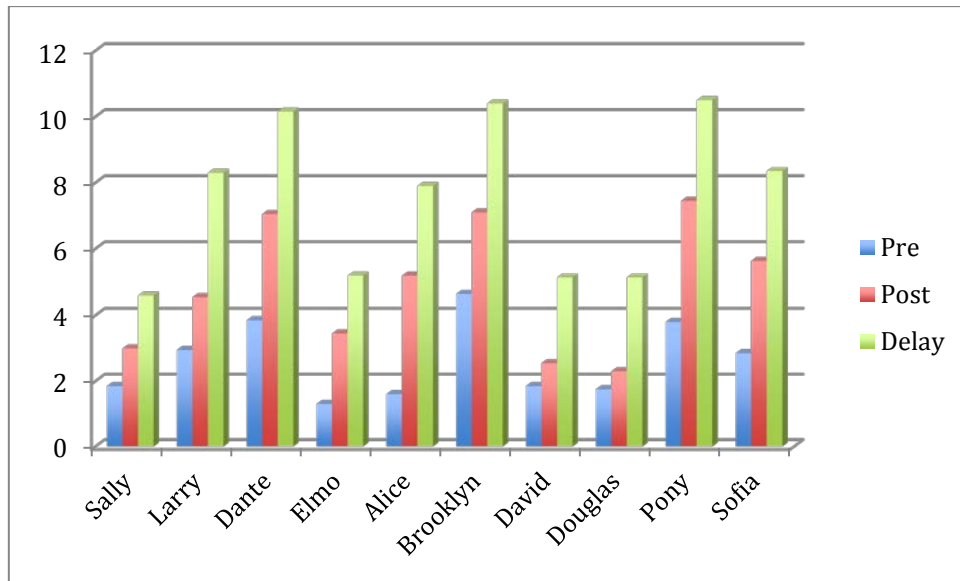
Figure 2. Differences of individual participant's scores of pretest, posttest, and delayed posttest

The participants' self-evaluation

Immediately after each DA session participants was asked to recall what they had done and to evaluate themselves in the diary. A summary of their self-evaluations is shown in Table 4. These reveal the shift of their self-perceived level of achievement. There is some evidence of more positive self-evaluations in some sessions. In the first session, only two of the participants felt positive about their work, but in the following sessions most participants seemed to be more satisfied with their speaking. It is interesting to see that no one rated themselves as "poor in DA session 3 and 4. However, following the last two DA sessions, many participants them seemed to become more dissatisfied with their work and so they thought that they did poorly. The participants explained in their diaries that they found the EI sentences in later sessions were longer and more complicated. Another reason was that these later sessions were closer to the final exam dates. Many of the participants easily lost concentration due to their concern about the study schedule for the final examinations.

Table 4. The participants' self-evaluation of their speaking in DA sessions

Session	Poor	Not so good	Good	Very good	n
DA#1	3	5	1	1	10
DA#2	2	4	4	-	10
DA#3	-	4	6	-	10
DA#4	-	7	2	1	10
DA#5	5	3	2	-	10
DA#6	4	4	1	1	10

Note: n refers to the number of the participants

The following excerpts from participants' diary entries show how some participants talked to themselves when their performance was poor:

Excerpt 1: Sally
At least, my spoken words will let me know how close I am to the original ones. This helps me guess the meaning of the sentence. I should neither be shy nor fear.

Excerpt 2: Elmo
I am very poor in communicating with foreigners. I have to pay more attention to the lessons in the classroom.

Excerpt 3: David
I must be brave and dare to speak. I have to grasp the meaning. I have to encourage myself to speak even though I know I will make a lot of mistakes.

These excerpts illustrate how the participants told themselves what to do in order to deal with their areas of weaknesses. This is related to the concept of self-directed learning which also arose in interviews where, for example, one participant clearly explained that he was aware that the responsibility to solve the problems lay with him. He reported that it changed the way he learned in the classroom. His learning objectives became clear. He knew what he wanted from the teacher. He understood that if he lets his teacher learn about this, the right kind of support can be given to him.

The mediation and the participants' learning
The mediation provided during DA sessions played an important role in assisting students who were at a basic level of English to improve their speaking. The participants stated that the test design pushed them to actually speak English face-to-face with the mediator for which they had to overcome their fear or shyness. Excerpts 4 and 5, which were from Alice's diary, illustrate this aspect:

Excerpt 4: Alice –following DA#1
I was very afraid that I would say something stupid. I was embarrassed. I got the words in my mind but I think it was wrong.

Excerpt 5: Alice –following DA#4
This activity brings me more confident to speak. I dare to take risk although I know it could be a wrong sentence.

This example shows the changes in Alice's perception of her way of speaking. It appears that she gained a more optimistic view of her own speaking ability.

The participants reflected that their learning was based on the gentle support that they got from the mediator. They said the mediator communicated with a non-judgmental attitude. This led to the feeling of trust that they had toward the mediator. The participants also felt comfortable providing more information about themselves that

enabled the mediator to understand them better. Douglas' account about his English accent was an example of a personal story that a participant shared with his mediator in his diary:

> Excerpt 6: Douglas
> My English accent is ugly. My friends often laugh at me when I speak. Because of this, I don't want to speak English.

After receiving this information, the mediator paid more attention to raising Douglas' confidence in his pronunciation and how to make it intelligible.

Research question 2: What are the students' attitudes toward DA?

The participants reported in their interviews that the experiences gained from participating in DA sessions motivated them to expose themselves to English. Some participants felt that the biggest challenge was the fact that the native speaker's pronunciation was very different from that of Thai people. Needing to overcome this challenge, one participant started watching Hollywood films without Thai dubbing, but with English subtitles. Another participant downloaded a singing application on his mobile phone to practice singing English songs along with the western singers. One participant asked her friend to help her do the mock EI test before coming to the DA sessions. Many of them recognized that the sentences in the EI test were taken from the English lessons in the classroom. Thus they paid more attention to the language focus and the example sentences in each lesson. One participant tried to answer the teacher in the class as often as he could. He stated that this is the way to increase the chance of making progress in his DA sessions.

The more the participants attended DA sessions, the more they got involved and wanted to achieve the goal of repeating the whole sentence successfully. They said that the interaction in the mediation was responsive and centred on their personal needs. This brought about an increase in self-esteem. They felt good about themselves even when they made a mistake. They stated that the important thing was that they had a chance to correct their mistakes with individual support. Therefore, they were not afraid or embarrassed of speaking incorrect English sentences in DA sessions.

Many participants reported that they liked to come to the DA sessions. A small competition occurred among them. This showed that their involvement became intense. A few of them could repeat sentences correctly. One participant stated that achieving a correct sentence was like getting a hard-to-win award. Once he achieved it, he felt that he was the winner. Thanks to this success, he felt good about himself. Another participant who could also make a complete sentence reported that he felt he was more proficient in English speaking than he had expected.

Conclusion

This study explored the impact of DA on university EFL students who had difficulties in speaking English. This is a small-scale study and its results need to be verified by other research. However, the study is significant because the use of DA in EFL speaking classes in Thailand has not previously been attempted. The results show that DA assisted the participants in improving their speaking skills in various ways. The students seemed to build up a sense of ownership in their own learning and they

purposively and actively became engaged in the test tasks. Their attempts to increase their opportunities to learn English both inside and outside the classroom demonstrated they were developing as self-directed learners. They tried to play their own part in the learning process in order to achieve their own goals. The participants developed and retained a positive attitude toward DA. This suggests that DA has promising potential as a classroom practice, especially for low proficient students in a remedial classroom.

About the authors
Ms. Prathana Siwathaworn is a Ph.D. candidate at Chulalongkorn University. She is an English language lecturer working in the Department of Foreign Languages, Faculty of Humanities, Kasetsart University. Her research interest is in sociocultural assessment.

Dr Jirada Wudhayagorn is an assistant professor of English at Chulalongkorn University Language Institute, and head of the Language Assessment and Evaluation Track of English as an International Language (EIL) program at Chulalongkorn Graduate School. Her research interest is in language policy, and language assessment and evaluation.

References

Abbott, S. P., Reed, E., Abbott, R. D., & Berninger, V. W. (1997). Year-long balanced reading/writing tutorial: A design experiment used for dynamic assessment. *Learning Disability Quarterly, 20*(3), 249-263

Baek, S.-G., & Kyoung, J. K. (2003). The effect of dynamic assessment based instruction on children's learning. *Asia Pacific Education Review, 4*(2), 189-198

Bavali, M., Yamini, M., & Sadighi, F. (2011). Dynamic assessment in perspective: Demarcating dynamic and non-dynamic boundaries. *Journal of Language Teaching and Research, 2*(4), 895-902

Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Boston: Allyn & Bacon.

Burger, S., & Chrétien, M. (2001). The development of oral production in content-based second language courses at the University of Ottawa. *Canadian Modern Language Review, 58*(1), 84-102

Dörfler, T., Golke, S., & Artelt, C. (2009). Dynamic assessment and its potential for the assessment of reading competence. *Studies in Educational Evaluation, 35*(2), 77-82

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics, 27*(3), 464-491

Gaillard, S. (2014). *The Elicited Imitation Task as a method for French proficiency assessment in institutional and research settings.* (Unpublished doctoral dissertation), University of Illinois, Urbana-Champaign, Illinois.

Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin, 124*(1), 75-111

Hamayan, E., Saegert, J., & Larudee, P. (1977). Elicited imitation in second language learners. *Language and Speech, 20*(1), 86-97

Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. Cambridge: Cambridge University Press.

Heritage, M., Walqui, A., & Linquanti, R. (2015). *English language learners and the new standards: Developing language, content knowledge, and analytical practices in the classroom*. Cambridge, MA: Harvard Education Press.

James, B. (2015, November, 6). Thailand's English skills lagging, says training company. *The Bangkok Post*. Retrieved from https://www.bangkokpost.com/learning/learning-news/756536/thai-english-proficiency-drops-now-3rd-worst-in-asia-ef

Jaturapitakkul, N. (2013). Students' perceptions of traditional English language testing in Thailand. *Academic Journal of Interdisciplinary Studies, 2*(3), 445

Khamkhien, A. (2010). Teaching English speaking and English speaking tests in the Thai context: A reflection from Thai perspective. *English Language Teaching, 3*(1), 184

Kozulin, A. (2001). *Mediated learning experience and cultural diversity* Paper presented at the Unlocking The Human Potential Conference (August 18-20, 2001), Royal Crown Conference Centre, Winnipeg, Manitoba, Canada. Retrieved from http://www.umanitoba.ca/unevoc/conference/papers/kozulin.pdf

Kozulin, A., & Garb, E. (2004). Dynamic assessment of literacy: English as a third language. *European Journal of Psychology of Education, 19*(1), 65-77

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

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

Lauchlan, F. (2012). Improving learning through dynamic assessment. *The Australian Educational and Developmental Psychologist, 29*(02), 95-106

Leung, C. (2007). Dynamic assessment: Assessment for and as teaching? *Language Assessment Quarterly, 4*(3), 257-278

Lidz, C. S. (1991). *Practitioner's guide to dynamic assessment*. New York: Guilford Press.

Mehri, E., & Amerian, M. (2015). Challenges to dynamic assessment in second language learning. *Theory and Practice in Language Studies, 5*(7), 1458-1466

Murphy, R. (2011). *Dynamic assessment, intelligence and measurement*. Chichester, West Sussex: Wiley-Blackwell.

Nazari, B. (2012). Teach-to-test instruction of dynamic assessment: A critical overview. *Bellaterra Journal of Teaching and Learning Language and Literature, 5*(4), 56-68

O'Sullivan, B. (2000). *Towards a model of performance in oral language testing.* (Unpublished Ph.D. thesis), University of Reading, Reading, UK.

Poehner, M. E. (2005). *Dynamic assessment of oral proficiency among advanced L2 learners of French.* (Unpublished doctoral dissertation), Pennsylvania State University, Pennsylvania.

Poehner, M. E., & Lantolf, J. P. (2010). Vygotsky's teaching-assessment dialectic and L2 education: The case for dynamic assessment. *Mind, Culture, and Activity, 17*(4), 312-330

Ratanapinyowong, P., Poonpon, K., & Honsa, S., Jnr. (2007). *Problems and solutions in teaching and assessing English skills in Thai higher education and the need for professional development* Paper presented at the Voices of Asia 2007 Symposium, MARA University of Technology, Malaysia.

Roehr, K., & Ganem-Gutiérrez, G. A. (Eds.). (2013). *The metalinguistic dimension in instructed second language learning*. London: Bloomsbury.

Sternberg, R. J. (2000). Prologue. In C. S. Lidz & J. G. Elliot (Eds.), *Dynamic assessment: Prevailing models and applications*. Amsterdam: JAI.

Sternberg, R. J., Grigorenko, E. L., Birney, D. P., Fredine, N., Jarvin, L., & Jeltova, I. (2007). *Dynamic instruction for and assessment of developing expertise in four ethnic groups*. Connecticut: National Research Center on the Gifted and Talented, University of Connecticut.

Teng, B., & Sinwongsuwat, K. (2015). Teaching and learning English in Thailand and the integration of Conversation Analysis (CA) into the classroom. *English Language Teaching, 8*(3), 13

Tzuriel, D. (2000). Dynamic assessment of young children: Educational and intervention perspectives. *Educational Psychology Review, 12*(4), 385-435

Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques*. Cambridge: Cambridge University Press.

van Compernolle, R. A., & Zhang, H. S. (2014). Dynamic assessment of elicited imitation: A case analysis of an advanced L2 English speaker. *Language Testing, 31*(4), 395–412

van der Veen, C., Dobber, M., & van Oers, B. (2016). Implementing dynamic assessment of vocabulary development as a trialogical learning process: A practice of teacher support in primary education schools. *Language Assessment Quarterly, 13*(4), 329-340

Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Basingstoke: Palgrave Macmillan.

Appendix 1: EI scoring rubric

This EI scoring rubric was adapted from Gaillard's (2014) rubric. A brief explanation of each criterion is provided below.

Meaning
The sentences contained varied content according to the lessons in the book. If two ideas were expressed in the sentence, but the participants failed to repeat one or both of them, then they did not succeed in demonstrating complete control of this criterion for the content of this message.

Syntax
The sentences in the test were designed to contain a particular syntax that reflected the course content. The grading focused on word order and grammatical category of the words that were arranged in the sentence.

Vocabulary
Specific vocabulary was used based on what the participants learned in the class. The participants' vocabulary knowledge was important for measuring the level of accuracy in their oral production.

Pronunciation
The grading focused on whether the examinees' pronunciation hindered their comprehension or not.

Fluency
The grading focused on the ease of production of the examinees, and on the occurrences of pause and hesitation.

Score	4	3	2	1	0
Meaning	This oral production expresses **exactly the same meaning** as the one in the original sentence.	This oral production expresses a **meaning very similar** to the one in the original sentence.	This oral production expresses a **meaning that is vague** and/or globally different from the one in the original sentence.	This oral production expresses the **beginning of a meaning sometimes different** from the one in the original sentence.	This oral production **does not express any meaning** corresponding to the one in the original sentence.
Syntax	This oral production contains **exactly the same syntactic structure** as the one in the original sentence and has **no syntactic mistakes**.	This oral production contains **the syntactic structures** copied the initial sentence with **only one syntactic mistake**.	This oral production contains **more than one/two syntactic structure(s)** more or less copied from the ones in the original sentence.	This oral production contains **one/two simple syntactic structure(s)** more or less copied from the ones in the original sentence.	This oral production contains **no syntactic structure.**

Vocabulary	This oral production contains **all the words** of the original sentence.	This oral production contains **the words** of the original sentence with only one **vocabulary mistake.**	This oral production contains **more than two words** of the original sentence.	This oral production contains **only one or two word(s)** of the original sentence.	This oral production contains **none of the words o**f the original sentence.
Pronunciation	This oral production is **perfectly intelligible** and perfectly copied from the original sentence **without any prosodic or segmental mistake.**	This oral production contains **prosodic and/or segmental elements** copied from the original sentence. There is **only one/two*mistake(s).** *Clearly intelligible, not hinder comprehension despite small articulatory errors or hesitation*	This oral production contains **more than two*prosodic and/or segmental elements** more or less copied from the original sentence. *In the best case, **half** of the elements **is present**.*	This oral production contains **only one/ two pro-sodic and/or segmental elements** more or less copied from the original sentence. *A lot of difficulty understanding the sentence. The repeated words are difficult to understand, due to poor phonemic articulation.*	This oral production **is not understandable** *The articulated phonemes **do not correspond** to the English phonological system **at all**.*
Fluency	This oral production copied from the initial sentence is expressed **with ease and no one hesitation nor pause.**	This oral production copied from the initial sentence is expressed **with ease and only one/two* hesitation(s) and/or pause(s)** or a missing word**. There is no break** in the sentence continuity. *The speech rhythm is slower, slightly more segmented than the one in the original sentence. The speed is not 'normal'*	This oral production, more or less copied from the initial sentence is expressed **with some ease but with a lot of breaks** in the sentence continuity (pauses and/or hesitations and/or missing words are present).	This oral production more or less copied from the initial sentence is expressed **with little ease and with a lot of breaks** in the sentence continuity (pauses and/or hesitations and/or missing words are present).	This oral production more or less copied from the initial sentence is expressed **with a lot of difficulties and has several breaks** in the sentence continuity(pauses and/or hesitations and/or onomato-poeias and/or English words insertion and/or missing words are present). *Nothing is clearly perceptible.*