

The Impact of Diagnostic Formative Assessment on Listening Comprehension Ability and Self-Regulation

Faezeh Ghazizadeh¹, Khali Motallebzadeh²

Received: 23 March 2017

Accepted: 23 August 2017

Abstract

Assessment plays a crucial role in any educational milieu (OECD, 2003). The researchers sought to investigate the possible impact of diagnostic formative assessment on students' listening comprehension ability and self-regulation. To this end, 46 participants were selected from among 78 students. A listening test was given to the participants to measure their level of listening comprehension ability. After 14 sessions (7 weeks) of instructing listening materials and experiencing formative assessment principles, the same listening test was administered as the posttest. Gathered data were analyzed and the results revealed that students in the experimental group outperformed those in the control group regarding their performances in listening test. Moreover, the results proved that students' self-regulation increased while they practiced diagnostic formative assessment based on the results obtained from pretest and post test. The results of the present study may be useful for those teachers who practice language for the sake of communication and forming students' knowledge by diagnosing their weak and strong points.

Keywords: *Diagnostic Formative Assessment, Listening Comprehension Ability, Self-Regulation, Classroom Assessment*

1. Introduction

There have always been some discussions regarding the true way to assess learners. In recent years much has been made of alternative forms of assessment. Learners get ready for the working life through active participation in assessment design, choices, criteria and making judgment. The methodology of evaluation is changing from a “culture of testing” to a “culture

¹ English Department, Islamic Azad University, Torbat-e-Heydareih Branch, Torbat-e-Heydareih, Iran.
(corresponding author) Email: faezeh.ghazizadeh.tefl@gmail.com

² English Department, Islamic Azad University, Torbat-e-Heydareih Branch, Torbat-e-Heydareih, Iran. Email: k.motallebz@iautorbat.ac.ir, kmotallebz@gmail.com

of assessment” (Birenbaum&Dochy, 1996;Sluijsmans et al., 2001, as cited in Birjandi&Siyyari, 2010).With the existence of educational assessment in contrary with psychometric testing, placing testing at the service of learning became one of the most important purposes to be followed in education (Gipps, 1994; Brown & Hudson, 1998; Lambert & Lines, 2000).The alternative means of assessment are the most effective ones among different methods and techniques through which the purposes of educational assessment could be achieved. The alternative means of assessment refer to the use of checklists, teacher observations, journals, videotapes, audiotapes, logs, conferences, portfolio, self-assessment, and peer-assessment (McKay, 2006; Brown & Hudson, 1998, 2002, as cited in Birjandi&Siyyari, 2010).Based on Brown et al. (1998), the alternative means of assessment is not intrusive and it provides the students with the chance of being assessed on everyday class activities. Also, the term “formative assessment” is an old term, but educational systems are trying to use it in more detailed and specific ways. Thus, there is a need for further research and theorizing on formative assessment (Black &Wiliam, 1998).Assessment is formative only if it leads to action by the teacher and students to increase students’ learning i.e. the wash back effect of formative assessment (Black, 2000). For instance, the outstanding characteristic of formative assessment is that the assessment results are used, by the teacher and students, to change or modify their work to make it better and more effective (Black, 2000).

Because summative assessment has not proved to be effective in discovering the learners’ areas of weaknesses, formative assessment is increasingly becoming the focus of concentration on educational assessment and in the professional improvement of teachers. Formative assessment is done via the teacher–student interactions during learning program (Newman, Griffin & Cole, 1989).Listening has also been an aspect of assessment that is under investigation. Based on the researches available in testing and assessment, the diagnostic assessment of listening is especially under-represented. Several studies have examined the diagnosis of listening post-hoc via the use of statistical models (e.g., Buck &Tatsuoka, 1998; Lee &Sawaki, 2009).

2. Literature Review

The researchers adopted a two-dimensional framework of formative assessment developed by Black and Wiliam (2009, cited in Huang, 2011) to organize the different dimensions of formative assessment. One dimension refers to the agent of learning, e.g. teacher, a peer, or the learner him/herself. The other dimension focuses upon the stages of learning, e.g. the goal -“where the learner is going”, the current status-“where the learner is right now”, and the bridge between the two-“how to get there.”

2.1. Diagnostic Formative Assessment

Formative assessment is defined as assessment for learning and not assessment of learning (Black, 1993). Formative assessment can possibly specifically enhance learning since it happens while direction is in advance and can serve as a reason for giving opportune input to build learner learning (Sadler, 1989). Besides, procedures that demonstrate to build learner learning ought to likewise improve the probability that learner learning will be reflected in all around planned substantial scale evaluations. Be that as it may, instructive specialists are just start to recognize powerful developmental evaluation rehearses and to specifically interface these practices to measures of learner learning (Black & Wiliam, 1998). Black (1993) emphasized that formative assessment is essential to effective teaching and learning. Formative assessment involves gathering, interpreting, and acting on information about students' learning so that it may be improved (Bell & Cowie, 2001). Classroom formative assessment can be seen as a continuum determined by the premeditation of the assessment moment, the formality of means used to make explicit what students know and can do, and the nature of the action taken by the teacher (the characteristics of the feedback). The continuum then goes from formal formative assessment in one end to informal formative assessment on the other (Bell & Cowie, 2001; Shavelson et al., 2003). Ross (2005) utilized a blended techniques approach, which included self-evaluation as a type of developmental appraisal, to study Japanese students (n = 2215) who were enlisted in a two-year, sixteen course English for scholarly purposes program. Examinations demonstrated that developmental appraisals yielded higher language capability development than the individuals who were evaluated by routine summative evaluations as it were. Ross' discoveries likewise uncovered that albeit developmental appraisal can create substantive increments in accomplishment and capability development, this effect might be space subordinate, e.g. language listening cognizance change.

2.2. Listening Comprehension Ability

It is important to recognize that listening is a distinct skill from reading. Listening involves real-time processing generally without the option of going back to earlier sections of the passage the listener may have missed (Flowerdew, 1994). In a conversation, listeners may be able to exert some control over the speech rate of their interlocutor, while listening to a radio program provides no opportunity for control over the speed of delivery, and attending a professional lecture does so only with certain restrictions and considerable effort. Aurally presented material also involves a number of phonological and lexical features that are not present in written material (ibid). Listening, which is defined as making sense of sounds heard

and reacting after this process, is the skill most used in the classroom environment (Taylor, as cited in Melanlioglu, 2013, p. 1177). This indicates that listening is as important as academic attitudes and reading skills for academic success (Conaway, 1982). For this reason, special importance is attached today to efforts oriented towards improving the listening skill, which is thought to have initially been ignored (Kline, 1996), in mother tongue curriculum. Although improving the listening skill has a significant part in curricula, it is a highly difficult process for teachers (Dawes, 2008), because numerous variables need to be taken into account while trying to improve this skill. Adding to the fact that learners recognize listening as the most difficult skill to learn, it is said that L2 listening remains the least researched of all four language skills (Vandergrift, as cited in Kurita, 2012, p. 31). In spite of being the least researched skill, L2 listening studies have addressed various issues; for example, cognitive issues such as bottom-up processes and top-down processes; linguistic issues such as linguistic factors that contribute to listening comprehension, for example lexis and phonology, and affective issues such as motivation and anxiety in listening have all been investigated (Kurita, 2012, p. 31).

2.3. Self-Regulation and Its Empirical Studies

According to Schwartz (2003) self-regulation comprises such processes as setting goals for learning, attending to and concentrating on instruction, using effective strategies to organize, code, and rehearse information to be remembered, establishing a productive managing time effectively, seeking assistance when needed, holding positive beliefs about one's capabilities, the value of learning, the factors learning and the anticipated outcomes of actions, and experiencing pride and satisfaction with one's Efforts. Based on Butler, Cartier, Schnellert, Gagnon, Giammarino(2011) self-regulated L2 learning strategies are defined as deliberate , Goal-directed attempts to manage and control efforts to learn the L2 these strategies are broad, teachable actions that learners choose from among alternatives and employ for L2 learning purposes (e.g., constructing, internalizing, storing, retrieving, and using information; completing short-term tasks; and/or developing L2 proficiency and self-efficacy in the long term).Also self-regulation refers to “the ability to flexibly activate, monitors, inhibits, persevere and/or adapt ones' behavior, attention, emotions and cognitive strategies in response to direction from internal cues, environmental stimuli and feedback from others, in an attempt to attain personally-relevant goals” (Moilanen, as cited in Heo,2014,p.93). Posner and Rothbart (2000) noted that self-regulation is the single most important factor in understanding human development. Shonkoff and Phillips (2002) also remarked that self-regulation connects to all aspects of adaptation in academic, social and career domains, as self-regulation denotes the alteration of one's behaviors and modulation of one's reactivity to the given environment. (As cited in Heo, 2014, p.76).Research has demonstrated that the self-regulation process takes place at both the unconscious and conscious level (Forgas, Baumeister&Tice, 2009).Obviously, self-regulation

has a voluntary effortful, motivated and conscious component. In terms of the, conscious component of self- regulation, scholars have used the term, *self-control* more frequently (Forgas et al., 2009).

According to Boekaerts and Corno (2005) all theorists assume that students who self-regulate their learning are engaged actively and constructively in a process of meaning generation and that they adapt their thoughts, feelings, and actions as needed to affect their learning and motivation. Similarly, models assume that biological, developmental, contextual, and individual difference constraints may all interfere with or support efforts at regulation. They believe that Theorists are in agreement that students have the capability to make use of standards to direct their learning, to set their own goals and sub-goals. Finally, all theorists assume that there are no direct linkages between achievement and personal or contextual characteristics; achievement effects are mediated by the self-regulatory activities that students engage to reach learning and performance goals. Self-regulation does not happen to learners, rather, it happens by them as they proactively monitor, regulate and control their thoughts, feelings, and behaviors with the objective to accomplish their goals (Cleary & Zimmerman, 2004).

Although it is argued that most of the learner's self- regulate their learning to some degree, the extent to which they consciously do so differentiate achievers from underachievers (Butler & Cartier; Randi; Zimmerman &Risemberg, as cited in Peeters, et al.2014, p.28). Teachers need to be able to learn in and from practice since the knowledge to teach can hardly be fully obtained before or apart from practice (Randi, 2004). They work in a rapidly changing environment and need to continuously update their teaching skills (Randi, Corno, & Johnson, as cited inPeeters, et al., 2014, p.32).

The present study seeks to address the following research questions:

Q1. Does diagnostic formative assessment have any significant effects on Iranian EFL learners listening comprehension ability?

Q2. Does diagnostic formative assessment have any significant effects on Iranian EFL learners' self-regulation?

Q3: What are participants' attitudes towards diagnostic formative assessment?

3. Methodology

3.1. Participants and Setting

Forty-six Iranian upper intermediate EFL learners who were chosen from among a population of EFL learners (about N=78) in three Language Institutes (Ayandesazan, Kian &Khorasan) in Torbat-e-Heydareih, Iran were selected. Sample selection was based on Quick

Placement Test (QPT). The participants' age ranged from 14 to 25 years. Only females took part in this study. These participants were randomly assigned to control (N=23) and experimental (N=23) groups.

3.2. Instrumentations

The following instruments were employed by the researchers to gather the necessary data.

3.2.1. Quick Placement Test (QPT)

This test was administered at the beginning of the study as a means for checking participants' homogeneity. Oxford University Press and University of Cambridge Local Examinations Syndicate (2001) developed the test containing 60 multiple-choice items; grammar, vocabulary, and cloze test. The test contains two parts. The first part is administered to all candidates and is aimed at those learners who are at or below intermediate level. The second part (items 41-60) is administered to those who score more than 31 out of 40 on the first part and can be used for those learners with higher ability. According to Geranpayeh (2003) QPT has been validated in 20 countries through administering to more than 2,000 learners. Furthermore, the reliability indices of the tests, which were calculated by the trial phases, were 0.9 for the 60 items test and 0.85 for the 40 item test.

3.2.2. Researcher-made Listening Comprehension Test

The test was developed by the researchers and was used as both pretest and posttest. The items were selected from the exercises of Tactics for Listening (Richards, 2010). It included 40 items (multiple-choice). Participants responded to the items in 40 minutes. The test reliability was checked through Cronbach's Alpha. The results revealed that the researcher-made listening comprehension test was reliable since Cronbach's Alpha was 0.72. The criterion-related validity of the test was also calculated and correlation index showed the test was valid ($r=.68$). The content validity of the test was also proved by two experts in Azad University of Torbat-e Heydarieh.

3.2.3. The Self-Regulation Questionnaire (SRQ)

In order to measure students' self-regulation, the participants of this study were given a 19-item questionnaire based on self-regulation learning. The validity of the test was reported by Bouffard, Boisvert, Vezeau and Larouche (1995) by Cronbach's coefficient Alfa. Also Joker (2001) by this way reported the validity of the test for 50 students of Institute as sample. To measure the reliability of the translated questionnaire, the researchers computed respondents'

answers and analyzed the obtained data via SPSS to ensure its reliability. The reliability index for the translated questionnaire was 0.903.

3.3. Procedure and Design

The first step of this study before commencing the formal experiment was selecting the participants. To achieve this end, 46 Iranian upper intermediate EFL learners, were selected. To make sure of participants' homogeneity, the quick placement test was administered. Based on the obtained results, they were randomly assigned to control (N=23) and experimental (N=23) groups. The two groups received the pretest, Listening and SRQ.

The participants were acquainted with the purpose of the study from the beginning. The goals and procedures of the study were presented to them. The same textbook – Tactics for Listening – was taught in both groups and they had English classes for the same amount of time. In both groups the same syllabus was used. The teacher of both groups was the researcher herself. The experiment lasted 14 sessions (7 weeks); it started on July 2th and finished on August 11th. At the outset of the semester, both groups had a pre-test aiming to test their listening comprehension and self-regulation for comparing participants' knowledge regarding their listening comprehension and self-regulation level. Since the subjects in both groups had no significant differences before the treatment, the formal experiment began.

Throughout the study which lasted for 14 sessions, twelve sessions for applying the experiments and two sessions devoted to pretest and posttest, participants in both groups received listening materials from Tactics for Listening (Richards, 2010). Both groups were taught by one EFL teacher twice a week (Mondays & Wednesdays). The techniques used were different in two groups. Experimental group received the treatment, diagnostic formative assessment. Accordingly, apart from their course materials, the participants in the experimental group received formative tests and quizzes, homework exercises, exercises with multiple-choice answers, fill-in-the blanks, and comprehension summary of an audio material. Throughout the instructional process the teachers/ researchers monitored students' progress and provided feedback on their strong and weak points. Feedback is the key element in formative assessment which allows students to correct conceptual errors and encourages teachers to modify instructional activities in light of their effectiveness. Hence, all the types of formative assessment applied during the experiment were always provided with feedback (by teacher, peer or self-assessment) and were discussed thoroughly with the students.

The main purpose of diagnostic formative assessment is to recognize the students' weaknesses and strengths. Participants' performance in each assessment was recorded and compared in pairs. Therefore, the researchers became aware of her students' progress. In addition to the assessment carried out in the experimental group, such activities as listening to the audio

materials, discussing, interpreting, answering to the teacher's questions were performed. The class time was divided into two parts. The first part was allocated for responding to questions of the previous lesson and practicing audio materials. Participants carried out the tasks within 45 minutes. In the second part of the class time (45 Min) these students were asked to answer to new audio materials of the new unit of the text-book. Participants were regularly assessed and this process informed them of the relation between the assessment and learning. Regular assessment required them to prepare themselves for receiving new approaches to listening and finding an effective way for analyzing what they heard.

To analyze the obtained data, the research employed qualitative and quantitative data. The qualitative data were obtained with the help of the semi-structured interview, in which the participants shared their attitudes towards the use of formative assessment in EFL listening comprehension; the quantitative data were obtained with the help of the pre-test and post-test results as well as the closed-ended items in the questionnaire.

Two variables were investigated in this study: independent and dependent. The independent variable of the study is the employment of diagnostic formative assessment applied to listening comprehension. The dependent variable is listening comprehension of the participants measured by pre- and post-test scores obtained from a researcher-made listening test. Moreover, the data gathered by means of self-regulation questionnaire were analyzed to check the effectiveness of diagnostic formative assessment on students' self-regulation. The results of the pretest and posttests were analyzed and tabulated with the help of the SPSS (21.00) program to address the research questions.

The researchers adopted a two-dimensional framework of formative assessment developed by Black and Wiliam (2009, cited in Huang, 2012) to organize the different dimensions of formative assessment. One dimension refers to the agent of learning, e.g. teacher, a peer, or the learner him/herself. The other dimension focuses upon the stages of learning, e.g. the goal — “where the learner is going”, the current status — “where the learner is right now”, and the bridge between the two — “how to get there.” Forged under Black and Wiliam's two dimensions, classroom learning based on formative assessment is believed to progress in the following temporal sequence (p. 28). First, considering the final goal of the learner, there are some steps to take into account which are as follows: the teacher explains about learning intentions and criteria for success, another student understands and shares learning intentions and criteria for success, and the learner himself understands learning intentions and criteria for success. The next step is the present status of the learner: the teacher manages effective classroom discussions and other learning tasks which extract evidence of student understanding and possible ways to get there. The third step is offering feedback which moves learners forward by the teacher. In the next step, other learners can help to activate students as instructional

resources for one another. Step five includes activating students as the owners of their own learning (learner) (Huang, 2012, p. 102).

The teacher followed the above framework in the experimental group throughout the study and checked the participants' gradual achievement each session. She provided the class with the opportunity to benefit from peer assessment, self-assessment, and teacher assessment. In addition, class discussion was held and effective feedbacks were provided for the participants. The students' errors were recognized and explained by the teacher and the students paid close attention to the oral materials given by the teacher. She sometimes gave them clear correct format and sometimes gave them some clues to let them correct themselves. To correct the errors, teacher-correction, peer-correction, and self-correction were used. Teacher corrected the errors both individually and collectively. During the process of self-correction and peer-correction, the teacher was responsible for supervising the task to prevent from subsequent problems.

Participants in the control group received the same materials as those practiced in the former group. However, they lacked any formative assessment. They were assessed on a summative form at the end of the final session of instruction. In other words, they were assessed via a linear and predictable system and the placebo included extra tasks without any feedbacks and diagnostic assessment, and teaching session lasted for a longer time compared to the experimental group. Listening to the audio materials via TV and CD player, answering to the teacher's questions, discussing about the materials, and note-taking were the activities which were used as teaching and learning tasks in control group. They were taught the same materials with traditional book exercises and activities (homework exercises, gap filling, matching, multiple-choice exercises). These types of activities were not provided with feedback.

4. Results and Discussion

4.1. Data Analysis

4.1. Data Analysis before the Experiment

Before commencing the experiment, the researchers analyzed the data to make sure of their normality. Table 1 shows the following data:

Table 1 Tests of Normality for Pretest and Posttest

	Groups	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Pretest	Experimental	.137	23	.200*	.965	23	.576
	Control	.157	23	.149	.907	23	.036

Posttest	Experimental	.135	23	.200*	.938	23	.162
	Control	.142	23	.200*	.939	23	.167

As seen in Table 1, since sig. (2-tailed) is higher than the significance level (sig. = 2) in both pretest and posttest, the data is considered to be normal. The test given to the participants showed to be both reliable (as mentioned in instrumentations section) and normal. An independent sample test was run to compare the performance of students in experimental and control groups.

Table 2 Independent Samples Test for the Pretest of Listening Comprehension Test

	Levene's Test for Equality of Variances		t-test for Equality of Means		
	F	Sig.	t	df	Sig. (2-tailed)
Pretest	2.05	.15	.62	44	.53

At the beginning of the study, after the participants were assigned to two groups, the researchers must check to make sure that they are at the same level regarding the variables of his research. Listening comprehension ability is going to be measured before and after the treatment and the performance of the participants are compared in each group. Table 2 shows that the performance of participants in experimental and control groups in pretest had no significant difference since Sig. (2-tailed) = .53 > .05.

Table 3 Tests of Normality for Self-Regulation Questionnaire in Pretest and Posttest

	Groups	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Pretest	control	.198	23	.120	.861	23	.094
	experimental	.126	23	.200*	.969	23	.662
posttest	control	.167	23	.094	.878	23	.079
	experimental	.086	23	.200*	.986	23	.979

To make sure that the data obtained through administering the questionnaire among the participants is normal in both pretest and posttest, a test of normality was run and as seen in Table 3, since Sig. is higher than 0.05 (.120, .200, .094, .979), there is no significant difference among the data entered in SPSS.

Table 4 Independent Samples Test for Pretest of Self-Regulation Questionnaire

	M	SD	Levene's Test for Equality of Variances		t-test for Equality of Means	
			Sig.	t	Sig. (2-tailed)	

Pretest	59.61	16.62	.54	.46	-.44	.66
	61.65	14.69			-.44	.66

To measure the possible difference between the mean scores of both groups in the first administration of the questionnaire, an independent sample test was run. Since the mean scores of both groups, they were very close to each other, and as shown in Table 4, Sig. (2-tailed) = .66 > .05, thus there is no significant difference between the performance of participants in the first administration of the questionnaire.

4.2. Data Analysis after the Experiment

Table 5 Independent Samples Test for Posttest of Listening Comprehension Test

	Levene's Test for Equality of Variances		t-test for Equality of Means		
	F	Sig.	t	df	Sig. (2-tailed)
Posttest	6.65	.013	5.23	44	.000

As seen in Table 5 Sig. (2-tailed) shows that the performance of students in the posttest had a significant difference from their performance in the posttest ($P=.000$).

Table 6 Group Statistics for Posttest of Self-Regulation Questionnaire

	Groups	N	M	SD	Sig. (2-tailed)
posttest	control	23	62.17	11.68	.006
	experimental	23	71.00	8.91	.006

Regarding students' answers to self-regulation questionnaire, the data in Table 6 were presented as the following: N=23, Experimental Group (N=71.00, SD=8.91) and Control Group (N=62.17, SD=11.68) and Sig. (2-tailed) = .006 which shows that their answers to the questionnaire in the first time had a significant difference from theirs in the second time. , thus the second null hypothesis was also rejected.

4.3. The Results of Semi-Structured Interview

In order to elicit participants' ideas about diagnostic formative assessment, the researchers asked three questions from the students in the experimental group. To this end, a three-minute interview was performed by the researchers and ten students in the experimental group answered their questions. They answered the questions and mostly had positive attitudes and considered it effective in the process of their learning.

To answer the first question of the interview, "Was diagnostic formative assessment effective in your learning?", nine out of ten interviewees said that "it was very effective since we could find our problems during taking listening tests and used more strategies taught by the

teacher during the sessions of instruction." One student talked about the exhausting process of taking tests during the period of teaching and refused the effectiveness of diagnostic formative assessment based on her reasons which will be presented in the following question.

In answer to the second question, "how did this type of assessment affect your learning of listening and self-regulation?", one student expressed that "it was a tool for learning rather than being a test or improvement in testing ability. It helped me to be aware of my weak points, in particular. It also made me aware of my strong points which I tried to focus and develop." To answer this question, another student said that "I learned to manage my learning when the teacher checked our progress session by session. At first, it was difficult for me to adapt myself with this method, but soon I found that this type of assessment increased my abilities in learning since I became aware of those parts which I did not learn or study completely. Moreover, after each test or quiz, I looked for my weak points to cover. To me, it was very effective because I could understand where in the lesson I have to focus on." On the other hand, one student in the experimental group had a very different idea. She said that "taking short tests and quizzes every session and asking to monitor our progress were much of a burden rather than the tasks I could perform. It was boring and frustrating for me to prepare myself for the test or quiz all the time. I preferred my teacher to tell me to study more rather than find it where I had to concentrate on."

In response to the third question of the interview dealing with "Will you apply this technique of assessment for better learning?" most of the students agreed and their responses were positive. They claimed that the most positive aspect of diagnostic formative assessment was its impact of knowing their weaknesses. Since it was a repetitive answer the researchers continued and asked "How will you apply it?" some of them said that they will try to make questions out of the content of their materials they try to cover and practice them in pairs or groups. For listening comprehension enhancement, some of them suggested the application of listening strategies and scoring their performances by a peer or identifying their errors after the test. Two of them had no idea about the application of diagnostic formative assessment and their self-study programs.

4.4. Results and Discussion

As mentioned at the beginning of this chapter, the researchers aimed at investigating the possible impact of applying diagnostic formative assessment on listening comprehension ability and self-regulation of Iranian EFL learners. To gain this end, she selected homogenized participants and administered a listening comprehension test and a self-regulation questionnaire. Then, the gathered data were analyzed to make sure that these participants were homogenized considering their performances in the test. Moreover, their ideas about applying diagnostic formative

assessment rules in the experimental group were drawn via a semi-structured interview. After working on their listening comprehension through sessions of instruction, the same test was administered to measure the possible impact of the instruction on their performances. As explained in the previous sections, the experimental group outperformed participants in the control group regarding their listening comprehension ability in the posttest. Considering students' self-regulation, the participants in the experimental group revealed to have a considerable progress after applying the principles of formative diagnostic assessment during working on listening comprehension tasks. Finally, participants in the experimental group agreed upon the usefulness and practicality of using diagnostic formative assessment for improving their listening comprehension ability and self-regulation.

Regarding formative assessment, Cizek, Fitzgerald and Rachor (1995) believed that teachers appear to be unaware of the assessment work of colleagues and do not trust or use their assessment results. In his paper Crook (1988) concluded that the summative function of evaluation--grading--has been too dominant and that more emphasis should be given to the potential of classroom assessments to assist learning. Feedback to students should focus on the task, should be given regularly and while still relevant, and should be specific to the task. However, in Crooks' view the 'most vital of all the messages emerging from this review' (p. 470) is that the assessments must emphasize the skills, knowledge and attitudes perceived to be most important, however difficult the technical problems that this may cause.

In a project, 25 Portuguese teachers of mathematics were trained in self-assessment methods on a 20-week part-time course, methods which they put into practice as the course progressed with 246 students of ages 8 and 9 and with 108 older students with ages between 10 and 14 (Fontana & Fernandes, 1994). The students of a further 20 Portuguese teachers who were taking another course in education at the time served as a control group. Both experimental and control groups were given pre- and post- tests of mathematics achievement, and both spent the same times in class on mathematics. Both groups showed significant gains over the period, but the experimental group's mean gain was about twice that of the control group's for the 8 and 9-year-old students--a clearly significant difference. Similar effects were obtained for the older students, but with a less clear outcome statistically because the pre-test, being too easy, could not identify any possible initial difference between the two groups. The focus of the assessment work was on regular--mainly daily--self-assessment by the pupils. This involved teaching them to understand both the learning objectives and the assessment criteria, giving them opportunity to choose learning tasks and using tasks which gave them scope to assess their own learning outcomes. The findings of their study were in line with the findings of the present study which revealed the significance of formative assessment in the process of learning.

In a study conducted by Whiting et al. (1995), the performances of 7000 students during 18 years were recorded. This involved regular testing and feedback to students, with a

requirement that they either achieve a high test score--at least 90%--before they were allowed to proceed to the next task, or, if the score were lower, they study the topic further until they could satisfy the mastery criterion. Whiting's final test scores and the grade point averages of his students were consistently high, and higher than those of students in the same course not taught by him. `Me students' learning styles were changed as a result of the method of teaching, so that the time taken for successive units was decreased and the numbers having to retake tests decreased. In addition, tests of their attitudes towards school and towards learning showed positive changes. The findings of semi-structured interview yielded the same results with the findings of their study considering both the impact of assessment and learners' ideas toward using this method of assessment in their learning.

The results of the present research indicated the positive impact of formative assessment on students' listening comprehension ability. Moreover, students' answer to the questions of the interview revealed that they had positive attitude about the application of diagnostic formative assessment for listening comprehension enhancement.

5. Conclusions

As explained in the previous section, students' performance in the experimental group was highly increased regarding their answers to listening comprehension test. Not only is it possible due to practicing the listening comprehension materials, but also it may be the consequences of applying diagnostic formative assessment which helped students understand the weaknesses in their listening comprehension ability to cover them. Students in the experimental group practiced listening comprehension content through implementing diagnostic formative assessment and by means of applying these methods of evaluation they enhanced their ability in the listening comprehension posttest.

The other aspect of this study deals with self-regulation. It is an internal factor which can be differed from time to time and situation to situation. As mentioned above, students' self-regulation was measured by a questionnaire before and after presenting the materials and practicing the rules of diagnostic formative assessment. Since diagnostic formative assessment deals with checking students' gradual progress during the process of learning, it could be a good reason for students' high performance in the second distribution of self-regulation questionnaire.

One more point which is worth considering is that most of students in the experimental group agreed upon the benefits of using diagnostic formative assessment while practicing listening comprehension materials during the sessions of instruction because it helped them increase their understanding of possible development from one session to another.

As the results of this study revealed, students in the experimental group who took advantage of working with formative assessment and understood their weak points and finally covered them outperformed students in the control group who did not receive any instructions about diagnostic formative assessment during the process of learning listening comprehension content.

References

- Bell, B. & Cowie, B. (2001). *Formative assessment and science education*. Dordrecht: Kluwer.
- Black, P. (2000). Research and the development of educational assessment. *Oxford Review of Education*, 26(34), 407–419.
- Black, P. (1993). Formative and summative assessment by teachers. *Studies in Science Education*, 21, 49–97.
- Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in education*, 5(1), 7-74.
- Birenbaum, M., & Dochy, F. J. (Eds.). (1996). *Alternatives in assessment of achievements, learning processes and prior knowledge*. Norwell, Massachusetts: Kluwer Academic Publishers.
- Birjandi, P., & Siyyari, M. (2010). Self-assessment and peer-assessment: a comparative study of their effect on writing performance and rating accuracy. *IJAL*, 13(1), 23-45.
- Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review*, 54(2), 199– 231.
- Bouffard, T., Boisvert, J., Vezeau, C. & Larouche, C., (1995). The impact of goal orientation on self-regulation and performance among college students. *British Journal of Educational Psychology*, 65, 317-329
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment, *TESOL Quarterly*, 32(4), 653-675.
- Buck, G., & Tatsuoka, K. (1998). Application of the rule-space procedure to language testing:
- Butler, D. L., Cartier, S. C., Schnellert, L., Gagnon, F., & Giammarino, M. (2011). Secondary students' self-regulated engagement in reading: Researching self-regulation as situated in context. *Psychological Test and Assessment Modeling*, 53(1), 73-105.

- Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1995). Teachers' assessment practices: *Preparation, isolation, and the kitchen sink. Educational Assessment, 3*(2) 159-179.
- Cleary, T. J., & Zimmerman, B. J. (2004). Self-regulation empowerment program: A school-based program to enhance self-regulated and self-motivated cycles of student learning. *Psychology in the Schools, 41*(5), 537-550.
- Conaway M (1982). Listening: Learning too and retention agent. In: Algier AS & Algier KW (Eds.). *Improving reading and study skills*. San Francisco: Jossey-Bass.
- Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of educational research, 58*(4), 438-481.
- Dawes L (2008). *The essential speaking and listening: Talk for learning*.
- Flowerdew, J. (1994). *Academic Listening: Research Perspectives*. New York: Cambridge University Press.
- Fontana, D. & Fernandes, M. (1994). Improvements in mathematics performance as a consequence of self-assessment in Portuguese primary school pupils. *British Journal of Educational Psychology, 64*, 407-417.
- Forgas, J. P., Baumeister, R. F., & Tice, D. M. (2009). The psychology of self-regulation: An introductory review. In J. P. Forgas, R. F. Baumeister, & D. M. Tice (Eds.), *The psychology of self-regulation* (pp. 1-17). New York: Psychology Press.
- Gipps, C. (1994). Beyond the test. *Improving Schools, 10*, 249-260.
- Heo, N. (2014). The relationship between creative potential and self-regulation among high achieving young adults with themoderating effects of parenting styles." PhD (Doctor of Philosophy) thesis, University of Iowa. <http://ir.uiowa.edu/etd/4643>.
- Huang, S.-C. (2011). Convergent vs. divergent assessment: Impact on college EFL students' motivation and self-regulated learning strategies. *Language Testing, 28*(2), 251-271.
- Kline JA (1996). *Listening effectively*. Alabama: Air University Press.
- Kurita, T. (2012). Issues in Second Language Listening Comprehension and the Pedagogical Implications. *Accents Asia, 5*(1), 30-44.
- Lambert, D., & Lines, D. (2000). *Understanding assessment: Purposes, Perceptions, Practice*. London: Routledge Falmer.
- Lee, Y. W., & Sawaki, Y. (2009). Cognitive diagnosis and Q-matrices in language assessment.
- McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.

- Melanlioglu, D. (2013). Impacts of authentic listening tasks upon listening anxiety and listening comprehension. *Educational Research and Reviews*, 8(14), 1177-1185.
- Newman, D., Griffin, P., & Cole, M. (1989). *The construction zone: Working for cognitive change in school*. Cambridge University Press.
- OECD (2003), *Learners for Life: Student Approaches to Learning: Results from PISA 2000*, OECD, Paris.
- Peeters, J., De Backer, F., Reina, V. R., Kindekens, A., Buffel, T., & Lombaerts, K. (2014). The Role of Teachers' Self-regulatory Capacities in the Implementation of Self-regulated Learning Practices. *Procedia-Social and Behavioral Sciences*, 116, 1963-1970.
- Posner, M. I., & Rothbart, M. K. (2000). Developing mechanisms of self-regulation. *Development and psychopathology*, 12(03), 427-441.
- Richards, J. C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly*, 17(2), 219–240.
- Ross, S. (2005). The impact of assessment method on foreign language proficiency growth. *Applied Linguistics*, 26, 317-342.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119–144.
- Schwartz, M. H. (2003). Teaching law students to be self-regulated learners. *L. Rev. MSU-DCL*, 447.
- Whiting, B., Van Burgh, J.W. & Render, G.F. (1995) Mastery learning in the classroom. Paper presented at the annual meeting of the American Educational Research Association. San Francisco, CA. ERIC Document ED382688.