



The Effects of Implementing Summative Assessment, Formative Assessment and Dynamic Assessment on Iranian EFL Learners' Listening Ability and Listening Strategy Use

Davood Ghahremani¹

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

Received: 23 April, 2011 Accepted: 13 January, 2012

Abstract

The present study was conducted to investigate the effects of implementing three forms of assessment namely , summative, formative and dynamic assessment on Iranian freshmen's listening ability and listening strategy use to fulfill the purposes of the study , 140 freshmen from Garmsar university and jame-elmikarbordi university who were majoring in English translation were selected . They formed randomly three experimental groups. Each group experienced a certain type of assessment. To study the effects of the assessment types, the learners took five teacher-made listening tests. Moreover, to observe any development in the learners' level of listening strategy use a questionnaire based on O'Malley Chamot and Kupper (1989), Young (1997) and Goh (2000, 2002) was used. The results indicated that the learners in dynamic group not only could outperform the other groups in terms of listening ability, but they also used more listening strategies.

Keywords: Summative assessment, Formative Assessment, Dynamic Assessment, Listening Strategy Use

Background

McNamara (2001) believes that the process of testing is a universal feature of social life. Plenty examples in history have shown that people have been put to the test to prove their capability or to establish their qualifications. Implementing tests to see how a person performs according to a certain level of performance has become an important social tradition and plays a key role in that it directs admission to many important social positions. Bachman (1990) has defined the effect of testing on teaching and learning as backwash, and believes that it can be harmful or beneficial. If the content of the test and testing techniques are inconsistent with the objectives of the course, the test may cause harmful backwash. The basis of traditional testing methods such as the translation method was considered subjective and the accuracy and fairness of such

evaluations were considered at best questionable. These inadequacies forced language testing specialists to apply more objective measures and develop new approaches. Discrete point approach appeared out of developments in linguistics and psychology which led to the appearance of structural linguistics and behavioristic psychology. Based on the principles of structural linguistics and behavioristic psychology the audio-lingual approach to language teaching was developed. It was believed that measurement of three levels of language i.e., sounds, words and sentences can manifest the language proficiency of the learners. Although these tests were highly reliable and valid, new developments in linguistics questioned the foundations of this type of tests (Fulcher, 2010).

The principles of generative – transformational linguistics along with cognitive

*Corresponding Author' Email: ghahremani2006@yahoo.com



psychology assumptions created a new approach in language teaching which was known as cognitive-code learning theory. In accordance with this teaching theory a new approach in language testing was popular which was called the integrative theory. According to this theory language is a holistic phenomenon and that knowledge of discrete items does not guarantee using language in real life successfully. Integrative tests such as oral interviews, reading comprehension tests, compositions, listening comprehension tests, dictation type tests and cloze procedures are more popular. This process led to a new approach called the notional-functional approach. The testing method developed based on this approach is called functional testing. The main purpose of this method is assessing learners' ability in carrying out language functions (Davidson, 2004 & Lynch, 2001).

Escalating criticism on implementing tests inappropriately for classroom assessment has made educators think about other alternative assessment methods. The new assessment paradigms view assessment as an ongoing process through which learning process includes both observing and learners' involvement in making decisions about their abilities. Assessment is considered as an interactive process through which both teachers and learners monitor the learners' performance (Buck, 1994).

Dynamic assessment takes into account results of intervention. The examiner teaches the examinee how to perform better on individual items or on the test as a whole. Final score may be a learning score representing the difference between pretest (before learning) and posttest (after learning) scores, or it may be the score on the posttest considered alone. Contrary to traditional methods of assessment which focused on product of past development, the focus of types of dynamic assessment is interventionist and is on interactionist's future development. On the other hand, in traditional assessment methods the relationship between examiner and the examinee is neutral and disinterested while in dynamic assessment the examiner is interested in the examinees' development and assessment is performed in a very helpful atmosphere. The most distinguishing feature which differentiates the traditional assessment methods and dynamic assessment is the process of providing feedback. In the traditional methods there was usually no

specific plan for giving feedback during the process of assessment meanwhile in dynamic assessment the process of assessment is mediated. (Sternberg & Grigorenko, 2002).

Statement of the Problem

Many ESL/EFL students' academic listening and speaking skills in English are not strong enough to cope with their academic study in English-medium universities, in particular understanding English lectures and expressing opinions and comments. Such students have not gained sufficient English language skills for their academic study (Brown, 2005).

On the other hand, McNamara (2001) has described traditional testing as a tendency to measure what language the students have learned as a result of teaching. Classroom teachers are normally more concerned with assessing achievement in both the formative and summative senses of the term. In these kinds of assessment, there is a strong emphasis on comparing students; and feedback to students comes in the form of marks or grades, with little direction or advice for improvement. These kinds of testing events indicate which students are doing best and which ones are doing poorly. Typically, they don't give much indication of mastery of particular ideas or concepts because the test content is generally too limited and the scoring is too simplistic to represent the broad range of skills and knowledge that have been covered. It seems that these types of assessment are not very useful for communicating complex data about a student's individual abilities (Fulcher & Davidson, 2007).

Research Questions and Hypotheses

Based on what was mentioned earlier and with regard to the focus of the present research paper, the researcher formed two questions and also formulated two hypotheses:

- 1- Which assessment method, formative, summative or dynamic has differential effects on the learners' listening ability?
- 2- Which assessment method, formative, summative or dynamic has differential effects on learners' listening strategy use?

Research Hypotheses

- 1- None of the assessment methods has differential effects on the learners' listening ability.

2- None of the assessment methods has differential effects on the learners' listening strategy use.

Participants

A total number of 140 students took part in this study. They were all freshmen majoring in English translation in Garmsar University and one of the colleges of the University of Applied Science and Technology (Elmi-Karbordi). The participants were both males and females, and enjoyed almost the same level of English language proficiency.

Sampling Procedure

In order to have a more homogeneous sample and measure listening proficiency of the subjects at the beginning of the study, a PET proficiency test was administered. After conducting the proficiency test, the participants fallen in one standard deviation above and below the mean were selected. This resulted in the selection of 120 (40 summative, 40 formative, 40 dynamic) participants for the study. The study lasted almost 14 sessions.

Instrumentation

The following data collection instruments were used in this study. The first instrument was Preliminary English Test known as PET. The listening part of the test played two roles in here. Firstly, the scores drawn from the test were considered as the whole for the proficiency level of the participants of the test which was used to judge the homogeneity of the groups and secondly, the learners' scores from the listening part of the test were recorded as the result of the pre- and post-test of the study. Therefore only the listening part was re-administered at the end of the treatment section. Moreover, because the reliability was jeopardized, a reliability analysis was applied indicating a reliability of 0.75 for the listening test which was satisfactory for the purposes of the present study. Because the interval between the two copies of the test was more than a four-week period, it was believed that repetition of the test would not harm the validity of the test. However, through a pilot study the validity of the listening test was checked against the PET test. The second instrument was the teacher-made tests. Five listening tests were made by the researcher from lower intermediate level of the New Interchange

series which were used for the treatment purpose of the study. The book was used as the main course book of the classes in both universities. Each unit of the book contained listening comprehension parts for lower intermediate students. The third instrument was a listening strategy use questionnaire which was based on O'Malley, Chamot & Küpper (1989), Young (1997) and Goh (2000, 2002). This questionnaire includes a series of listening comprehension strategies which are classified under three headings; metacognitive, cognitive and socio-affective strategies.

Design

The design of this research is a comparison group design, a subcategory of quasi-experimental design. Measuring the effect of treatment was conducted through a pretest /post- test design. The subjects were non-randomly assigned to five intact classes (one summative with 40 subjects, two formative with 40 subjects and two dynamic with 40 subjects). Yet the classes were randomly assigned to three different treatment groups. For the purpose of this study, the PET was used both as pre-test and post-test to measure the learners' listening proficiency before and after the treatment. To see effects of the treatment, the researcher constructed 5 listening tests based on New Interchange series which were the learners' course book. During the 14 session treatment, the learners in three groups took these tests but they were treated differently. The learners in summative classes received no specific feedback from their teacher. They merely took the test and were informed of their scores. The learners in formative classes, on the other hand, not only took the test, but also received feedback on their possible problems from their teacher. In formative class, the learners' problems were identified and they were provided with some hints on how to overcome their problems. In the dynamic class, the learners took the tests and received specific types of feedback. For this group, the learners were helped to see their problems and also based on dynamic assessment types i.e., interventionism and interactionism, they were provided by detailed explanation on the nature of their problems and the strategies that they could use to avoid or overcome the difficulties.

As it was already mentioned, in order to see the effects of implementing the assessment types on the learners' listening strategy use, a listening comprehension strategy use questionnaire based on

O'Malley, Chamot & Küpper (1989), Young (1997) and Goh (2000, 2002) was used. The learners checked the items of the questionnaire three times based on the researcher's predetermined

plan. The process of checking the questionnaire of listening strategies in the three groups was the same i.e., all the learners checked the questionnaire after taking the teacher made listening tests.

Table 1: Listening comprehension strategies

Activities for metacognitive strategies	Activities for cognitive strategies	Activities for socio-affective strategies
1. Preview the content in different forms.	1. Use prior knowledge and knowledge about the target language to elaborate and complete interpretation.	1. Paraphrase what speakers say to check understanding
2. Rehearse the pronunciation of potential content words.	2. Infer missing or unfamiliar words using contextual clues, familiar content words, visual clues.	2. Ask speaker for clarification and repetition.
3. Establishing the purpose for listening.	3. Draw on knowledge of the world.	3. Learn to relax before and during listening.
4. Practice perception regularly.	4. Apply knowledge about the target language.	4. Encourage oneself to continue listening.
5. Take short notes of important content words.	5. Visualize scenes, objects, events, etc. being described.	
6. Check current comprehension with context of the message and prior knowledge.	6. Reconstruct meaning using words heard.	
7. Continue to listen for clarification in spite of difficulty.	7. Relate one part of the text to another.	
8. Evaluate comprehension using contexts, prior knowledge and external resources.	8. Relate limited interpretation to a wider social/linguistic context.	
9. Determine potential value of subsequent parts of input.	9. Assess the importance of problematic parts and decide whether to ignore them or actively seek clarification.	
10. Listen selectively according to purpose.	10. Find L1 equivalents for selected key words.	
11. Listen for gist.	11. Translate a sequence of utterance.	
12. Determine the potential value of subsequent parts and vary intensity of attention accordingly.	12. Predict general contents before listening using contexts and prior knowledge.	
13. Memorize words or phrases for later processing.	13. Predict details and unfinished utterances using contexts and prior knowledge.	
14. Pay attention to discourse markers, visuals and body language, tones and pauses.		

Based on O'Malley, Chamot & Küpper (1989), Young (1997) and Goh (2000, 2002).

Procedure and Data Analysis

This research pursued to study the effects of implementing three forms of assessment; formative, summative and dynamic on Iranian EFL learners' listening ability and listening strategy use. In order to answer the research questions of the study the following steps were taken.

In order to see the effects of applying three forms of assessment on the learners' listening ability before and after the treatment, three one-way repeated- measurement ANOVA were run.

Moreover, to study the effect of applying three forms of assessment on the learners' listening strategy use three one- way repeated- measurement ANOVA were run.

Results and Discussion

In order to see any differential effect of the three assessment methods on the learners' listening ability the following statistical measures were adopted. First the descriptive statistics were computed based on the learners' scores obtained from the post test.

Table 2. Descriptive Statistics of Applying Different Assessment Methods on the Learners' Listening Ability

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Summative posttest	40	17.6750	4.54825	.71914	16.2204	19.1296	7.00	25.00
Formative posttest	40	19.5750	3.53653	.55917	18.4440	20.7060	13.00	25.00
Dynamic posttest	40	21.9500	2.60128	.41130	21.1181	22.7819	15.00	25.00
Total	120	19.7333	4.02249	.36720	19.0062	20.4604	7.00	25.00

As indicated in Table.2, the reported means are considerably different (17.67, 19.57, 21.95).

Later, in order to find any meaningful difference among the means of the groups ANOVA was used.

The results from ANOVA tables.3&4 reveal that the mean differences between the three groups is significant ($F_2 = 13.77$, $P = 0.00$).

Finally, to find the highest difference among the means of the groups the Scheffe Test was used. (Table 5)

The results of the Scheffetest Table 6 reveals that the highest differences are between dynamic and summative groups (mean difference= 4.27500, $P = 0.000$) with dynamic and formative coming next.

Table3. Test of Homogeneity of Variances

Levene Statistic	df1	df2	Sig.
7.887	2	117	.001

Table 4. ANOVA Comparing Three groups for Listening Ability

	Sum of Squares	df	Mean Square	F2	Sig.
Between Groups	367.017	2	183.508	13.777	.000
Within Groups	1558.450	117	13.320		
Total	1925.467	119			

Table 5. Multiple Comparisons for Three Groups for Listening Ability

	(I) listening posttest factor	(J) listening posttest factor	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Tukey HSD	Summative posttest	Formative posttest	-1.90000	.81609	.056	-3.8373	.0373
		Dynamic posttest	-4.27500*	.81609	.000	-6.2123	-2.3377
	Formative posttest	Summative posttest	1.90000	.81609	.056	-.0373	3.8373
		Dynamic posttest	-2.37500*	.81609	.012	-4.3123	-.4377
	Dynamic posttest	Summative posttest	4.27500*	.81609	.000	2.3377	6.2123
		Formative posttest	2.37500*	.81609	.012	.4377	4.3123
Scheffe	Summative posttest	Formative posttest	-1.90000	.81609	.071	-3.9234	.1234
		Dynamic Posttest	-4.27500*	.81609	.000	-6.2984	-2.2516
	Formative posttest	Summative posttest	1.90000	.81609	.071	-.1234	3.9234
		Dynamic posttest	-2.37500*	.81609	.017	-4.3984	-.3516
	Dynamic posttest	Summative posttest	4.27500*	.81609	.000	2.2516	6.2984
		Formative posttest	2.37500*	.81609	.017	.3516	4.3984

Table 6. Scheffe Test Results for Three Groups Listening Ability

Listening posttest factor	N	Subset for alpha = 0.05	
		1	2
TukeyHSD ^a	Summative posttest	40	17.6750
	Formative posttest	40	19.5750
	Dynamic posttest	40	21.9500
	Sig.		.056
Scheffe ^a	Summative posttest	40	17.6750
	Formative posttest	40	19.5750
	Dynamic posttest	40	21.9500
	Sig.		.071

Discussion

The results reported through the Tables 2,4,5 and 6 confirm the previous discussions for this question. The results indicate weighty development of the listening ability in the dynamic group. It is worth mentioning that this development can be related to both the quality and quantity of feedback that the learners have received throughout the treatment. As it was already mentioned, to examine the effect of assessment methods on the learners' listening strategies, three separate one-way repeated-measurement ANOVA were run. Based on the

results of table 7 which displays the means of the learners' listening strategy use, it is obvious that the listening strategy use of the learners in summative group has improved. As indicated in table 4.14, the means of the summative group strategy use are different (13.77, 15.30, 14.95).

The results obtained through repeated measure ANOVA table 8 signifies that this improvement is statistically significant with the F_2 value of 8.08 and the P value of 0.00. Since Mauchly's test of Sphericity is not found to be significant (P= 0.29), the part of table that corresponds to Sphericity assumed was used

Table 7. Descriptive Statistics for Summative Group (Listening Strategy Use)

	Mean	Std. Deviation	N
Summative Strategy Score1	13.7750	2.23592	40
Summative Strategy Score2	15.3000	2.16262	40
Summative Strategy Score3	14.9500	2.22975	40

Table 8. Repeated Measure ANOVA for Summative Group(Listening StrategyUse)

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
Sum straseries	.938	2.417	2	.299	.942	.988	.500

Table 9:Repeated Measure ANOVA for Summative Group

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
SumStraSeries	Sphericity Assumed	51.050	2	25.525	8.084	.001	.172
	Greenhouse-Geisser	51.050	1.884	27.098	8.084	.001	.172
	Huynh-Feldt	51.050	1.976	25.829	8.084	.001	.172
	Lower-bound	51.050	1.000	51.050	8.084	.007	.172
Error(SumStraSeries)	Sphericity Assumed	246.283	78	3.157			
	Greenhouse-Geisser	246.283	73.473	3.352			
	Huynh-Feldt	246.283	77.080	3.195			
	Lower-bound	246.283	39.000	6.315			

Table 10:Pairwise Comparison for Summative Group

(I) SumStraSeries	(J) SumStraSeries	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1	2	-1.525*	.345	.000	-2.389	-.661
	3	-1.175*	.414	.022	-2.211	-.139
2	1	1.525*	.345	.000	.661	2.389
	3	.350	.428	1.000	-.719	1.419
3	1	1.175*	.414	.022	.139	2.211
	2	-.350	.428	1.000	-1.419	.719

The results of pair wise comparisons revealed that more significant difference can be observed between the first and the second time with the mean of (1.52) and the P value of 0.00. The difference between time 1 and 3 was significant only at 0.05 level. Similar studies performed by other researchers (Chamot & Kupper, 1989, Chamot, 1995, Byrnes, 2002), have reported significant improvement in learners' listening strategy use caused by strategy training and awareness raising. The results of these studies have implicitly stated the role of assessment type in improving learning strategy use. From Table 4.18, we can judge that the means of the formative group are considerably different (12.60, 15.75, 19.55), indicating the development of listening strategy use due to

applying formative assessment be significant (P= 0.00), the part of table that corresponds to Greenhouse-Geisser was used. Since based on Table 12 and 13, Mauchly's test of Sphericity is found to as displayed in Table 13 the difference between the means are statistically significant ($F_{1,45} = 239.454$, P= 0.00). The results of pairwise comparisons in Table 14 reveals that the difference between each pair is significant with the P value of 0.00.

The results reported through the Tables 11, 13 & 14, indicate significant increase in learning strategy use by the learners in formative group. According to Douglas (2000) the assessment perspective adopted by the teacher and practiced by the learners can significantly influence both the quality and quantity of the actions that the learners take in the classroom.

Table 11. Descriptive Statistics for Formative Group (Listening Strategy Use)

	Mean	Std. Deviation	N
FormativeStrategyScore1	12.6000	1.64551	40
FormativeStrategyScore2	15.7500	1.42775	40
FormativeStrategyScore3	19.5500	1.88040	40

Table 12. Mauchly's Test of Sphericity for Formative Group

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
FormStraSeries	.622	18.029	2	.000	.726	.747	.500

Table 13: Tests of Within-Subjects Effects for Formative Group

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Forms strategy	Sphericity Assumed	968.867	2	484.433	239.454	.000	.860
	Greenhouse-Geisser	968.867	1.452	667.441	239.454	.000	.860
	Huynh-Feldt	968.867	1.493	648.884	239.454	.000	.860
	Lower-bound	968.867	1.000	968.867	239.454	.000	.860

Table 14. Pairwise Comparisons for Formative Group

(J) FormStraSeries	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a		
				Lower Bound	Upper Bound	
1	2	-3.150*	.279	.000	-3.847	-2.453
	3	-6.950*	.403	.000	-7.959	-5.941
2	1	3.150*	.279	.000	2.453	3.847
	3	-3.800*	.251	.000	-4.428	-3.172
3	1	6.950*	.403	.000	5.941	7.959
	2	3.800*	.251	.000	3.172	4.428

The result of Table 15 indicates a significant development in dynamic group's listening strategy use (13.40, 17.17, 21.37). Since Mauchly's test of Sphericity, as it can be seen from Table 16 is not found to be significant ($P= 0.350$), the part of table that corresponds to Sphericity Assumed was used. As displayed in Table 17 the difference between the means are statistically significant ($F_2= 359.518$, $P= 0.00$).

The results of pairwise comparisons, Tables.17&18 reveal that the difference between each pair is significant with the P value of 0.00. The results of this part according to Tables 15, 17 and 18 indicate that the learners in dynamic group could outperform the learners in summative and

formative groups. As it was discussed in the previous section, feedback has an important and key role in shaping the process of learning. Dynamic assessment means that interaction can take place, and feedback can be given, during the assessment or examination, which separates it from more "traditional assessments" (McKay, 2006).On the other hand, the feedback that the learners usually receive in dynamic assessment classes are more timely and personal. It seems that student have found this mode of examination interesting and instructive, especially the comparison part. This form of giving feedback is sure to enhance motivation and engagement in learning process (Falchikov, 2001).

Table 15. Descriptive Statistics for Dynamic Group Strategy Use

	Mean	Std. Deviation	N
Dynamic Strategyscore1	13.4000	1.19400	40
Dynamic Strategyscore2	17.1750	1.55064	40
Dynamic Strategyscore3	21.3750	1.68990	40

Table 16. Mauchly's Test of Sphericity for Dynamic Group

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	Df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
DynStraSeries	.946	2.099	2	.350	.949	.996	.500

Table 17. Tests of Within-Subjects Effects for Dynamic Group

Source	Type III Sum of Squares	df	Mean Square	F2	Sig.	Partial Eta Squared	
DynStraSeries	Sphericity Assumed	1273.217	2	636.608	359.518	.000	.902
	Greenhouse-Geisser	1273.217	1.898	670.824	359.518	.000	.902
	Huynh-Feldt	1273.217	1.992	639.058	359.518	.000	.902
	Lower-bound	1273.217	1.000	1273.217	359.518	.000	.902

Table 18. Pairwise Comparisons for Dynamic Group

(I) DynStraSeries	(J) DynStraSeries	Mean Difference (I-J)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^a	
					Lower Bound	Upper Bound
1	2	-3.775*	.267	.000	-4.442	-3.108
	3	-7.975*	.296	.000	-8.716	-7.234
2	1	3.775*	.267	.000	3.108	4.442
	3	-4.200*	.327	.000	-5.017	-3.383
3	1	7.975*	.296	.000	7.234	8.716
	2	4.200*	.327	.000	3.383	5.017

Summary and Discussion of Findings

The present study aimed at investigating the effects of implementing three forms of assessment namely formative, summative and dynamic on Iranian freshmen university students' listening ability and listening strategy use.

Considering the effects of the three assessment methods on the learners' listening ability, the obtained results indicated considerable difference between the means of the groups (Summative (17.67), Formative (19.57), Dynamic (21.95)).

Assessments is usually used to measure how much students have learned up to a particular point in time (Cheng, 2005; Ableeva, 2008). In summative assessment which is known as assessment of learning, the teacher intends to see whether students are meeting standards set by the state, the district, or the classroom teacher. According to Black & William (1998), summative assessments are conducted after a unit or certain time period to determine how much learning has taken place. These authors also believe that if the purpose of evaluation is assigning grades to students and providing accountability, these forms of assessment are important but they do not support learning during the learning process which is a considerable deficit for this kind of assessment. It may, however, be argued that learners usually receive feedback whether directly or indirectly from their peers or teachers, but the point which should be taken into consideration is that in summative classes usually there is no planned and continuous support and feedback (Leung, 2007).

On the other hand, the results of the present study indicated considerable improvement of the learners in the formative group. As it was mentioned above, this development happened due to the learners' more active role in understanding success and the support and feedback that they received from their teacher after assessment. Similar findings are reported in Campione (1989), Buck (2001) and Cassidy, (2007). These authors have reported that it will be most useful if teachers provide some feedback to the learners, perhaps in the form of a brief comment or, at the very least, a check, check-plus or check-minus, with a brief verbal explanation about what each symbol indicates (e.g. "You have mastered the skill, You need more practice", etc.). They have also emphasized that using different forms of assessment, teachers can arrive at a more accurate picture of what students know and understand. Formative

assessment weighs understanding along the process of learning and directs teacher decision making about future instruction. Formative assessments also provide feedback to students so they can improve their performance. Rea-Dickins, (2001) has described a formative class in which learners try to form a picture of success and to use each assessment to learn how to improve their learning. Shohamy, (2001) believes that the least quantity use of formative assessment enables teachers evaluate the process of learning in their classes both qualitatively and quantitatively. Teachers will also be able to check their learners' development and decide on what they need to develop their mastery.

The reported results of this part are in keeping with similar studies done by Cioffi, & Carney, (1983), Ableeva, (2008) & Cheng, (2005). The results reported by these authors have highlighted significant influence of dynamic assessment on learners, whether children or adult, their cognitive development, school performance, motivation and overcoming their learning disabilities.

References

- Ableeva, R. (2008). The effects of dynamic assessment on L2 listening comprehension. In J. P. Lantolf & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 57-86). London: Equinox.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.
- Black, P. & William, D. (1998). Assessment and classroom learning. *Educational Assessment: Principles, Policy and Practice*. 5(1), 7-74.
- Brown, J. D. (2005). *Testing in language programs: A comprehensive guide to English language assessment*. McGraw-Hill.
- Buck, G. (1994) The appropriacy of psychometric measurement models for testing second language listening comprehension. *Language Testing*, 11, 145-170.
- Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press.
- Byrnes, H. (2002). Toward academic-level foreign language abilities: Reconsidering foundational assumptions, expanding pedagogical options. In B. L. Leaver & B. Shekhtman (Eds.), *Developing professional-level language proficiency* (pp. 34-58). Cambridge, UK: Cambridge University Press.

- Campione, J. C. (1989). Assisted assessment: a taxonomy of approaches and an outline of strengths and weaknesses. *Journal of Learning Disabilities* 22, 151–65.
- Cassidy, S. (2007). Assessing “inexperienced” students’ ability to self-assess: Exploring links with learning style and academic personal control. *Assessment and Evaluation in Higher Education*, 32(3), 313–330.
- Chamot, A. U., & Kupper, L. (1989). Learning strategies in foreign language instruction. *Foreign Language Annals*, 22/1, 13–21.
- Chamot, A. U. (1995). Learning strategies and listening comprehension. In D. H. Mendelsohn & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 13–30). San Diego, CA: Dominic Press.
- Cheng, L. (2005). *Changing Language Teaching through Language Testing: A Washback Study*. Cambridge: Cambridge University Press.
- Cioffi, G., & Carney, J. (1983). Dynamic assessment of reading disabilities. *The Reading Teacher*, 36, 764–768.
- Cioffi, G., & Carney, J. (1983). Dynamic assessment of reading disabilities. *The Reading Teacher*, 36, 764–768.
- Davidson, C. (2004). The contradictory culture of teacher-based assessment: ESL assessment practices in Australian and Hong Kong. *Language Testing*, 21(3), 305–334.
- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge, UK: Cambridge University Press.
- Falchikov, N. (2001). *Learning together. Peer tutoring in higher education*. London & New York: Routledge.
- Fulcher, G. (2010). *Practical language testing*. London, UK: Hodder Education.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Goh, C. C. M., 2000. A Cognitive Perspective on Language Learners “Listening Comprehension Problems. *System*, 28, pp.55–75.
- Goh, C. C. M., (2002). *Teaching Listening in the Language Classroom*. Singapore: SEAMEO Regional Language Centre.
- Leung, C. (2007). Dynamic assessment: Assessment for and as teaching? *Language Assessment Quarterly: An International Journal* 4: 257–278.
- Lidz, C., & Elliott, J. (Ed.), (2000). *Dynamic assessment: prevailing models and applications*. Oxford: Elsevier Science.
- Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language Testing* 18: 351–372.
- McNamara, T. F. (2001). Language assessment as social practice: challenges for research. *Language Testing* 18: 333–349.
- McKay, P. (2006). *Assessing young language learners*. Cambridge, UK: Cambridge University Press.
- O'Malley, J. M., Chamot, A. U. & Küpper, L., (1989). Listening Comprehension Strategies in Second Language Acquisition. *Applied Linguistics*, 10 (4), pp.418–437.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18(4), 429–462.
- Shohamy, E. (2001). *The Power of Tests: A Critical Perspective on the Uses of Language Tests*. New York: Longman.
- Sternberg, R.J., & Grigorenko, E. L., (2002). *Dynamic Testing. The Nature and Measurement of Learning Potential*. Cambridge: Cambridge University Press.
- Young, M. Y. C., (1997). A Serial Ordering of Listening Comprehension Strategies Used by Advanced ESL Learners in Hong Kong. *Asian Journal of English Language Teaching*, 7, pp.35–53.
- Davood Ghahremani** has a PhD in English language teaching from Islamic Azad University, Tehran, Iran. He has experienced teaching English in both schools and universities. At present, he is a faculty member of Islamic Azad University, Garmsar Branch and the chairman of Elm-Gostar higher education institute, Tehran, Iran. He is interested in language testing and research.