Journal of Studies in Learning
and Teaching English
Volume. 5, Issue. 2, Ser. 10, (2016), 105-122

The Effect of Portfolio Assessment on the Development of Metadiscourse Markers Awareness in EFL Learners' Oral Performance

Tooba Mardani
MA in Translation
Studies Young Researchers and Elite Club
Isfahan (Khorasgan) Branch, Islamic Azad University
Isfahan, Iran
Mardani_mars@yahoo.com

Abstract. Portfolio assessment as one of the alternatives to testing is defined as the systematic collection of student work measured against predetermined scoring criteria. This paper aimed to investigate the effect of portfolio assessment in the oral performance of EFL learners in an attempt to examine its impact on their metadiscourse awareness. To determine the impact of portfolio assessment on the oral performance and metadiscourse awareness of EFL learners, 60 university students majoring in English language teaching were initially selected. After being homogenized for their proficiency level, they were randomly divided into an experimental group (EG) and a control group (CG). As the treatment, portfolio assessment was employed as the teaching technique for the experimental group. Data was then subjected to different statistical procedures. The results of data analysis revealed that the participants in the EG outperformed those in the CG with regard to the achievement in their overall oral performance. Second, the findings showed that experimental groups' awareness to metadiscourse markers significantly increased after instruction. The findings of the present study have implications for learners and teachers in the realm of TEFL in particular and education in general.

Keywords: Metadiscourse awareness, portfolio assessment, traditional assessment, oral performance, portfolio-based instruction

Received: December 2015; Accepted: May 2016

1. Introduction

The theoretical foundation that underpins this study is the notion of "constructivism" which refers to more recent views on teaching and learning proposing that all individuals learn by constructing information about the world and by using active and dynamic mental processes (O' Malley & Valdez Pierce,1996). Following this line of investigation, many researchers have come to recognize that alternative assessment is an important means of gaining a dynamic picture of students' academic and linguistic development. "Alternative assessment refers to procedures and techniques used within the context of instruction which can be easily incorporated into the daily activities of the school or classroom" (Hamayan, 1995, p. 213). In contrast to traditional testing, through alternative assessment, students are evaluated on what they integrate and produce rather than on what they are able to reproduce and recall (Hamp-Lyons, 1996).

Johnson (1996) defines portfolios as a cumulative collection of work students do. In other words, portfolios show a student's work from beginning of the term to the end. Yang (2003) defined portfolio as a compilation of students' work, which documented their effort, progress and achievement in their learning, and their reflection on the materials negotiated for the portfolio. Crosby (1997) indicated that the primary purpose of portfolios in EFL context was to increase the level of students' motivation and to give them a sense autonomous learning.

Portfolio assessment as one of the alternatives to testing is defined as the systematic collection of student work measured against predetermined scoring criteria. These benchmarks may include scoring guides, rubrics, check lists, or rating scales (O'Malley & Valdez Pierce, 1996). Portfolio assessment is a systematic collection of a variety of teacher observations and student products, collected over time, that reflect a student's developmental status and progress (cited in Shabban, 2001). Among such alternatives to testing, portfolio assessment was singled out in this study because it might have potential effect on instruction. So, this paper aimed to explore the effect of portfolio assessment on the students' oral performance, especially on their knowledge and use of

metadiscourse markers.

According to Douglas (2000), portfolio assessment is particularly applicable to foreign-language assessment. Standardized tests, noted O'Mall- ey and Chamot (1990), provide foreign language teachers with an incomplete picture of student needs and learning. Hamp-Lyons & Condon (2000) argued that the "greatest theoretical and practical strength of a portfolio, used as an assessment instrument, is the way it reveals and informs teaching and learning" (p.4). Lee (2001) pointed out that portfolio assessment prioritized student-centered over conventional concept of teaching.

Hamp-Lyons & Condon (2000) believed that in portfolio assessment both reliability and validity were necessary and must be addressed. Williams (1998, 2000) argued that without standards for implementation and outcomes, portfolio assessment will become unfair because "it increases the subjectivity teachers bring to evaluation" (2000, p.136). Reliability needs to be based on performance instead of idiosyncratic scores that have no preset criteria (O'Malley & Pierce, 1996). One of the main problems of reliability in L2 portfolio assessment is inter-rater reliability, consistency of scores because teacher are not used to this new concept of assessment. The psychometric view of reliability is too narrow to take into account the "less standard forms of assessment" such as the portfolio. Portfolio assessment requires that readers be trained to agree and to score papers based on a common rubric that describes numerical points. Rubrics should include development and organization, fluency of idea (problem) description, and mechanics (O'Malley & Pierce, 1996). If readers agree, there is a reliable rate of agreement. If readers do not agree, there is low inter-rater reliability. Without a sufficiently high rate of reliability, scores cannot be considered valid. Portfolios can be used to support or even determine a grade when a well-defined scoring guide or rubric is developed by both the teacher and the learner (Defina, 1992, P. 37).

Portfolios as teaching tools demand of teachers and students' rich concepts and a careful consideration about what kind of good works to be put into them. As students and their teachers work together on the development of the portfolio over the term, the teacher is able to

assess the student's growth and learning in the course as well as the excellence reached by the end of the course. Portfolios, then, are 'a tool for thoughtful classroom assessment' (Hamp-Lyons, 1996). Portfolios have been associated with alternative assessment not only in general education but more particularly in second language education as well (Darling-Hammond, 1994; Hamayan, 1995; Shohamy, 2001). The literature reveals a degree of controversy and confusion concerning the use of portfolio assessment as an alternative to traditional testing. It is argued that portfolio assessment is more than merely one of many homogeneous alternatives in assessment (Brown & Hudson, 1998). As further argued by Lynch and Shaw (2005, p. 264),"the portfolio, as an exemplar of alternative assessment, represents a different paradigm or culture that requires an approach to validity evidence (i.e., to establishing the trustworthiness of the inferences made from the assessment process) differing in certain critical aspects from the approach used in traditional testing."

Many teachers and experts have offered suggestions on portfolio contents based on their experience with using PA. Portfolios are often grouped into four types according to Mandell and Michelson (1990):

* Showcase-student only puts best example or best product in for each objective

* Cumulative-Student place all work relevant to each objective into portfolio

* Process-Student places pre/post-samples of work for each objective into the portfolio

* Each type of portfolio should include all of the essential components of a portfolio listed above

Crockette(1998) considers portfolio contents to fall into five categories:

• Found samples, which refer to pieces done to fulfill class assignment

• Processed samples, or the students' analyses and self-assessment of a work previously graded by the teacher

• Revisions or samples of student work that have been graded and then revised, edited, and rewritten

• Reflections, which are related to the processed samples but are applied to the portfolio as a whole, providing a chance for students to think about who they are, what strengths and weaknesses are

• Portfolio projects, which cover work designed for the sole purpose of inclusion in student portfolios.

Crockette (1998) suggests that the portfolio should include the contents mentioned above as well as other items considered relevant to its specific purpose.

And essential components of a portfolio suggested by Campell, Melenyzer and Nettles, D. & Wyman, R. (2000) should be:

* Table of contents with page numbers

* Personal introduction describing the student's background and capstone experience

* Program outcomes, with artifacts linked to the outcomes

Each artifact should be described in a short narrative(included with artifact) reflecting upon what it is, how it demonstrates obtainment of the objective, and what the student learned as a result (self-reflective). Requiring more than one artifact/objective increases scoring reliability (Campell, Melenyzer, Nettles & Wyman, 2000).

Portfolio assessment is a systematic collection of a variety of teacher observations and student products, collected over time, that reflect a student's developmental status and progress (cited in Shabban, 2001). Genesee and Upshur (1996) provided a plausible reason for the application of portfolio assessment: "A portfolio is purposeful collection of students' work that demonstrates to the students and others their efforts, progress, and achievements in given areas" (p. 99). They maintained that the primary value of portfolios was in the assessment of student achievement because they provide a continuous record of students' language development that can be shared with others.

Was O'Malley and Chamot (1990) indicated that a key element of portfolios is student self-assessment; without self-assessment and reflection on the part of the student, a portfolio is not a portfolio. Moreover,

many researchers (e.g., Genesee & Upshur, 1996; Upshur and Turner 1998; Kormos 1999; Papajohn, 1999; Lynch, 2001; Khoshsima, 2006, to just mention a few) investigated the effect of portfolio assessment on learning process in classroom context. Their findings depicted plausible impact on the assessment of students' classroom performance.

The implementation of a portfolio assessment system in many classrooms today is of great interest to teachers and researchers at every level of education. Murphy (1999) stated that they are implemented in classrooms in order to accomplish various goals. These goals range from providing students with a sense of ownership, motivation, accomplishment to assessing curriculum needs for demonstrating competency. Because of the number of goals associated with the implementation of portfolios, Murphy categorized them into three groups: "teaching tools, professional development, and assessment purposes" (Murphy, S. 1999, p. 4). Following this line of investigations, the current study has taken both the formative and summative functions of portfolio assessment in the classroom so as to clearly portray the oral performance of the students through their real performances during the course and at the end of the instruction. This research specifically focuses on participants' use of "metadiscourse markers". As Hyland (2005) stated in his book on metadiscourse, "the term metadiscourse was coined by Zelling Harris in 1959 to offer a way of understanding language in use, representing a writer's or reader's attempts to guide a receiver's perception of a text" (P. 3). The concept has been further developed by writers such as Williams (1981), Vande Kopple (1985) and Crismore (1989).

Metadiscourse was first defined by Williams (1981) as 'writing about writing'. Vande Kopple (1985) also referred to metadiscourse as 'discourse about discourse or communication about communication'. Hyland (2004) maintained that metadiscourse helped authors interact with their audience in order to communicate successfully with them. Moreover, Hyland (1998) contended that metadiscourse markers were integral to the text. In other words, they cannot be removed or changed at will. In a quantitative study, Hyland (1998) examined metadiscourse markers in 28 research articles and found 373 instances of metadiscourse in each research article. In another textual analysis, Hyland (1999) ex-

plored metadiscourse markers in 21 textbooks and found 405 instances of metadiscourse markers in each text, around one per 15 words. Hyland has concluded that metadiscourse play an important part in communication. Crismore (1984) defined metadiscourse as "discoursing about spoken or written discourse" (p.66). She has added that metadiscourse provides readers or listeners with direction rather than information.

Indubitably, the advantages of metadiscourse are many. For instance, discourse structuring functions of metadiscourse guide readers through a text and help them organize content while reading, thus creating global comprehension (Crismore, 1989). Metadiscourse has been recognized as one of the major rhetorical features and strategies in the production of a text (Hyland, 1998). In fact, it "is not indispensable stylistic device which authors can vary at will. It is integral to the contexts in which it occurs and is intimately linked to the norms and expectations of particular cultural and professional communities" (Hyland, 1998).

According to Vande Kopple (2002), metadiscourse refers to elements in texts that convey meanings other than those that are primarily referential. As Hyland (2004) stated, the importance of metadiscourse lies in its underlying rhetorical dynamics which relate it to the contexts in which it occurs.

The present study proposes Hyland's (2004) divisions of metadiscourse markers as the framework of itself.

Hyland (2004) divides metadiscourse markers into two broad categories, each one with a set of subcategories, interactive (textual) and interactional (interpersonal):

I. Interactive Markers: They enable the writer to manage the information flow so as to provide his preferred interpretations. They include the following subtypes:

1. Transitions: These markers mainly indicate: additive, contrastive, and consequential steps in the discourse. Some examples are: in addition, but, thus, and, etc.

2. Frame markers: They indicate text boundaries or elements of schematic text structure, like: my purpose here is to, finally, to conclude, etc.

3. Endophoric markers: They refer to information in other parts of the text and make the additional material available to the readers. Some examples are: in section, see figure, noted above, etc.

4. Evidentials: They refer to sources of information from other texts, such as: X states, (Y, 2010), According to X, etc.

5. Code glosses: They help readers grasp functions of ideational material. They show the restatements of ideational information, like: namely, such as, in other words, e.g., etc.

II. Interactional Markers: They involve the reader in the argument. They 'focus on the participants of the interaction and seek to display the writer's persona and a tenor consistent with the norms of the disciplinary community' [Hyland 2004, p.139]. The interactional resources include:

1. Hedges: They withhold writer's full commitment to proposition. Examples: might, about, perhaps, possibly, etc.

2. Boosters: They emphasize force or the writer's certainty in proposition.Examples: it is clear that, in fact, definitely, etc.

3. Attitude markers: They indicate the writer's appraisal or attitude to propositional information. Some examples are: unfortunately, surprisingly, I agree, etc.

4. Engagement markers: They explicitly refer to or build a relationship with the reader. Examples: consider, you can see that, note that, etc.

5. Self-mentions: They explicitly refer to authors' presence in terms of first person pronouns and possessives. Examples: I, we, our, my, your, etc.

Following the above-mentioned studies, the present research tried to investigate the effect of portfolio assessment on using metadiscourse markers. The present study aimed at investigating whether metadiscourse markers can help English learners to improve their oral performance and whether the participants who received portfolio assessment could develop a plausible understanding of that knowledge or awareness in speaking English.

2. Research Questions

The purpose of this study was to investigate the implementation of portfolio assessment in an academic context in an attempt to examine its effect on the students' oral performance, especially their metadiscourse awareness.

Taking the above purpose into consideration, the present study addressed the following research questions:

1) Is portfolio assessment an effective way to improve English language learners' oral performance?

2) To what extent do EFL students develop metadiscourse markers knowledge (awareness) in oral performance by the treatment of portfolio assessment?

3. Methodology

This study is intended to determine the effect of portfolio assessment on the development of metadiscourse markers awareness in EFL learners' oral performance. The design for this study is Quasi-experimental in nature, since the classroom groups are already in place and had to be intact. In order to have a strong quasi-experimental design, internal threats to validity were controlled by use of pretesting. To be confident that there were no significant difference among the subjects of the Experimental Group (EG) and Control Group (CG) regarding the variables under investigation, both groups were tested at the beginning of the experiment. Both group's knowledge of metadiscourse markers were addressed in this testing.

4. Participants

The participants in this study were 60 students, (30 males and 30 females) and aged between 20 to 23 years. They were majoring in English language teaching at the University of Najafabad. They were randomly divided into two groups: a control group and an experimental one. The control group was taught and assessed based on the more traditional

method used before, but the experimental one was exposed to a portfolio based teaching and assessment method with the metadiscourse awareness in focus.

5. Data Collection Procedures

The 60 selected students were randomly divided into two groups. They were tested in advance to be sure about their homogeneity in oral performance. The instructional methods, textbooks, and assignments in both the experimental and the control groups were identical, and both groups were taught by the same instructor. Students in portfolio-based group were advised on format of portfolio- based instructional procedures. These portfolios contained all samples of their works including both classroom assignments and homework as well as self-assessment records. Self-assessment records were the students' self-ratings of their own works based on the criteria defined by the teacher at the beginning of the treatment. The study extended over a timeframe of 16 weeks. Scores obtained on the scale of 1-6 comprised students' course grades. In compliance with the class syllabus, the instructor taught students how to do their translation and how to use the metadiscourse markers to improve their oral performance.

In order to eliminate instructor bias resulting from factors other than the work submitted (e.g., attendance, participation, student disposition toward the class or teacher) (Baker, 1993) an independent scorer were invited to evaluate final assessment of portfolios.

6. Results

As mentioned above, this study aimed at examining the effect of portfolio assessment on the participants' achievement in oral performance in academic context, especially focusing on metadiscourse knowledge (awareness). It should be noted that, portfolio assessment here was used as a technique for teaching speaking to EFL students and it was not considered in its broader sense as a separate system of assessment.
In order to investigate the null hypotheses, a series of statistical procedures such as *t*-test and chi-square were run. The descriptive statistics

were computed in order to provide the average mean scores for both experimental and control groups. The obtained data are described and summarized in table 1 below.

Table 1: Descriptive statistics: pre-test and post-test concerning CG and EG

Groups	Mean	SD
CG Pre-test	19.5	4.32
CG Post-test	20.10	**3.40**
EG Pre-test	19.7	4.54
EG Post-test	22.53	3.56

The descriptive statistics in Table 1 indicates that there is a difference between the mean scores of the pre-test and the post-test in the control and experimental groups. The participants' performance on the second sample was better than their performance on the first one. Specifically, Table 1 reveals that the participants' oral performances in the experimental group seem to have improved after being exposed to the treatment in the classroom. In order to investigate the first null hypothesis, an independent sample *t*-test was run. The *t*-observed value for the comparison of the control and experimental groups' mean scores on the test is 2.70. As shown in Table 2, this amount of *t*-value exceeds the *t*-critical. It can be claimed that there is a significant difference between the two groups mean scores on the test, so the first null hypothesis is rejected. The findings imply that the experimental group with a mean of 22.53 outperformed the control group whose mean is 20.10.

Table 2: A Comparison of the Post-test Mean Scores: Control and Experimental Groups

Observed t	D.F.	t-Critical
2.70	58	2.01

N = 60; *p = < 0.05

As shown in Table 2, this amount of *t*-value exceeds the *t*-critical. It can be claimed that there is a significant difference between the two groups mean scores on the test, so the first null hypothesis is rejected.

As a result, based on the analyses, it can be said that the application of portfolio assessment as the teaching procedure in oral performance proved to be significantly effective and helped the participants in their overall speaking in the experimental group.

As for the second question, which aimed at investigating the effect of portfolio assessment on metadiscourse awareness in the participants' oral performance, the chi-square test was applied. This statistical procedure was used in order to examine the frequency of metadiscourse markers used correctly and appropriately by the participants.

Through the analysis of the oral performances of the students in this study, the number of metadiscourse markers appropriately used by the participants was counted. The obtained data were described according to frequencies and percentages (See Table 3 below). Then, a comparison was made between frequencies and percentages of metadiscourse markers in order to observe if there would be any meaningful difference between the control and experimental groups in terms of proper use of them. To do so, one chi-square was run to find the difference between CG and EG concerning the appropriate use of metadiscourse markers.

Table 3: Frequency and percentage of metadiscourse markers

Groups- Meta-discourse	Total frequency	Percentage
CG	331	46.2
EG	385	53.8

The chi-square observed value for comparing the experimental and control groups' appropriate use of metadiscourse markers is 4.07. As Table 4 below depicts, this amount of chi-square value exceeds the critical value of chi-square, i.e., 3.84. It can be claimed that there is a significant difference between the numbers of metadiscourse markers produced by the two groups. As shown in Table 3 above, the experimental group pro-

duced 53.8% of correct metadiscourse markers while the control group produced 46.2%.

Table 4: Chi-square metadiscourse markers

Observed chi-square	D.F.	Critical chi-square
4.07	1	3.84

N = 60; *p = 0.05

7. Discussion and Conclusions

Concerning the first null hypothesis, the findings confirmed a significant difference between experimental and control groups. That is, portfolio assessment affected the participants' overall oral performance to a large extent in the experimental group. By further observing the portfolio assessment, it can be said that it not only provides improved information about students' achievement in oral performance but also makes a positive influence on teaching and student learning. This explanation confirms the argument by Hancock (1994) and Omalley & Valdez Pierce (1996) that alternative assessment has a useful backwash effect on teaching and learning. The results can also be explicated in the sense raised by Genesee and Upshur (1996). That is, using portfolio assessment in second language classroom can have a very specific focus, such as speaking, or broad focus that includes examples of all aspects of language development. Furthermore, step by step observation of the portfolios of the students in the experimental group indicates that the holistic ratings they received after the instructor's assessment were highly correlated to their scores obtained on the last oral performance of them as posttest. This finding implies that participants in the experimental group had a significant achievement in their oral performance through using portfolio assessment.

It is noteworthy that the metadiscourse markers tuition seemed to have activated the treatment class to deploy or sustain in using a wider spectrum of markers as compared with that of the control class. This confirms the value of metadiscourse markers tuition in raising learners'

awareness in trying more types of markers. To further enhance the effectiveness of strategies-based instruction, it may, therefore, be desirable that the teacher focuses on helping learners explore a variety of strategies used by their classmates as this may motivate them to deploy not just their own, existing non-target strategies but try out new strategies recommended by their peers. This way, they may become better 'orchestrators' of a range of strategies at their disposal (Macaro & Erler 2008).

Moreover, through qualitative analysis of the participants' oral performance obtained at the end of the experiment, a number of interesting points concerning the use of metadiscourse markers can be raised. It was found that good oral performance, which favored higher scores, included more metadiscourse markers.

Based on findings of this study and the related discussion, it can be concluded that:

1) The use of portfolio assessment in the classroom in an EFL context affected participants' achievements in their overall oral performances.

2) The metadiscourse awareness of the participants in the experimental group proved to be significantly enhanced compared with the control group.

Two immediate implications are implied by the findings obtained in this study. First, the application of alternative assessment procedures such as portfolio assessment in classroom can be highly beneficial; this is when assessment is integrated with instruction. Second, portfolio assessment is really an authentic assessment and is a productive and useful tool for assessing the students' progress in class performance.

In fact, findings of this study would propose an integrative model of assessment for classroom application with performance testing such as speaking. This can be a significant starting point toward the integration of instruction and assessment. In order to come to sound and unbiased decisions regarding the learners' classroom language behaviors, language testers and assessors should move toward a multi-level system of evaluation that can provide multiple sources of information. This has been, indeed, the concern of most researchers at the turn of the century

(e.g., Teasdale & Leung, 2000; Shohamy, 2001; Lynch, 2005; Leung & Lewkowicz, 2006; Lam & Lee, 2009). In other words, educators would require both quantitative information and qualitative description about language ability in order to better understand the meaning of scores obtained by students.

References

Brown, D. H. (2004). *Language assessment:* Principles and classroom practices. Place: Pearson education, Inc.

Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOLQuarterly, 32,* 653-675.

Campell, D., Melenyzer, B., Nettles, D. & Wyman, R. (2000). *Portfolio and performance assessment in teacher education.* Needham Heights, MA: Allyn & Bacon.

Crismore, A. (1984). The rhetoric of social studies textbooks: Metadiscourse. *Journal of Curriculum Studies, 16*(3), 279-296.

Crismore, A. (1989). *Talking with readers: Metadiscourse as rhetorical act.* New York: Peter Lang.

Crockett, T. (1998). *Strategies in learning and using a second language.* London: Longman.

Crosby, C. (1997). Portfolio assessment in the Korean ESL writing classroom. Thai TESOL Bulletin, 10(2). *Retrieved October 13,* 2010, from http://www.thaitesol.org/bulletin/1002/100204.html

Darling-Hammond, L. (1994). Performance-based assessment and educational equity. *Harvard Educational Review, 64,* 5-30.

Defina, A. A. (1992). *Portfolio Assessment: Getting Started.* New York: Scholastic Professional Books.

Douglas, D. (2000). *Assessing language for specific purposes.* Cambridge: Cambridge University Press.

Hut Chings, P. (1990). Learning over time: Portfolio assessment. *AAHE Bulletin (April),* 6-8.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new era of cognitive development inquiry. *American Psychologist, 34,* 906-911.

Fuertes-Olivera P. A., Velasco-Sacristán M., Arribas- Bañño A., Samaniego-Fernández E. (2001). Persuasion and advertising English: Metadiscourse in slogans and headlines. *Journal of Pragmatics, 33,* 148-154.

Genesee, F. & Upshur, J. A. (1996). *Classroom-based evaluation in second language education.* Cambridge: Cambridge University Press.

Hamayan, E. V. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics, 15,* 212-226.

Hamp-Lyons, L. (1996). Applying ethical standards to portfolio assessment of writing in English as a second language. In M.Milanovic and N. Saville (Eds.), *Performance testing, cognition and assessment* (151-162). Cambridge: Cambridge University Press.

Hamp-Lyons, L., & Condon, W. (2000). *Assessing the portfolio: Principles for practice, theory, and research.* Cresskill, NJ: Hampton.

Hancock, C. R. (1994). Alternative assessment and second language study: What and why? *The Ohio State University.* Retrieved from http://www.ERIC_Digests.com

Hutchings, P. (1990). Learning over time: Portfolio assessment. *AAHE Bulletin (April),* 6-8.

Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics, 30,* 437-445.

Hyland, K. K. (1999). Talking to students: Metadiscourse in introductory course books. *English for Specific Purposes, 18*(1), 3-26.

Hyland, K. (2004). Disciplinary discourses: *Social interactions in academic writing.* Longman/Pearson.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing.* London: Continuum.

Johnson, B. (1996). *Performance Assessment Handbook: Volume 1 Portfolios and Socratic Seminars.* Princeton: Eye On Education, Inc.

Kormos, J. (1999). Simulating conversations in oral proficiency assessment: A conversation analysis of role play and non-scripted interviews in language exams. *Language Testing, 16,* 163-188.

Lam, R., & Lee, I. (2009). Balancing the dual functions of portfolio assessment. *ELT Journal, 63,*4, 1-12.

Lee, K. C. (2001). *Teaching materials and methods of comprehensive activity fields.* Taipei: Shin-Lee.

Leung C., & Lewkowicz, J. (2006). Expanding Horizons and Unresolved Conundrums: Language Testing and Assessment. *TESOL Quarterly, 40*(1), 211-234.

Lynch, B. (2001). Rethinking assessment from a critical perspective. *Language Testing, 18,* 351-72.

Lynch, B. & Shaw, P. (2005). Portfolios, Power, and Ethics. *TESOL Quarterly, 39,* 263-297.

Mandell, A. & Michelson, E. (1990). *Portfolio Development and Adult Learning: Purposes & Strategies.* Chicago: Council for Adult and Experimental Learning.

Murphy, S. (1999). "Assessing Portfolio." In C. Cooper & L. Odell (Eds.), Evaluating writing (pp. 114-136). Urbana, IL: National Council of Teachers of English.

O'Malley, J. & Pierce, L. (1996). *Authentic assessment for English language learners.* US: Addison-Wesley Publishing.

Papajohn, D. (1999). The effect of topic variation in performance testing: case of the chemistry TEACH test for international teaching assistance. *Language Testing, 16,* 1, 52-81.

Schiffrin, D. (1980). Metatalk: Organisational and evaluative brackets in discourse Social Inquiry: *Language and Social Interaction,* 50, 1980.

Shaaban, K. (2001). Assessment of young learners. *English Teaching Forum, 39,* 4, 16-22.

Shohamy, E. (2001). Democratic assessment as an alternative. *Language Testing, 18,* 4, 373-391.

Ruetten, M. K. (1994). Evaluating ESL students' performance on proficiency exams. *Journal of Second Language Writing, 3,* 85-96.

Song, B., & August, B. (2002). "Using portfolio to Assess the Writing of ESL Students: A Powerful Alternative?" *Journal of Second Language Writing, 11*, 49-72.

Teasdale A., & Leung, C. (2000). Teacher assessment and psychometric theory: a case of paradigm crossing? *Language Testing, 17,* 2, 163-184.

Tierney, R. J., Carter, M. A., & Laura, E. D. (1991). *Portfolio Assessment in the Reading-Writing Classroom.* Norwood: Christopher-Gordon Publishers, Inc.

Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College Composition Communication, 1*, (36), 1-15.

Vande Kopple, W. J., & Metadiscourse, d. (2002). In E. Barton and G. Stygall (Eds.), *Discourse studies in composition,* (91-113).

Williams, J. M. (1981). Style: Ten lessons in clarity and grace. Glenview, Illinois: Scott, Foresman and Co.

Williams, J. D. (1998). *Preparing to teach writing.* Mahwah, NJ: Lawrence Erlbaum.

Williams, J. D. (2000). Identity and reliability in portfolio assessment. In B. Sun stein & J. Lovell (Eds.), *The portfolio standard,* (135-148).

Yang, N. D. (2003). Integrating portfolios into learning strategy-based instruction for EFL college students. *IRAL, 41*(4), 293-317.