The Effect of Dynamic Assessment on EFL Learners' Acquisition of Request and Apology

Z. Tajeddin

Associate Professor
Allameh Tabataba'i University, Tehran
email: zia_tajeddin@yahoo.com

F. Tayebipour *

Ph.D. Candidate
Allameh Tabataba'i University, Tehran
email: tayebipour@yahoo.com

Abstract
Originating from Vygotsky's Sociocultural Theory, dynamic assessment (DA) proposes a novel approach to second language acquisition (SLA) research according to which a dialectical relationship is envisaged between instruction and assessment. Although DA has been applied to some areas of SLA, there are areas, such as interlanguage pragmatics (ILP), that have been neglected. To address the issue, 40 university students of two proficiency levels were selected and assigned randomly to 1 of 4 groups consisting of two DA and two Non-DA groups. Each group received instruction concerning how to use appropriate request and apology strategies in the English language. However, following Lantolf and Poehner (2011), DA groups received ZPD-sensitive feedback whereas NDA groups received no more feedback but were assessed according to their independent performance. All participants completed a pretest, a posttest, and a delayed posttest of request and apology speech acts that were rated by two native speakers of English on a 6-point Likert scale. The findings revealed that DA groups outperformed NDA groups and that DA groups of both high and low proficiency levels differed significantly from pretest to posttest to delayed posttest. However, high NDA group didn't show such a difference. The results also showed no interaction between proficiency and instruction indicating that instruction, but not proficiency, had a significant effect on posttest and delayed posttest performance of the students. The findings may be revealing in that they support DA and its applicability to ILP instruction.

Keywords: interlanguage pragmatics, dynamic assessment, non-dynamic assessment, zone of proximal development

1. Introduction

According to some interlanguage pragmatics (ILP) researchers (e.g. Bardovi-Harlig & Dörnyei, 1998; Bardovi-Harlig & Hartford, 1996, cited in Takimoto, 2006), there is a need for instruction of L2 pragmatics on the assumption that many aspects of pragmatic competence may not develop without instruction. For these researchers, the type of instruction does not matter that much. Some others (e.g. House, 1996; Rose & Ng, 2001; Takahashi, 2001; Tateyama, 2001) believe in a particular type of instruction, i.e. explicit instruction, on the grounds that providing learners with explicit metapragmatic information rather than with implicit input results in much better performance. For example, Takahashi (2005) contends that explicit instruction leads to more effective learning as it is associated with deeper level of noticing by and higher level of awareness of second/foreign language learners.

The point is that results of the studies carried out so far have been conflicting in that they have been either explicit or implicit-oriented. Such being the case, it seems necessary to apply more promising approaches to ILP instruction where explicit-implicit is dealt with not in terms of polarity but on the basis of modality making the instruction more learner-friendly and ZPD-sensitive. The purpose of this study was to measure the effects of dynamic assessment vs. non-dynamic assessment on EFL learners' acquisition of request and apology speech acts.

2. Literature Review

2.1 DA approach to instruction

DA is drawn from Vygotsky's Sociocultural Theory of Mind and its seminal concept of the zone of proximal development. It refers to an integrated process of teaching and assessment in an unpredictable, ever-changing, dynamic manner. In Williams and Burden's (1997, p.42) words, DA is a process where "assessment and learning are seen as inextricably linked and not separate." In the opinion of Lidz and Gindis (2003, p. 99), DA is an "approach to understanding individual differences

and their implications for instruction that embeds intervention within the assessment procedure." Finally, according to Haywood and Lidz (2007, p. 1) DA is "an interactive approach to conducting assessments… that focuses on the ability of the learner to respond to intervention."

As these definitions may indicate, DA is against any dualistic view of instruction and assessment because it erases well-entrenched, long-lasting boundaries between instruction and assessment, unifies the two, and prepares the grounds for more learning-friendly interactions between an examiner (mediator) and his/her examinees (learners). According to Poehner (2008), DA proposes an essentially different assessment paradigm from what we may have traditionally come to think of it. This paradigm is characterized by features such as process-oriented, interactive, and ZPD-sensitive making DA not only an instruction with a focus on what an individual has learnt up to now but a future-oriented instruction that explores the individual's potentialities for further learning.

According to Caffrey, Fuchs, and Fuchs (2008), DA differs from NDA, i.e. non-dynamic, traditional, static assessment, in many respects such as relationships between an examiner and learners, nature of the feedback provided, and nature of learning. As for the relationships, they point out that while in NDA the atmosphere between an examiner and the examinees has been a threatening one, in DA it is a supportive one that focuses on a joint activity towards learners' learning. With regard to feedback, whereas in NDA no or very little learning-friendly feedback may be provided, in DA feedback has to be fine-tuned to match person(s)-specific ZPDs. Finally, while in NDA the emphasis has been exclusively on the product of learning, in DA it is, over and above the product, on the process(es) of learning.

For these reasons, in almost all ZPD-oriented research, an indispensable element has been the variable of relevant help. In this respect, Aljaafreh and Lantolf's (1994) mechanisms of effective help may be noteworthy here. The first mechanism is graduation which means that help should be offered in a gradual process. The second mechanism

is called contingency implying that feedback should be provided at an exact time. The third is that help should be given in a two-way street and in an interactive, dialogic manner. These mechanisms are indeed the building blocks of DA and have been applied to many academic disciplines.

DA has been interpreted differently and various models have been proposed: Budoff's (1974) Test-Train-Test Assessment, Feuerstein's (1979) Learning Potential Assessment Device, Carlson and Wiedl's (1978) Testing-the-Limits Assessment, Vygotsky's (1978) Graduated Prompting Assessment, and A Continuum of Assessment Model-Mediated and Graduated Prompting (Bransford, Delclos, Vye, Burns, & Hasselbring, 1987; Burns, Haywood, Delclos, & Siewart, 1987; Vye, Burns, Delclos, & Bransford, 1987) (all cited in Jitendra & Kameenui, 1993). However, the point is that, as Kozulin and Garb (2001) state, what makes these interpretations similar is their reliance on test-teach-test paradigm; what makes them different is the type of instruction given between pre-and post-tests.

What Kozulin and Garb (2001) imply by difference in the type of instruction given is reminiscent of Lantolf and Poehner's (2004) division of labor in terms of which DA is divided into interventioinist and interactionist. Poehner (2005) states that interventionist DA refers to a psychometric-oriented approach where administration procedures are standardized and learning products are quantified to be measurable afterwards whereas interactionist DA is not concerned about standardization of the interaction, quantification of the learning and its measurement. Instead, in interactionist DA, feedback is emergent rather than preplanned and learning is interpreted rather than measured. This means that "type of instruction" provided makes the difference with the interventionist one being more practical with group of learners and the interactionist one more practical with individual learners.

2.2 Previous research on DA in language education

Although DA is a newcomer in SLA, it is by no means the case in other disciplines. As Haywood and Lidz (2007, p. 2) contend, DA "is no longer

a new approach to psychological and educational assessment [as] some of its current applications have been around for more than a half century." Thus, Haywood and Lidz enumerate the studies conducted on the basis of DA such as the one on reading by Gettinger (1984); on mathematics by Jitendra and Kameenui (1993); and on speech and language by Kozulin and Garb (2001), to name a few.

Nonetheless, L2 DA studies in general and ILP DA studies in particular do not have such a robust literature. That is why Poehner (2008, p. 5) states that "to date, few studies have examined L2 performance from a Dynamic Assessment perspective, although the growing interest in Vygotskyan theory among applied linguists has led to some exploration of how DA principles might be used in L2 contexts." However, the following studies fall within L2 DA studies: Aljaafreh and Lantolf (1994), Nassaji and Swain (2000), Kozulin and Garb (2001), Poehner (2005), and Ableeva (2010).

Aljaafreh and Lantolf's (1994) study was a pioneering study where a mediator collaborated with three learners who were trying to produce grammatical features such as use of tense, modal verbs, prepositions, and articles in their compositions. The mediator met them individually in the writing class and made revisions on their previously written work. This revision was done on the basis of a regulatory scale which would change from most implicit to most explicit. As Poehner (2005) rightly notes, although Aljaafreh and Lantolf's (1994) study was not specifically framed as DA, it is possible to consider the study within the framework of DA due to the fact that the mediator attempted to co-construct a ZPD with the learners, interacting with them to diagnose areas of difficulty and to help them gain control over the relevant structures.

Nassaji and Swain (2000) followed Aljaafreh and Lantolf's (1994) study and attempted to find out whether ZPD-sensitive mediation was necessary to improve performance or if any kind of mediation could be sufficient to help the learners move beyond what they could do independently. They were also interested in knowing which one of the mediations was more effective to promote development. So, Nassaji and

Swain (2000) paired a mediator with two ESL learners. With one, the mediation was ZPD-sensitive; with the other it was quite random, i.e. without any attempt to adjust the level of assistance to the learner's responsiveness. The results showed that the one receiving ZPD-sensitive mediation had actually been less accurate when independently producing the initial composition but nevertheless showed greater improvement as a result of the mediation given outperforming the non-ZPD student on the final composition task.

Another study is that of Kozulin and Garb (2001) which was named interventionist DA after the pre-test, post-test format it followed. The results were clearly in favor of DA in that it proved to be significantly effective in promoting learners reading comprehension skill. Although Poehner (2005) argues against the dynamic nature of Kozulin and Garb's (2001) study, the study obviously falls within DA framework though obviously on the interventionist side of the continuum.

In another study by Poehner (2005), the focus was on learners' oral abilities. Six advanced students of L2 French were asked to orally construct a series of narratives in French based on short videos. The learners had to create the first narrative independently while in a second narrative they received some help. The results of these two narratives were used to develop a program of individualized instruction. Following the instruction, participants' learning was assessed by the repetition of the original assessment tasks and introduction of some newly developed but similar tasks. According to Poehner, "the findings suggest that DA is an effective means of understanding learners' abilities and helping them to overcome linguistic problems. The approach is especially relevant to L2 classrooms as a method for rendering formative assessment practices more systematic" (p. iv). Finally, Ableeva (2010) investigated the effects of DA on improving listening comprehension of students learning French as a foreign language and compared the results to a traditional test of listening comprehension. The results indicated that DA illuminated the sources of poor performance that were hidden during traditional assessments. The results also showed that, through interactions in the

ZPD, DA was able to establish not only the actual level of learners' listening ability but also to diagnose/assess their potential development while at the same time promoting this development.

As it may have been noticed, L2 DA studies seem to have been growing in depth and breadth. However, such is not the case with ILP making it necessary to carry a piece of research in this particular domain.

3. Purpose of the Study

A review of literature indicates that DA has been a promising strand of research. That is why authors such as Poehner (2005, p. iv) maintain that "DA should be taken seriously by Applied Linguistic researchers interested in language assessment, teaching, and learning." The point is that no study, except for few exceptions (Tajeddin, Alemi, & Pakzadian, 2011; Tajeddin & Mosleh, 2011) has ever been conducted on ILP development from the perspective of DA. Thus, in line with Ohta's (2005) seminal article where she encourages SLA researchers to apply Vygotsky's sociocultural theory and his zone of proximal development to ILP instruction, this study was carried out to find out the effects of DA on ILP instruction through addressing the following questions:

1. Is there any significant difference between the effects of DA vs. NDA on low and high proficiency EFL learners' acquisition of the speech acts of request and apology?

2. Is there any significant difference within each of DA and NDA groups in terms of their acquisition of the speech acts of request and apology from pretest to posttest to delayed posttest?

3. What are the effects of instructional approaches and learners' levels of proficiency on DA groups and NDA groups' acquisition of the speech acts of request and apology from pretest to posttest to delayed posttest?

4. Method

4.1 Participants

As many as 40 participants were selected from among 160 male and female undergraduate university students who majored in Teaching English as a Foreign Language (TEFL). They were assigned randomly into 4 groups, each consisting of 10 learners. Two groups were low proficiency students selected from among first-semester students and two groups were high proficiency students selected from among eighth-semester students. Then, based on the design of the study, 20 symmetrical scores from semester 8 students and 20 from semester 1 students were selected. They were then assigned randomly into DA and NDA groups, i.e. two DA groups (high vs. low) and two NDA groups (high vs. low). The mean age of the participants was 22, their first language was Persian, and none had studied the English language abroad. Attempts were made to select as homogeneous participants as possible because, according to Haywood and Lidz (2007, p. 223), "homogeneous grouping helps to minimize variability in work efficiency and the resulting boredom and off-task behavior that can be expected when some individuals finish with a given part of the task well before others do."

4.2 Instruments

Two instruments were employed: (1) a general proficiency test, and (2) a written discourse completion test. As for the former, Oxford Quick Placement Test (Geranpayeh, 2003) was administered. The test consists of three parts: Part one (1-40) deals with simple grammar and vocabulary items. Part two (40-60) is concerned with a bit more difficult multiple choice items and cloze test. Part three comprises a writing section where candidates are required to write a well-organized paragraph of 150-200 words. From the three parts, only the first was administered due to the nature of the test which requires second and third parts to be administered only when the testees are able to answer more than 35 items out of 40 items correctly. The rational for administering the test to both high and low level students was to ensure that there was a significant

difference between the two groups before beginning the treatment sessions.

The second test was a discourse completion test (DCT) about request and apology speech acts. The test consisted of 16 items, i.e. 8 items on request and 8 items on apology speech acts. The items were given to the participants to be answered as pretest, as posttest and as delayed posttest. They required the participants to read short descriptions in English of some situations and write what they would say in the English language for each situation considering the interlocutors' power and distance.

4.3 Treatment materials

Treatment materials consisted of 12 discourse completion test (DCT) items, including 6 items about request and 6 about apology speech acts. The items were answered by both DA and NDA participants during the treatment sessions. However, following Lantolf and Poehner's (2004) interventionist model, DA groups answered the items while they were provided with preplanned mediation whereas NDA groups received no such mediation but were tested on the basis of their independent performance in answering the items. To provide DA groups with preplanned mediation, Lantolf and Poehner's (2011) scale was adopted on the basis of which 8 forms of mediation are provided depending on each and every particular learner's responsiveness. If a learner's response is correct, the mediator gives no mediation. However, if the learner's response is not correct and/or appropriate, the mediator moves one or more steps further till the last step where s/he has to provide the learner with full explanations. In fact, the forms of mediation given are as follows: 1.Teacher pauses 2. Teacher repeats the whole phrase questioningly 3. Teacher repeats just the part of the sentence with the error 4. Teacher asks: what is wrong with that sentence? 5. Teacher points out the incorrect word 6. Teacher asks either or questions 7. Teacher identifies the correct answer and 8. Teacher explains why. As it may have been noticed, step 1 is the most implicit while step 8 is the most explicit mediation which may be provided. To run DA and NDA sessions, the 4 groups had to have 4

meetings of 30 minutes, 4 days a week and for a total of 6 weeks. In other words, treatment sessions lasted for 6 weeks, i.e., 3 weeks for request and 3 weeks for apology speech acts in an every other week order.

4.4 Data collection

The data for the study were collected using the two tests mentioned above, i.e. the general proficiency test and the discourse completion test (DCT). General proficiency test results provided us with scores while the DCT results provided us with some sentences produced by the learners and rated by two native speakers of English on the basis of a 6-point Likert scale adopted from Taguchi (2006). The scale ranged from 0 to 5: (1) 0 meant no performance at all; (2) 1 meant very poor performance, i.e. expressions were very difficult to understand; (3) 2 meant poor performance, i.e. due to the interference from grammatical and sociolinguistic errors correctness and appropriateness were difficult to determine; (4) 3 meant fair performance, meaning that expressions were only somewhat correct and appropriate; (5) 4 meant good, i.e. expressions were mostly correct and appropriate; and (6) 5 meant excellent performance, meaning that expressions were fully correct and appropriate for the situation. The DCT data were collected as DA as well as NDA participants attempted to answer the 16 pragmatic items as the pretest, posttest and delayed posttest. The pretest was administered one week after a preliminary instruction was given on the two speech acts, the posttest was given one week after the treatment sessions were over, and after an interval of two weeks the delayed posttest was administered.

4.5 Data analysis

The data collected through the general proficiency test and the discourse completion test (DCT) were analyzed using the statistical software SPSS. Descriptive statistics, including minimum, maximum, mean, and standard deviation were calculated. As for inferential statistics, an independent samples t-test, a repeated measure ANOVA, and a Two-way ANOVA were employed. The reasons for utilizing parametric tests were the four

assumptions of interval data, independence, and normality and homogeneity of variances (Field, 2009) which were met in the study. The present data were interval in nature and none of the subjects participated in more than one group. Thus, the assumptions of interval data and independence were met. However, the assumptions of normality and homogeneity of variances should also be probed empirically. In order to meet the assumption of normality, the ratios of skewness and kurtosis over their respective standard errors should be within the ranges of + / - 1.96. As displayed in Table 1, under the columns of normality tests, the students' scores on the pretest, posttest and delayed posttest of Dynamic Assessment (DA) and Non-Dynamic assessment (NDA) groups enjoyed both normality of skewness and kurtosis.

Table 1. Normality test of pretest, posttest and delayed posttest by group

			INSTRUCTION	Number	Skewness Statistic	Std. Error	Normality Test of Skewness	Kurtosis Statistic	Std. Error	Normality Test of Kurtosis
DA	Low		Pretest		0.889	0.687	1.29	0.119	1.334	0.089
			Posttest	10	-0.458	0.687	-0.66	-0.071	1.334	-0.053
			DPosttest		-1.318	0.687	-1.91	1.890	1.334	1.42
	High		Pretest		-0.348	0.687	0.506	0.224	1.334	0.168
			Posttest	10	0.388	0.687	0.565	-0.701	1.334	0.525
			DPosttest		-1.343	0.687	1.95	2.001	1.334	1.5
ND A	Lo w		Pretest	10	-0.216	0.687	-0.314	-0.658	1.334	-0.932

INSTRUCTION		Number	Skewness		Normality Test of Skewness	Kurtosis		Normality Test of Kurtosis
			Statistic	Std. Error		Statistic	Std. Error	
	Posttest		-0.114	0.687	-0.166	-0.836	1.334	-0.627
	DPosttest		-1.107	0.687	-1.611	1.047	1.334	0.784
	Pretest		-1.111	0.687	-1.62	2.353	1.334	1.76
High	Posttest	10	-1.292	0.687	-1.88	0.812	1.334	0.608
	DPosttest		-1.202	0.687	-1.75	0.800	1.334	0.599

Likewise, as displayed in Table 2, under the columns of normality tests, the students' scores on the proficiency test enjoyed both normality of skewness and kurtosis.

Table 2. Normality test of proficiency by level

Level	Number	Skewness		Normality Test of Skewness	Kurtosis		Normality Test of Kurtosis
		Statistic	Std. Error		Statistic	Std. Error	
Low	20	1.003	0.512	1.95	1.384	0.992	1.39
High	20	-0.451	0.512	0.88	0.374	0.992	0.37

Therefore, based on the results of Table 1, Table 2, and two other tests of normality, i.e. Kolmogorov-Smirnov and Shapiro-Wilk in terms of which *p*-values were greater than significance level ($p<0.05$), it can then safely be concluded that the assumption of normality is met and

parametric tests can be employed to probe the research questions. The assumption of homogeneity of variances, however, will be tested later when reporting the main results. With regard to inter-rater reliability of pretests, posttests, and delayed posttests of high and low groups, the following indices were obtained respectively: high group (pretest r =.88; posttest r=.79; and delayed posttest r=.93) and low group (pretest r =.95; posttest r=.92; and delayed posttest r=.94). The total reliability of the proficiency test was K-R21: .85.

5. Results

The results are presented in two sections: While section one reports the results on the differences between high and low groups in terms of language proficiency, section two is concerned with the findings on the three research questions.

5.1 Group differences by proficiency

Table 3 shows the descriptive statistics of the language proficiency test in terms of the level of proficiency and the instructions, i.e. DA vs.NDA. As the table shows, the mean scores for the high and low levels of DA group were 25.7 and 17.7, respectively. The mean score for the high NDA group was 24.9, and that for the low NDA group was 17.9. The total mean scores of language proficiency for high and low groups were 25.3 and 17.8, respectively.

Table3. Descriptive statistics of the language proficiency test

Level	Instruction	N	Minimum	Maximum	Mean	Std. Deviation
High	DA	10	16.00	37.00	25.7000	6.25478
	NDA	10	9.00	32.00	24.9000	8.04777
	Total	20	9.00	37.00	25.3000	7.02702
Low	DA	10	9.00	32.00	17.7000	6.41266
	NDA	10	9.00	32.00	17.9000	6.20842
	Total	20	9.00	32.00	17.8000	6.14389
Total	DA	20	9.00	37.00	21.7000	7.40626
	NDA	20	9.00	37.00	21.4000	7.86331
	Total	40	9.00	37.00	21.5500	7.54117

To see if there was any significant difference between the mean scores of high and low proficiency groups on the proficiency test, an independent t-test was run. The t-observed value was 3.593 (Table 4). This amount of t-value is greater than the critical value of 2.02 at 38 degrees of freedom. Based on these results, it can be concluded that there was a significant difference between high and low proficiency groups' mean scores on the proficiency test

Table 4. Independent samples *t*-test for the language proficiency (LP) test

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
									Lower	Upper
LP	Equal variances assumed	.436	.513	3.593	38	.001	7.50000	2.08718	3.27473	11.72527
	Equal vriances not assumed			3.593	37.335	.001	7.50000	2.08718	3.27225	11.72775

5.2 Pragmatic difference between DA and NDA groups

The first research question was concerned with the difference between the effects of DA and NDA on low and high proficiency groups. To answer this question, four sub-questions had to be addressed:

1. Sub-question 1: the difference between high DA and NDA in the posttest

2. Sub-question 2: the difference between low DA and NDA in the posttest

3. Sub-question 3: the difference between high DA and NDA in the delayed posttest

4. Sub-question 4: the difference between low DA and NDA in the delayed posttest

Regarding sub-question 1, the mean score in high DA was 4.65 and in high NDA it was 4.05. To investigate the difference between the means of high DA and high NDA on the pragmatic posttest, an independent samples t-test was run. The *t*-observed value was 3.286 (Table 5). This amount of *t*-value is greater than the critical value of 2.23 at 10 degrees of freedom. Based on the results, it can be concluded that there is a significant difference between high DA and high NDA groups' mean scores on the posttest ($p < .008$). Thus it can be concluded that the high DA group outperformed the high NDA group on the pragmatic posttest.

Table 5: Independent samples *t*-test for the pragmatic posttest in high DA and NDA groups

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Post	Equal variances assumed	5.790	.027	3.286	18	.004	.5938	.18353	.21755	.98870
	Equal variances not assumed			3.286	10.179	.008	.5938	.18353	.19518	1.01107

As to sub-question 2, the mean score of low DA was 4.50 and that of low NDA was 3.74. An independent samples *t*-test was run to compare the two groups' mean scores on the pragmatic posttest. The *t*-observed value was 3.765 (Table 6). This amount of *t*-value is greater than the

critical value of 2.17 at 12 degrees of freedom. Based on the results, it can be concluded that there was a significant difference between low DA and low NDA groups' mean scores on the posttest ($p < .003$). Thus it can be concluded that DA has been more effective than NDA in pragmatic acquisition in low-proficiency groups.

Table 6: Independent samples *t*-test for the pragmatic posttest in low DA and low NDA groups

	Levene's Test for Equality of Variances		t-test for Equality of Means					95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
post Equal variances assumed	6.774	.018	3.765	18	.001	.76250	.20250	.33707	1.18793
Equal variances not assumed			3.765	11.664	.003	.76250	.20250	.31988	1.20512

Sub-question 3 was concerned with the difference between high DA and high NDA groups in the pragmatic delayed posttest where the mean score in high DA was 4.85, and that in high NDA was 4.49. To see if the difference between the means of the two groups was significant, an independent samples *t*-test was run. The *t*-observed value was 2.098 (Table 7), not exceeding the critical value of 2.23 at 10 degrees of freedom. As a result, it can be concluded that there was not a significant difference between high DA and high NDA groups' mean scores on the pragmatic delayed posttest ($p<.061$) This means that there was no

difference between DA and NDA in their long-run effects on pragmatic acquisition in high-proficiency groups.

Table 7. Independent samples *t*-test for the pragmatic delayed posttest in high DA and high NDA groups

		Levene's Test for Equality of Variances		t-test for Equality of Means					
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Dpost	Equal variances assumed	6.855	.017	2.098	18	.050	.36250	.17278	-.00051	.72551
	Equal variances not assumed			2.098	10.346	.061	.36250	.17278	-.02075	.74575

Considering sub-question 4, the mean score of the low DA group was 4.63, and that of the low NDA group was 4. This shows that the DA group performed better than the NDA one. To investigate the significance of the mean difference, an independent samples *t*-test was run. The t-observed value was 2.746 (Table 8). This amount of *t*-value was found to be greater than the critical value of 2.2 at 11 degrees of freedom. Thus, the findings evidence a significant difference between low DA and low NDA groups' mean scores on the delayed posttest at *p* < .018, indicating that the delayed effect of DA was greater than that of NDA in low-proficiency groups.

Table 8. Independent samples *t*-test for the pragmatic delayed posttest in low
DA and low NDA groups

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Dpost	Equal variances assumed	6.510	.020	2.746	18	.013	.63438	.23099	.14908	1.11967
	Equal variances not assumed			2.746	11.438	.018	.63438	.23099	.12834	1.14041

With regard to question one, it is concluded that in many respects there were significant differences between the effects of DA vs. NDA on low- and high-proficiency EFL learners' acquisition of request and apology speech acts. However, it should be noted that no significant difference was found between the effects of high DA and high NDA groups in delayed pragmatic performance.

5.3 Pragmatic gains from the pretest to posttest to delayed posttest

The second research question aimed at investigating within-group differences among pretest, posttest, and delayed posttest performance of each of DA and NDA groups. To address the question, four sub-questions were developed:

Sub-question 1: the difference among the pretest, posttest, and delayed posttest in high DA group

Sub-question 2: the difference among the pretest, posttest, and delayed posttest in low DA group

Sub-question 3: the difference among the pretest, posttest, and delayed posttest in high NDA group

Sub-question 4: the difference among the pretest, posttest, and delayed posttest in low NDA group

As to sub-question 1, the mean scores of the high DA group on the pragmatic pretest, posttest, and delayed posttest were 3.79, 4.65, and 4.85 respectively. A repeated measure ANOVA was run to compare the mean scores of the pretest, posttest, and delayed posttest in the high DA group. The *F*-observed value for comparing the mean scores on the pretest, posttest and delayed posttest was 37.845 ($p = .000$) (Table 9). Based on these results, it can be concluded that there is a significant difference from the pretest to posttest to delayed posttest in the high DA group.

Table 9. Tests of within-subjects effects for the high DA group

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	6.335	2	3.168	37.845	.000
	Greenhouse-Geisser	6.335	1.182	5.360	37.845	.000
	Huynh-Feldt	6.335	1.256	5.044	37.845	.000
	Lower-bound	6.335	1.000	6.335	37.845	.000
Error(factor1)	Sphericity Assumed	1.507	18	.084		
	Greenhouse-Geisser	1.507	10.638	.142		
	Huynh-Feldt	1.507	11.305	.133		
	Lower-bound	1.507	9.000	.167		

Concerning sub-question 2, the mean scores of the low DA group on the pretest, posttest and delayed posttest were 2.34, 4.5, and 4.63 respectively, showing a linear increase from the pretest to posttest to delayed posttest. A repeated measure ANOVA revealed the differences between the mean scores on the pretest, posttest, and delayed posttest. The *F*-observed value is 60.719 ($p = .000$) (Table 10). This leads to the conclusion that there was a significant difference among the low DA groups means.

Table 10. Tests of within-subjects effects for the low DA group

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	33.048	2	16.524	60.719	.000
	Greenhouse-Geisser	33.048	1.070	30.892	60.719	.000
	Huynh-Feldt	33.048	1.097	30.131	60.719	.000
	Lower-bound	33.048	1.000	33.048	60.719	.000
Error(factor1)	Sphericity Assumed	4.899	18	.272		
	Greenhouse-Geisser	4.899	9.628	.509		
	Huynh-Feldt	4.899	9.871	.496		
	Lower-bound	4.899	9.000	.544		

Sub-question 3 was related to the high-proficiency NDA group the mean scores of which on the pragmatic pretest, posttest and delayed posttest were 4.04, 4.05, and 4.49, respectively. The developmental trend in the pragmatic acquisition is indicative of a linear progression from the pretest to posttest. A comparison of the mean scores on the pretest, posttest, and delayed posttest in the high NDA group was made based on a repeated measure ANOVA. The *F*-observed value was 3.419 ($p < .055$) (Table 11). It follows that no significant difference was found among the performance of the three tests.

Table 11. Tests of within-subjects effects for high NDA group

	Source	Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	1.322	2	.661	3.419	.055
	Greenhouse-Geisser	1.322	1.761	.751	3.419	.063
	Huynh-Feldt	1.322	2.000	.661	3.419	.055
	Lower-bound	1.322	1.000	1.322	3.419	.098
Error(factor1)	Sphericity Assumed	3.482	18	.193		
	Greenhouse-Geisser	3.482	15.851	.220		
	Huynh-Feldt	3.482	18.000	.193		
	Lower-bound	3.482	9.000	.387		

Finally, in the case of sub-question 4, the mean scores of the low NDA group on the pretest, posttest, and delayed posttest were 2.77, 3.74,

and 4, respectively. Like the other groups, these means evidence a linear trend in pragmatic acquisition. As the result of a repeated measure ANOVA on the mean scores of the pretest, posttest, and delayed posttest in the low NDA group shows, the *F*-observed value was 60.191 (*p* = .000) (Table 12). This significance level confirms the significant difference among the group's means on the three tests.

Table 12. Tests of within-subjects effects for the low NDA group

	Source	Type III Sum of Squares	df	Mean Square	F	Sig.
factor1	Sphericity Assumed	8.320	2	4.160	60.191	.000
	Greenhouse-Geisser	8.320	1.987	4.188	60.191	.000
	Huynh-Feldt	8.320	2.000	4.160	60.191	.000
	Lower-bound	8.320	1.000	8.320	60.191	.000
Error(factor1)	Sphericity Assumed	1.244	18	.069		
	Greenhouse-Geisser	1.244	17.881	.070		
	Huynh-Feldt	1.244	18.000	.069		
	Lower-bound	1.244	9.000	.138		

Altogether, as far as question two is concerned, it can be concluded that there were significant differences between the performances of DA and NDA groups, except for high NDA, from the pretest to posttest and to delayed posttest.

5.4 The effects of instruction and level of proficiency on pragmatic gains

The purpose of the third research question was to explore the immediate and delayed effects of instruction and level of proficiency on pragmatic gains. To address the question, the immediate and delayed effects of DA and NDA were studied. As to the immediate effects of instruction and level of proficiency, the highest mean was observed in the high proficiency group which received the DA-based instruction (4.65). By contrast, the NDA group falling within the low-proficiency group had the poorest performance on the immediate pragmatic test (3.74). In addition, in both proficiency levels, DA groups outperformed NDA ones (4.5766>3.8938).

A two way ANOVA functioned to measure the effects of instruction and level of proficiency on the posttest performance of the groups. The

F-observed value was .34 (*p*<.563) (Table 13). Based on the result, it can be concluded that there was not any significant interaction effect on the posttest results. Furthermore, the *F*-observed value for the effect of level of proficiency on the immediate posttest was 2.749 (p<.106), which means that there was not any significant effect by the level of proficiency on the posttest performance of the groups. Nevertheless, the *F*-observed value for the effect of instruction on the posttest was 24.97 (p = .000), which shows the effect of instruction on posttest performance.

Table 13. Tests of within-subjects effects for the immediate pragmatic posttest by instruction and proficiency level

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5.239[a]	3	1.746	9.353	.000
Intercept	717.462	1	717.462	3842.530	.000
instruction	4.662	1	4.662	24.970	.000
level	.513	1	.513	2.749	.106
instruction* level	.064	1	.064	.340	.563
Error	6.722	36	.187		
Total	729.423	40			
Corrected Total	11.961	39			
a. R Squared = .438 (Adjusted R Squared = .391)					

With regard to the delayed effects of instruction and level of proficiency, the highest mean score was observed in the high proficiency group which received the DA-based instruction (4.86). By contrast, the NDA group falling within the low-proficiency group had the poorest performance on the immediate pragmatic test (4.00). In addition, in both proficiency levels, DA groups outperformed NDA ones (4.7484>4.25).

A two way ANOVA was employed to test the effect of instruction and level of proficiency on the delayed posttest performance of the groups. The *F*-observed value for interaction effect of instruction and level of proficiency on the delayed posttest was found to be .888, which was significant at *p*<.352 (Table 14). This result documents the effect of interaction of high language proficiency and DA-based instruction on delayed pragmatic gains. Concerning the effect of each of the two variables in separation, the *F*-observed value for the effect of level of proficiency on the delayed posttest was 6.154, so it was significant at p <

.018. In the same vein, the *F*-observed value for the effect of instruction on delayed posttest was 11.943, significant at *p*<.001.

Table 14. Tests of within-subjects effects for the delayed pragmatic posttest by instruction and proficiency level

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3.949[a]	3	1.316	6.328	.001
Intercept	809.719	1	809.719	3892.377	.000
instr	2.484	1	2.484	11.943	.001
level	1.280	1	1.280	6.154	.018
instr * level	.185	1	.185	.888	.352
Error	7.489	36	.208		
Total	821.157	40			
Corrected Total	11.438	39			
a. R Squared = .345 (Adjusted R Squared = .291)					

To sum up, it can be concluded that instruction has been effective on immediate and delayed pragmatic posttests. However, it should be noted that level of proficiency was not effective on the posttests results, though it was effective on delayed posttest performance.

6. Discussion

In this study, three findings were obtained. The first was that DA groups performed significantly better than NDA groups on the pragmatic posttest. However, only in the delayed posttest results, the difference between high DA and high NDA was not significant. This should by no means imply that DA has had the same effects as NDA on EFL learners' acquisition of request and apology because when it comes to their performance on the posttest there is obviously a significant difference between DA and NDA. In this regard, it may be argued that although delayed posttest results may also be important, what matters most is posttest results because delayed posttest results are more a matter of recall rather than learning. Such being the case, it is true to say that DA has been more effective than NDA.

As for the explanation that might be given for the delayed posttest performance of the students of high DA and high NDA, one may say that DA may be in need of more time to raise learners' awareness of learning

so that their learning may last longer. The point is that DA learners had only 3 sessions for each particular speech act to develop and this period may not have been enough for learning to last long, meaning that more sessions may be needed as far as long-term learning is concerned.

The second finding was that performance of DA groups and NDA groups from the pretest to posttest to delayed posttest was significantly different. All groups except for one, i.e. high NDA, made significant changes from the pretest to posttest to delayed posttest. This developmental trend once again highlights DA's effect on learning because the results clearly indicate how effective DA has been to bring about stable changes from pretest to posttest to delayed posttest. As to pragmatic gains by the high NDA group, it should be noted that this group actually received no fine-tuned instruction so there is no wonder if they made no significant changes from the pretest to posttest and to delayed posttest. In other words, being within NDA group, the participants had no opportunity to be provided with appropriate feedback. Moreover, since they were semester 8 students, their previous learning may have been active as an intervening variable in the sense that the previous learning may have been responsible for the way they performed. Now a question may be raised as to why low NDA group did not show the same results. The answer might be that since participants of low NDA group were true beginners, they may have been more motivated to learn although they did not receive any fine-tuned feedback either. In this respect, it should be pointed out that although no motivation tests were used to measure the low NDA participants' motivation, when it comes to a comparison between low and high NDA groups' performance, low NDA group's results prove that they were more cooperative and more willing to learn English pragmatics in spite of the fact that their proficiency level was much lower than their high NDA counterparts. In other words, since they were true beginners they may have been more interested by nature, more willing, and more motivated to learn the materials they were exposed to while this was not necessarily the case with high NDA. One reason may be that the high NDA group had

already had an opportunity to learn for seven semesters and their previous learning may have been an intervening variable working against the instruction while this was not the case with the low NDA group.

With respect to the third finding, i.e. no interaction between instruction and level of proficiency, one may argue that once again DA, as the independent variable, has been effective on posttest results.

Regarding the above findings, the following explanations might be given. As for DA groups' better performance, ZPD-sensitive interactions may have been responsible for DA groups' superiority over NDA groups. In these interactions every learner seems to have had two types of opportunity to learn: One as a direct interlocutor and one as an indirect one. In other words, each learner may have benefited from two types of interactions: One when s/he was addressed directly and one when s/he was not addressed directly but through listening to others. In this regard, it is important, as Poehner's (2009) notes, to bear in mind that in any linguistic interaction based on DA, one may be either the first or the second interactant. When a learner is the first, s/he receives fine-tuned assistance directly. However, when the same learner is the second, the third, etc., the fine-tuned feedback is given indirectly and through manipulation by other interactants, s/he may have had opportunity to benefit from it. That is why learning in DA groups emerged more than that of the NDA groups.

As for the differences from the pretest to posttest to delayed posttest, the effect of DA is once again highlighted. In other words, it was the effect of DA that led to the difference which was not observed in the high NDA group. Regarding the question as to why low NDA revealed a significant difference from the pretest to posttest to delayed posttest, it might have been the case that they were more motivated to learn even without receiving fine-tuned feedback.

Considering the fact that no proficiency effect was found on the posttest, the pragmatics-in-spite-of-grammar hypothesis might be the explanation. According to Rueda (2006), two claims have been made regarding the relationships between pragmatic competence and

grammatical proficiency. The first is that L2 learners cannot learn pragmatics without having already learned a good amount of grammar. This claim is termed as Grammar-then-Pragmatics. The second claim states that L2 learners can manage L2 pragmatics in one way or another even if they have no command of grammar. This claim is named as Pragmatics-in-spite-of-Grammar. According to the second hypothesis, grammatical competence is independent from pragmatic competence. The hypothesis also emphasizes L2 learners' pragmatic competence in their L1. Finally, the hypothesis underscores universal pragmatic competence in terms of which L2 learners may already be aware of some socially accepted norms in the L2 though they may have to redefine the newly encountered social settings and learn L2 social norms.

As far as this study is concerned, Pragmatics-in-spite-of-Grammar hypothesis seems to be the explanation especially when it comes to learners with low grammatical and/or linguistic competence because they provided appropriate answers to the DCT items. Such a state of affairs may have been due to their developing pragmatic competence which is independent from their linguistic competence. Moreover, they may have benefited from their L1 pragmatic competence as well. This may imply that EFL learners may not be necessarily in need of grammatical competence to develop first to be able to acquire L2 pragmatics as they may resort to their L1 pragmatic norms when the need arises. In this regard, Kasper and Rose (2002) also refer to the primacy of pragmatics by citing Schmidt (1983) as well as Koike (1989). Finally, Reuda (2006) refers to researchers such as Salsbury and Bardovi-Harlig (2001), Eisenstein and Bodman (1986, 1993), and Walters (1980) and offers evidence for the primacy of pragmatics over grammatical competence.

7. Conclusion

Three questions were raised in this study. The first was to see if there was any significant difference between DA vs. NDA on low and high proficiency EFL learners' acquisition of request and apology. The answer was yes because DA groups differed significantly from NDA groups in

all cases in posttest and in all but one case, i.e. high NDA, in the delayed posttest. These results are evidence of learning as well as recall in all groups though not evidence of recall in high NDA. The conclusion one comes up with is that DA has been effective when learning rather than recall of the speech acts of request and apology.

The second question was raised to see if there was any significant difference in the performance of DA groups and NDA groups from the pretest to posttest to delayed posttest. The answer was positive as significant difference was found on the performance of DA and NDA groups from the pretest to posttest to delayed posttest. However, on the basis of the results obtained, one group, i.e. high NDA, did not differ significantly from the pretest to posttest to delayed posttest. The explanation was in favor of DA again because the groups that received fine-tuned feedback and were instructed on the basis of DA changed significantly from pretest to posttest to delayed posttest while such was not the case with high NDA group. The third question was raised to see if instruction and level of proficiency had any joint effect on students' posttest and delayed posttest performance. The answer was both yes and no. Yes because instruction was found to be effective on both post and delayed posttest. No, because level of proficiency was not found to be effective on posttest though it was found to be so on delayed posttest results. Therefore, considering the overall results, instruction, i.e. DA is found to be the only effective variable on both post test as well as delayed posttest test results. As to a comparison between the findings of this study and other similar studies, it should be pointed out that findings of this study do confirm other studies as this study has come up with similar finding, indicating that DA is an effective approach that needs to be taken into consideration by language teachers and other applied linguistics professionals.

The pedagogical implication drawn from this study is twofold. The first implication is that ZPD-oriented, DA-based, interactive activities lead to better learning of L2 pragmatics on the part of the EFL learners. Therefore, it is suggested that more ZPD-oriented activities be

incorporated into the EFL lessons since every one of the learners may find their own particular chance of meaningful interactions in one way or another thereby fostering their ILP development. Secondly, it is suggested that speech acts be included within the materials EFL learners are exposed to from the very low levels to higher levels of proficiency because as low level learners' ILP development in this study indicated true beginners may also learn L2 pragmatics without necessarily having already developed an acceptable level of general proficiency. Hence, it is recommended that textbook writers as well as instructors pay to the speech acts attention they really deserve.

References

Ableeva, R. (2010). *Dynamic assessment of listening comprehension in second language learning.* Unpublished Ph. D dissertation, The Pennsylvania State University.

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, *78*, 463-483.

Bardovi-Harlig, K., & Dörnyei, Z. (1998) Do language learners recognize pragmatic violations? Pragmatic vs. grammatical awareness in instructed L2 learning. *TESOL Quarterly, 32*, 233-59.

Bardovi-Harlig, K., & Hartford, B. (1996) Input in an institutional setting. *Studies in Second Language Acquisition, 17*, 171-188.

Bransford, J.D., Delclos, V.R., Vye, N.J., Burns, M.S., & Hasselbring, T.S. (1987). State of the art and future directions. In C.S. Lidz (Ed.). *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 479-496). New York: Guilford Press.

Budoff, M. (1974). *Learning potential and educability among the educable mentally retarded* (Final Report Project No.312312). Cambridge, MA: Research Institute for Educational Problems, Cambridge Mental Health Association.

Burns, M.S., Haywood, H.C., Delclos, V.R., & Siewart, L. (1987). Young children's problem-solving strategies: An observational study. *Journal of Applied Developmental Psychology, 8*, 113-121.

Caffrey, E., Fuchs, D., & Fuchs, L.S. (2008). The predictive validity of dynamic assessment: A Review. *The Journal of Special Education*, *41*(4), 254-270.

Carlson, J.S., & Wiedl, K.H. (1978). Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. *American Journal of Mental Deficiency, 82,* 559-564.

Eisenstein, M., & Bodman, J. (1986). "I very appreciate": Expressions of gratitude by native and non-native speakers of American English. *Applied Linguistics, 7*, 167-185.

Eisenstein, M., & Bodman, J. (1993). Expressing gratitude in American English. In G. Kasper & S. Blum-Kulka (Eds.). *Interlanguage pragmatics* (pp. 64-81). Oxford: Oxford University Press.

Feuerstein, R. (1979). *The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques.* Baltimore: University Park Press.

Field. A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: SAGE.

Geranpayeh, A. (2003). A quick review of the English Quick Placement test. *Research Notes*, *12*, 8-10.

Gettinger, M. (1984). Measuring time needed for learning to predict learning outcomes. *Exceptional Children, 51,* 244-248.

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

House, J. (1996). Developing pragmatic fluency in English as a foreign language: Routines and metapragmatic awareness. *Studies in Second Language acquisition*, *18*, 225-252.

Jitendra, A.K., & Kameenui, E.J. (1993). Dynamic assessment as a compensatory assessment procedure: A description and analysis. *Remedial and Special Education, 14*, 6-18.

Kasper, G., & Rose, K.R. (2002). Pragmatic development in a second language. *Language Learning,* 52, Supplement 1.

Koike, D.A. (1989). Pragmatic competence and adult L2 acquisition: Speech acts in interlanguage. *The Modern Language Journal*, *73*, 79-89.

Kozulin, A., & Garb, E. (2001). *Dynamic assessment of EFL text comprehension of at-risk students,* Paper presented at 9th Conference of the European Association for Research on Learning and Instruction, Fribourg, Switzerland.

Lantolf, J.P., & Poehner, M.E. (2004). *Dynamic assessment in the language classroom.* CALPER Professional Development Document, 0411, Center for Advanced Language Proficiency Education and Research.

Lantolf, J.P., & Poehner, M.E. (2011). Dynamic assessment in the classroom: Vygotskian praxis for second language development, *Language Teaching Research*, *15*, 11-33.

Lidz, C.S., & Gindis, B. (2003). Dynamic assessment of the evolving cognitive functions in children. In A. Kozulin, B. Gindis, V.S. Ageyev, & S.M. Miller (Eds.). *Vygotsky's educational theory in cultural context.* Cambridge: Cambridge University Press.

Nassaji H., & Swain, M. (2000). A Vygotskyan perspective on corrective feedback in L2: The effect of random versus negotiated help on the learning of English articles. *Language Awareness*, *9*, 34-51.

Ohta, A.S. (2005). Interlanguage pragmatics in the zone of proximal development. *System, 33*, 503-517.

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

Poehner M.E. (2008). *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development.* Berlin: Springer.

Rose, K.R., & Ng Kwai-fun, C. (2001). Inductive and deductive teaching of compliments and compliment responses. In K.R. Rose & G. Kasper (Eds.). *Pragmatics in language teaching,* (pp.145-170). Cambridge: Cambridge University Press.

Rueda, Y.T. (2006). Developing pragmatic competence in a foreign language. *Colombian Applied Linguistics Journal*, *8*, 169-182.

Salsbury, T., & Bardovi-Harlig, K. (2001). "I know you mean, but I don't think so:" Disagreements in L2 English. In L.F. Bouton (Ed.). *Pragmatics and language Learning: v*ol. 10 (pp. 131-151). University of Illinois, Urbana-Champaign: Division of English as an International Language.

Schmidt, R. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of one adult. In N. Wolfson & E. Judd (Eds.). *Sociolinguistics and language acquisition* (pp. 137-174). Rowley, MA: Newbury House.

Taguchi, N. (2006). Analysis of appropriateness in a speech act of request in L2 English. *Pragmatics, 16*, 513-533.

Tajeddin Z., Alemi, M., and Pakzadian, S. (2011). *Group dynamic assessment: Application in interlanguage pragmatics*. Unpublished manuscript, Sharif University of Technology, Iran.

Tajeddin Z., & Mosleh, M. (2011). *A dynamic assessment-based vs. an output-based approach to the instruction of the speech acts of apology, complaint, and request*. Unpublished manuscript, Allameh Tabataba'i University, Iran.

Takahashi, S. (2001). The role of input enhancement in developing pragmatic competence. In K.R. Rose and G. Kasper (Eds.). *Pragmatics in language teaching*. Cambridge: Cambridge University Press.

Takahashi, S. (2005). Noticing in task performance and learning outcomes: A qualitative analysis of instructional effects in interlanguage pragmatics. *System*, *33*, 437-461.

Takimoto, M. (2006). The effects of explicit feedback on the development of pragmatic proficiency. *Language Teaching Research, 10*, 393-417.

Tateyama, Y. (2001). Explicit and implicit teaching of pragmatic routines: Japanese *sumimasen*. In K.R. Rose & G. Kasper (Eds.). *Pragmatics in language teaching*. Cambridge: Cambridge University Press.

Vye, N.J., Burns, S., Delclos, V.R., & Bransford, J.D. (1987). A comprehensive approach to assessing intellectually handicapped children. In C.S. Lidz (Ed.). *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 327-359). New York: Guilford Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walters, J. (1980). Grammar, meaning, and sociological appropriateness in second language acquisition. *Canadian Journal of Psychology, 34*, 337-345.

Williams, M., & Burden, L.R. (1997). *Psychology for language teachers: A social constructivist approach.* Cambridge: Cambridge University press.