International Conference on Current Trends in ELT

Dynamic Assessment of L2 Listening Comprehension in Transcendence Tasks

Karim Shabani*

Allameh Mohaddes Nouri University, Nour, 46517, Mazandaran, Iran.

Abstract

L2 assessment researcher have long been preoccupied with the question of whether the learners can reapply their newly gained abilities in the non-assessment contexts, a concept traditionally known as generalizability (Bachman, 2004; Cronbach, 1990; Messick, 1995; Poehner, 2008) The present study was designed to flesh out the concept of generalizability from a qualitatively different perspective, namely, Vygotsky's (1978) Socio-cultural Theory (SCT) of Mind under a new framework known as *transcendence* (TR) (Poehner, 2010). This study, inspired by Poehner's (2009) interactionist group dynamic assessment G-DA, and Feuerstein et al.'s (1976) Mediated Learning Experience (MLE) concept, has set out to track the developmental trajectories of L2 learners' listening comprehension ability within a microgenetic framework in hopes to bring into perspective learners' qualitative changes during the interaction and mediation collaboratively negotiated in their ZPD across a set of different innovative tasks. The data for this G-DA study were collected from the classroom context (Poehner, 2009). The results, including both the quantitative and qualitative, indicated that non-dynamic assessment (NDA) procedure stops short of fully capturing the learners' underlying potential and leaves aside the abilities which are in the state of ripening. It was shown that the learners' ability to recognize an unrecognized word of the pretest transcended beyond the posttest task to the TR session, an improvement signaling their progressive trajectories towards higher levels of ZPD. On implication side, the paper recommends the use of DA as a development-oriented procedure to assess the learners' abilities, a procedure which focuses on the learners' emerging abilities in constantly innovative tasks.

Keywords: Group dynamic assessment (G-DA); Zone of Proximal Development (ZPD); transcendence; listening; interactionist

1. Introduction

Traditionally, assessment has been dichotomized into two forms: summative and formative. The former attempts to capture the learners' products of past learning and occurs at the end of an instructional program to ensure the success of an instructional course and learners' achievement. The most well-known widely-used type of summative

* Corresponding author. Tel.: +989113954613; fax: +981226255245.
  E-mail address: shabanikarim@gmail.com

assessment is achievement test which is intended to determine the extent to which learners have mastered specific course content. Quite contrariwise, the formative assessment is administered throughout the life of an educational program while the course is still in progress. According to Fulcher and Davidson (2007), formative assessment is carried out during the learning process as an intervention that is designed to encourage further learning and change. Indeed, practitioners' departure from summative towards formative assessment resides in the inherent potential of the latter which provides feedback into teaching and supports more reliable classroom-based decisions. Examples of formative assessment include observation, quizzes, questioning, self/peer assessment, graphic organizers, etc. (Bachman, 2004).

Poehne (2007) contends that these two forms of assessment, summative and formative, rest on the ideology of Cartesian dualism which assumes a separation between instruction and assessment. These forms take the testees' solo performance as an indication of their full potential. Such an ideology is dismissed altogether in the context of dynamic assessment (DA) where the main ambition is to materialize the twin function of assessing *and* promoting learners' developmental processes. In the DA camp, it is assumed that observation of individuals' independent performance limits the scope of assessment and falls short of revealing the full range of person's abilities. DA proponents further postulate that mediated interaction and, then, tracing individuals' responsiveness to mediation provides a site for surfacing the full range of their abilities (Lantolf, 2004; Poehner, 2008). Even the concept of generalizability is redefined in DA context as 'transcendence' i.e. the capacity on the part of learners to recontextualize their learning in new contexts with the mediation being continually present in the course of assessment. he main problem the present study is going to address is the question of whether the learners can reapply the abilities they gain through DA in the innovative contexts. The study is also an attempt to test the applicability of DA in uncovering the learners' emerging abilities in both familiar and unfamiliar tasks. It is assumed that the DA can fill the gap in the literature on the L2 testing as a revolutionary assessment procedure which looks upon evaluation and instruction as a unified procedure.

2. Generalizability reconceptualized as transcendence in DA context

Generalizability refers to the extent to which one can make inferences about individuals' future performance in non-assessment contexts based on his performance in assessment context (Bachman, 2004). Generalizability as a traditional psychometric property rests on the assumption that contexts are homogeneous and contextual variables being the same across different instructional and assessment situations. In other words, it assumes a match between instructional tasks and assessment tasks. At the time of assessment, to increase the reliability of procedures, NDA practitioners try to control the sources of variances and the individuals perform in isolation; they forbid the use of all subsidiary materials like computers, calculators, reference materials and other aids, so that standardization is secured and the risk that interpretation of the results is contaminated by factors external to the examinee is weakened. Quite paradoxically, such contrived procedures during assessment are never found during the instruction and learners benefit from all kinds of resources including cues or support from the teacher, collaboration with peers, subsidiary textbooks or internet, etc. Therefore, any claim about generalizability of learners' performance from their assessment context to non-assessment context seems problematic (Poehner, 2008).

According to McNamara and Roever (2006) the oddity in NDA arises when it sets to observe learners under assessment conditions but, then, draw inferences about their underlying abilities and make statements about their probable performance in non-assessment situations. McNamara (2004) referred to this problematic scenario as Labov's Observer's Paradox because learner's performance under assessment conditions can never be a true indicator of his performance in non-assessment situations mainly due to its change as a result of being observed.

According to DA researchers, any attempt to generalize the examinees' scores taken from their independent, unassisted performance in assessment tasks to a broader domain without identifying adequately the breadth and depth of the underlying constructs underestimates learners' potential and introduces a major threat to the construct validity known as 'construct under-representation' (McNamara & Roever, 2006).

What seems more relevant in DA context is Feuerstein's concept of transcendence. Unlike the concept of generalizability, transcendence rests on the assumption that contexts are highly variable. Transcendence (TR) refers to the learner's ability to transcend their learning to new contexts and the extent which they can recontextualize their learning (Poehner, 2008). TR is not considered as training oriented toward a specific content but rather as "a series of procedures designed to help learners continuously move beyond their abilities and the here-and-now demands of a given problem" (Poehner, 2007, p.327). Therefore, the issue in DA, as Poehner states, "is not to generalize to hypothetical contexts but to track learner development from one DA interaction to the next" (Poehner, 2008, p. 75). According to Poehner, (2007:334), "transcendence is fundamental to tracking development because it involves going beyond the test as learners endeavor to recontextualize their abilities while engaging in new tasks" because the difference between learners become manifest as they ask for different kinds of support during transcendence tasks that become increasingly difficult.

The hallmark of the concept of *transcendence* is the implied assumption that mediation is never withdrawn in the new contexts where the learners' capacity to tackle over new tasks is being assessed. Thus, the DA sessions are never closed as soon as the intervention stage is finished; learners can be further discriminated according to "how well they are able to sustain their performance when variations are introduced and tasks become more difficult" (Poehner, 2007, p. 327).

Proceeding from a Vygotskian perspective, Poehner (2007) contends that generalizing an interpretation of learners' abilities based on one-off performance makes little sense and that the goal of classroom interactions is connecting each teaching-assessment episode to the next in order to continually negotiate within learners' ZPD, promote the their development and track their progress over time.

DA researchers insist that teachers should further examine their learners' capacity to transcend their abilities beyond a given DA task. Williams and Burden (1997) argue that the learners need to understand the transcendence of the learning experience, identifying rules and recurring themes, leading to learning beyond the immediate time and place. The teachers should notify the learners that they learn a skill not merely for the completion of the immediate classroom task but more importantly for their 'real world' applications in the future. Teachers are responsible for explaining the purpose of the exercise and to make learners realize its inherent value to them.

3. Group dynamic assessment (G-DA)

Advocates of socio-cultural theory in applied linguistics suggest that collaborative interaction assists co-construction of knowledge in ways that guarantee remarkable achievement in different aspects of language learning culminating in knowledge far beyond the individuals' current status.

Krashen' input hypothesis has been tagged as inadequate to explain the nature of second language acquisition mainly due to its lip service to the primary role of social mediation in language development (Donato, 1994). Standing in striking contrast to Krashen, Vygotsky (1978) viewed development (including cognitive and linguistic) as a function of interpsychological processing experienced by individuals during concrete social interactions and then at a later stage a matter of intrapsychological and intrapersonal processing. Put it more plainly, Vygotsky conceptualized learning first as social and then individual.

Following Vygotsky's socio-cultural theory of mind and specially his concept of ZPD, group dynamic assessment or G-DA placing respectable premium on the collaborative negotiation of language tokens rests on the assumption that it is possible to engage a group of learners in collaboratively co-constructing a group's ZPD while catering to each individual's ZPD. Moreover, G-DA assumes that through joint efforts the group might function in ways that are beyond the present capabilities of any individual member (Gibbons, 2003; Poehner, 2009).

Poehner (2009), a pioneering researcher of G-DA, argues that group-based and one-to-one dynamic assessment

follow the same principle of offering learners mediation to help them co-construct a ZPD, but they differ in that G-DA must also take account of group's ZPD. As an illustration, Lantolf and Poehner (in press cited in Poehner, 2009) report the results of a G-DA study in which the teacher offers mediating support to a group of learners. The teacher does not run through the full range of mediating prompts with a single learner before beginning again with another individual. Instead, the teacher's focus remains fixed on the entire class but by calling on a particular student to answer a question and then addressing another to continue the preceding contribution, the teacher moves the entire class forward in its ZPD through co-constructing ZPDs with individuals.

Poehner (2009) identifies two types of interactants in the classroom; namely, the primary interactants including the teacher and one of the students with whom he negotiates his linguistic support/mediation and the secondary interactants including other students who listen and benefit from the teacher-student exchanges. He argues that because the exchange occurs in the social space of the class and before the other group members it has mediating potential for the rest of the group as well. Moreover, he identifies two forms of G-DA; concurrent and cumulative. In the *concurrent* G-DA, the teacher dialogues with the entire group. He may provide mediation in response to an individual but the interaction shifts rapidly between primary and secondary interactants as one learner's question, struggle, or comment sets the stage for another's contribution. In the *cumulative* G-DA, the teacher "conducts a series of one-to-one DA interactions as the group works toward the mastery of a problem; that is, the individuals take turns engaging directly as primary interactants with the teacher, with the understanding that each subsequent one-to-one exchange will have the advantage of building on earlier interactions that the class witnessed. This approach is cumulative in that the goal is to move the entire group forward in its ZPD through negotiations with individual learners in their respective ZPDs. Cumulative G-DA attempts to move the group forward through co-constructing ZPDs with individuals, but concurrent G-DA supports the development of each individual by working within the group's ZPD" (Poehner, 2009, p.476). The main purpose of this study is to examine the role of DA in uncovering the learners' emerging abilities and their use in the innovative contexts. To address the problem, the following research questions were raised:

1. To what extent can DA uncover the learners' listening comprehension abilities?

How best can DA assess the learners' listening comprehension abilities in the transcendence tasks?

4. Methodology

4.1. Participants

This research involved a group of 17 undergraduate students of English major with an intermediate level of linguistic proficiency recruited from Allameh Mohaddes Nouri College. The learners' status being 'intermediate' refers to the number of semesters they have spent studying English at the college. Thus, the participants are named intermediate by virtue of their enrollment in a fifth semester undergraduate university language course. They were both males and females ranging in age from 20 to 25. An announcement was made earlier on holding a reinforcement course of listening and 23 students volunteered to take part from whom 17 were selected based on their availability and time schedule.

4.2. Instrumentation

The following considerations were taken into account when selecting the listening texts for this study:

1. The texts should be authentic and genre specific. To meet this condition, all the texts were selected from the VOA news broadcasts.

2. All the texts used in the pretest were similar in terms of topic, pace of delivery and level of difficulty but not identical.

3. The texts should be relevant to the requirements of the course. The texts were selected from authentic news broadcasts. The main reason behind the selection of media discourse as the main text type for classroom practice was the observation that for many learners of English the ability to understand news broadcasts is the 'holy grail' of second language listening performance (Cross, 2009). The author's focus on this type of genre had the advantage of recruiting subjects consisting of highly motivated students, a theoretical and methodological prerequisite for implementation of G-DA.

4.3. Design

This study aligned itself with an interactionist G-DA approach and a mixed methods research methodology comprised of a propositional (idea unit) analysis and a microgenetic method (an SCT-based method). The combination of these two methods allowed a quantitative and qualitative examination of the data. The repeated analysis of the data through quantitative and qualitative tools helped scrutinize the data from different perspectives and establish methodological *triangulation*, a property that was felt essential for securing the validity (trustworthiness) and reliability (accuracy) of the findings (Mackey & Gass, 2005). The study followed a 'quan→QUAL' approach with an initial quantitative phase followed by a dominating qualitative phase (Dornyei, 2007). Moreover, it is of an 'intrinsic' case study type (Stake, 1995) with the aim to closely examine a particular group's performances in the real-life context for an in-depth analysis of its nuances.

The schematic representation of the design of the present study is as follows:

Week 1 ⟶ pretest (NDA + DA)
Week 2-7 ⟶ enrichment program (EP)
Week 8-9 ⟶ posttest (NDA + DA)
Week 10 ⟶ transfer 1 (transcendence; TR1)
Week 11 ⟶ transfer 2 (transcendence; TR2)
Week 12 ⟶ transfer 3 (transcendence; TR3)

As can be observed above, following the pretest, EP and posttest, the TR sessions were held at the end of the study. These tasks were designed to understand how far the students could extend their L2 listening ability attained as a result of G-DA intervention to new and more demanding tasks. For this reason, the three TR sessions administered were based on more complex texts. The three TR texts i.e. TR1, TR2 and TR3 were different from the pre/posttest texts in the format (audio) and topic. Unlike the pre/posttest texts that were audiovisual and treated almost the same topic (revolution and riots), the TR texts deprived students of the benefit of visual information and were thematically different but typical of the news genre (TR1 was an interview including three participants and TR2 and TR3 were speeches delivered by high-ranking officials). It is noteworthy that the enrichment program which lasted about 5 weeks aimed at addressing and fastening the most recurrent listening comprehension problems observed in the pretest.

4.4. Procedure

Each assessment session was characterized by two phases, an NDA elicitation stage which aimed at surfacing the class's ZAD or current status and a G-DA intervention phase. The decision to administer the NDA and DA sessions simultaneously reflects Poehner's (2005) suggestion that in DA we do not need to hold an NDA test separately since the DA procedure has the dual function of detecting students' ZAD and ZPD. At the outset, the teacher initiated the assessment non-dynamically asking the students to provide the content of the heard segments or the meaning of selected words and phrases independently without assistance. The students were also checked at times on the meaning of specific words and phrases. The NDA elicitation stage served as a diagnostic procedure to gain insights into the students' independent performance abilities. Immediately following this stage, a mediation phase began aiming at promoting learners' listening comprehension processes. As in the NDA, the mediated (G-DA) phase performed the same diagnostic function but delved more into the learners' cognitive processes and challenges involved aiming at enhancing their abilities.

After the data collection, all recorded data were studied, transcribed (as needed) and analyzed. The related data were transcribed using the transcription conventions adapted from Johnson (1995).
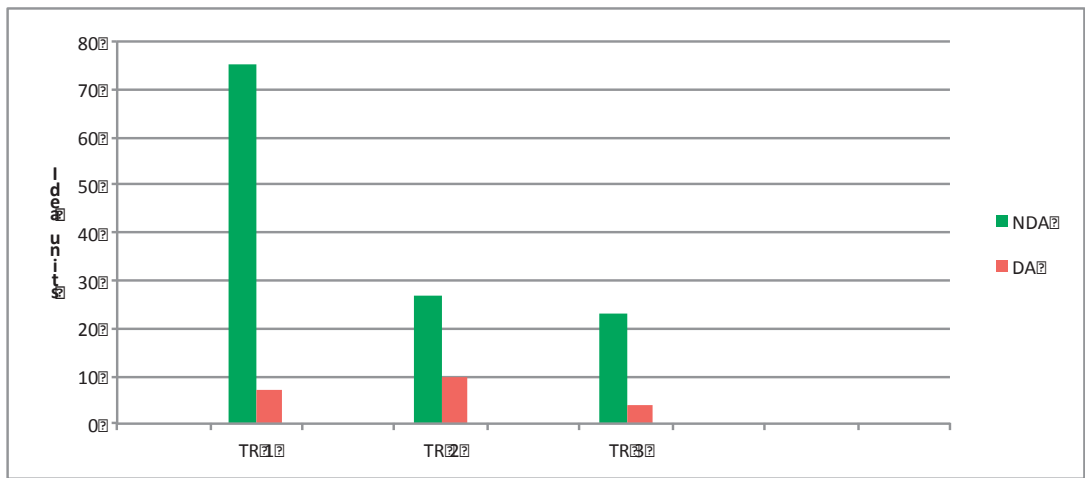
5. Results

5.1. Quantitative data analysis

The following figure displays the results of students' IU recalls in the three TR sessions:

Table 1. Results of IU recalls in the TR sessions

TR 1		TR2		TR3	
86		36		26	
NDA	DA	NDA	DA	NDA	DA
75	7	27	10	23	4

Total IUs: 145
Total IUs (NDA): 121
Total IUs (DA): 19

Note: NDA = non-dynamic assessment; DA = dynamic assessment

Fig 1. Chart of idea units in the TR sessions

Students' performances on the three TR tasks show almost the same extent of increase in the number of IU recalls as in the pretest, a finding which confirms the impact of G-DA interactions on the students' performances. This observation also demonstrates an inherent property of DA that assessment and instruction for development occur concurrently as a unified enterprise and that when instruction is incorporated into the assessment procedure a more complete picture of the learners' underlying abilities is obtained.

An interesting observation was that although the TR texts differed in format (audio) and topic from the pretest texts, they did not cause unexpected breakdowns. On encountering new listening tasks, students displayed signs of improvement in the form of a move away from relying upon explicit prompts towards more implicit ones. Learners' developmental trajectories and improvement were traced in the form of reduced demands for explicit mediations and high frequency of implicit mediational moves (for more detail see 'qualitative analysis' below). However, as in the pretest and posttest sessions, students faced a certain number of insurmountable IUs (1 in TR2 and 7 in TR3) which precluded any kind of recall, unmediated and mediated. Observation of learners' enhanced performance in the TR

tasks confirms their ability to extend their learning to innovative contexts.

It's time now we probed much deeper into the learners' performances through a qualitative analysis of the G-DA interactions.

5.2. Qualitative data analysis

The following protocols portray the learners' listening performance across different assessment sessions (i.e. pretest, posttest, and TR sessions). The qualitative analysis takes into account mediated and, when necessary, unmediated (independent) performances to illustrate learners' microgenetic gains. Examples of TR sessions are given to illustrate learners' developmental path and emerging capacity to reapply their enhanced performance in innovative contexts.

The following excerpt taken from the TR1 session demonstrates learners' enhanced ability to recognize a familiar word in new context.

Protocol 1 [TR1 session]
[He says there is a period of chaos]
1. T: well, let's go over this sentence. What did you hear?
2. S1: he says
3. Ss: [silent]
Listening
4. S2: he says there is a period of chaos
5. T: chaos means?
6. S2: *ashub* ^ (chaos)

Learners' improvement in this episode was traced in the form of reduced demands for explicit types of mediation. A known lexical item which lent itself to a strenuous and effortful recognition during the pretest was easily identified after a very implicit mediation (i.e. *replaying*), an observation which shows learners' departure from dependent performance towards independent performance and task control. In this example, the ability to recognize a lexical item with implicit types of feedback/mediation was observed in a new listening context.

The example above provides evidence of learners' improvement of lexical knowledge as a result of G-DA instruction in the pretest. The next excerpt taken from the TR session provides evidence of students' improved ability to recognize a new word in a totally new context:

Protocol 2 [TR1 session]
[We are also very concerned about the looting, pillaging and destruction of civilian property]
1. T: ok, now let's see what the speaker said
2. S1: we are very concerned about looting, and destruction
3. T: good, what else?
4. S2: looting and destruction of civilian property
5. T: what is the meaning of 'civilian'?
6. S3: *shahrvandi* ^ (of citizens)
7. T: you're right. What was the word after 'looting'?
8. S5: [silent]
9. T: it is 'pillaging'. Pillaging p-i-l-l-a-g-e means 'to rob'

As can be understood from this excerpt, the lexical item (i.e. '*looting*') that learners were not able to pick up in the pretest was immediately and correctly recalled by one of the students in the TR session (line 2). The ability to comprehend a lexical item embedded in a new context could be viewed as the climax of learners' developmental growth (Feurestein et al., 1976; Poehner, 2005) and evidence of *internalization* or *appropriation* (Vygotsky, 1978). This excerpt demonstrates that the knowledge gained through G-DA can be deployed and used in innovative listening contexts. However, the observation of learners' inability to decode a new lexical item in

this episode further demonstrates that learners' lexical knowledge is still incomplete and students might encounter new words in new contexts.

6. Discussion and conclusion

This study was an attempt to provide an analysis of L2 learners' listening performance quantitatively and qualitatively. The results afforded insights into a two-month L2 listening microgenesis. The quantitative method (propositional analysis) provided a *product* view of the learners' listening performances while qualitative (microgenetic) method helped examine the *process* of learners' listening comprehension abilities. The microgenetic analysis was used to both verify and complement the quantitative findings. The method, in addition to the identification of unmediated and mediated abilities, allowed for diagnosing the sources of learners' listening comprehension difficulties and their microgenetic changes and the quality (type) of the teacher's mediation. This mixed methods approach comprising quantitative and qualitative procedures yielded a more elaborate and comprehensive understanding of the learners' listening abilities, thereby enhancing the validity of the conclusions.

The observation of increased IU recalls during mediated sessions provided support for the significant effect of G-DA interactions on the learners' listening comprehension processes. The G-DA did much more than to assess the abilities that were fully matured; it actively intervened to develop those abilities by unifying assessment and instruction as one single activity. Besides getting access to the students' ZAD, G-DA helped identify the learners' ZPD. The results, including both the quantitative and qualitative, indicated that NDA procedure stops short of fully capturing the learners' underlying potential and leaves aside the abilities which are in the state of ripening.

The quantitative analysis of IU production in the TR sessions also revealed a progression of group ZPD. The listening improvement in the TR sessions was observed almost with the same size of IU recalls as in the pretest. It bears repeating that in the TR sessions learners were tested on more difficult tasks in innovative contexts. The TR assessment sessions involved the learners in listening to a number of texts that were different in topic and modality. During unmediated phases the number of learners' IU recalls showed a noticeable increase and in the mediated phases a tendency towards reliance on implicit types of feedback was noticed; the frequency of the first two implicit strategies on the scale rose dramatically while the explicit strategies showed an abrupt drop.

G-DA protocols were provided to show the effects of G-DA on the learners' listening abilities at the levels of phonology, lexis and grammar. The examples provided support for enhancement of learners' independent performance both in familiar (posttest) and innovative (TR) tasks. It was shown that the learners' ability to recognize an unrecognized word of the pretest transcended beyond the posttest task to the TR session, an improvement signaling their progressive trajectories towards higher levels of ZPD in terms of phonological functioning. The lexical improvement manifested itself in the form of reduced demand for explicit mediations and, in some cases, increased independent recalls. Finally, on the area of grammar, learners' improvements were demonstrated through examples of protocols from the TR session which confirmed their improved ability to draw upon the necessary grammatical knowledge to facilitate the process of listening comprehension. On implication side, the paper recommends the use of DA as a development-oriented procedure to assess the learners' abilities, a procedure that focuses on the learners' emerging abilities in constantly innovative tasks. Unlike, the static assessment, the DA envisages no endpoint for the learners' development because it considers the learners' abilities as dynamic and changeable.

A major focus of this study was to gain insights into the group ZPD of a group of L2 learners and more technically its intragroup ZPD. A new line of research worth exploring would be the simultaneous use of G-DA by the mediator in another class with a different group of L2 learners. The findings would help compare the intragroup ZPD and the intergroup ZPD (Cervania, 2003) of the two groups. The inclusion of several randomly selected groups in the design would help identify the objective ZPD of L2 learners at a certain age period.

References

Bachman, L.F. (2004).*Statistical Analyses for Language Assessment*. Cambridge: Cambridge University Press.
Cervania, R. (2003). A Qualitative Study of Group Work in the Development of Filipino as a Second Language. Unpublished Doctoral Dissertation. University of Hawaii.
Cronbach, L. J. (1985). Essentials of psycological testing. In Ed./Druk: 5. Paperback. Pp: 726. Cond./Kwaliteit: Goed.
Donato, R. (1994). Collective scaffolding in second language learning. In J. Lantolf & G. Appel (Eds.), *Vygotskian approaches to second language learning research,* Norwood, NJ: Ablex (pp. 33-56).
Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford: Oxford University Press.
Feuerstein, R. (1976). Dynamic assessment of cognitive modifiability in retarded performers: The learning potential assessment device. In B. B. Wolman (Ed.), *International encyclopedia of neurology, psychiatry, psychoanalysis and psychology*, Section XII. New York.
Fulcher, G. & Davidson, F. (2007). *Language Testing and Assessment*. London & New York: Routledge.
Gibbons, P. (2003). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly*, *37*(2), 247-273.
Johnson, K.E. (1995). *Understanding Communication in Second Language Classrooms*. Cambridge: Cambridge University Press.
Lantolf, J. P. (2004). Overview of sociocultural theory. In O. S. John, K. van Esch & E. Schalkwijk (Eds.), *New insights in second language learning and teaching* (pp. 13-34). Frankfurt: Peter Lang Verlag.
Mackey, A., & Gass, S. M. (2005). *Second Language Research: Methodology and Design*. Mahwah, NJ: Lawrence Erlbaum Associates.
McNamara, T. F, & Roever, K. (2006). *Language Testing: The Social Dimension*. Oxford: Blackwell.
Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50,* 741–749.
Poehner, M.E. (2005). Dynamic assessment of oral proficiency among advanced L2 learners of French. Unpublished doctoral dissertation. The Pennsylvania State University, University Park, PA.
Poehner, M.E. (2007). Beyond the Test: L2 Dynamic Assessment and the Transcendence of Mediated Learning. *Modern Language Journal, 91*, 323-340.
Poehner, M.E. (2008). Both sides of the conversation: The interplay between mediation and learner reciprocity in dynamic assessment. In J.P. Lantolf & M.E. Poehner. (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 33-56). London, Equinox Publishing.
Poehner, M. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly*, *43*(3), 471-491.
Stake, R. E. (1995).*The Art of Case Study Research*. Thousand Oaks, CA: Sage Publications.
Vygotsky, L. (1978). Mind in society: The development higher psychological processes.(M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, (Eds. and Trans.) Cambridge, MA: Harvard University Press. (Original work published in 1955).
Williams, M. & Burden, R.L. (1997). *Psychology for Language Teachers: A Social Constructivist Approach.* Cambridge: Cambridge University Press.