

Second Language Task Difficulty: Reflections on the Current Psycholinguistic Models

Zhengdong Gan

The Hong Kong Institute of Education, Hong Kong

Email: zdgan@ied.edu.hk

Abstract—One of the unresolved issues currently being debated not only in the area of second language (L2) pedagogy but also in language testing and assessment community is what makes an L2 task difficult or complex. This paper serves three purposes. It first presents a review of historical perspectives on task difficulty, and discusses the recent developments in the testing and assessment field's understanding of a psycholinguistic approach to characterizing task difficulty proposed by some second language pedagogy researchers. The paper then makes the case that this psycholinguistic approach to the investigation of task difficulty focuses mainly on the cognitive processing of a given task, obscuring the role of linguistic demand and social dimensions in defining task difficulty and developing tasks for the L2 classroom. The paper points out that what is lacking in the literature is how interpersonal dynamics and affective dimensions of a given task may influence the cognitive processing of the task. The paper ends by suggesting that future research needs to examine how different psycholinguistic factors, linguistic demands, learner attributes and social conditions interact to make a given task more or less 'difficult' for different learners through use of different research methods. Such research will be useful as it is in line with the central objective of language teaching to produce learners who are able to communicate effectively in the target language of a particular speech community.

Index Terms — task, task difficulty, task-based language teaching, psycholinguistic models

I. INTRODUCTION

As a consequence of policy decisions to increase the number of people in their population who can communicate effectively in English, a great deal of investment has been made by governments worldwide in programs designed to improve the teaching and learning of English. For example, Nunan (2003) observes that national policies and syllabuses on English language education in Asian countries have been moving increasingly towards task-based language teaching. In the literature of L2 acquisition and pedagogy, task-based language teaching and learning is often advocated as the latest methodological realization of communicative pedagogy (Hu, 2005; Littlewood, 2007). There is now widespread acceptance that it is by engaging learners in doing valued tasks that relevant declarative and procedural knowledge is developed, and learners further develop deep linkages between what they learn and how that learning can be put to use beyond the classroom (Norris, 2009). However, despite the fact that task-based language teaching has now achieved something of the status of a new orthodoxy (Littlewood, 2004), researchers and practitioners still seem uncertain as to what exactly a task is. Breen (1987), for example, sees tasks as a range of learning activities from the simple and brief exercise type to more complex and lengthy activities such as group problem-solving or simulations. Nunan (1993) associates tasks as activities that involve speakers in using language for the purpose of achieving a particular goal or objective in a particular speaking situation, that is, goal-oriented language use. In spite of an apparent lack of consensus about what constitutes a task, L2 acquisition and pedagogy researchers generally agree that pedagogic tasks provide a means of engaging learners in meaningful use of the target language (Brumfit, 1984). The concept of 'learning through tasks' has thus become an intrinsic part of the professional discourse and various innovations with task-based language teaching are frequently introduced (Littlewood, 2007).

This paper focuses on L2 task difficulty which is now firmly established as a central concern in task-based language teaching and remains a contested domain of inquiry and practice. In the discussion that follows, I first present a review of the historical perspectives on task difficulty, and discuss the recent developments in the testing and assessment field's understanding of a psycholinguistic approach to characterizing task difficulty proposed by some second language pedagogy researchers. I then make the case that this psycholinguistic approach to the investigation of task difficulty may obscure the role of linguistic and social dimensions of a given task in learners' processing of the task.

II. HISTORICAL PERSPECTIVES ON TASK DIFFICULTY

An important issue which underscores much of the enquiry in the fields of L2 acquisition and pedagogy, and more recently L2 testing and assessment is what makes an L2 task difficult. Investigations into L2 task difficulty and its impact upon interlanguage developmental date to the early 1980s when SLA researchers tended to classify tasks according to the quality and quantity of interaction that result from a particular task. Hence it is assumed that two-way

tasks are more difficult than one-way tasks (Pica et al., 1993; Fulcher, 2003). Crookes (1986, Cit. in Norris, Brown, Hudson, and Yoshioka, 1998) lists the following factors that can affect task difficulty: 1) Number of steps needed; 2) Number of parties involved; 3) Presupposed knowledge; 4) Intellectual challenge; 5) Spatio-temporal displacement.

Candlin (1987) proposed a general scheme of task sequencing features. This scheme involves description of some factors that are supposed to affect the level of difficulty of a given task. For example:

1) Cognitive load. Tasks which require learners to follow a clear chronological sequence, referring to individual actions of individual characters, will clearly be cognitively less demanding than a task in which there is no such clear development and where the picture is complicated by multiple actions and multiple actors.

2) Communicative stress. Tasks where other interlocutors are less communicatively competent, or tasks where other interlocutors know less about the subject-matter will be less challenging than tasks where other interlocutors are more communicatively competent, or tasks where other interlocutors know more about the subject-matter

3) Particularity and generalizability. Tasks which follow some generalized pattern, or some ritualized interpretive schema, will be easier to manage than those where the order of assembly or the norms of interpretation are unclear and to be negotiated.

4) Code complexity and interpretive density. Tasks that involve description of textually elaborate texts would be easier than tasks that involve description of textually simple texts.

Nunan (1989) proposed a set of criteria for identifying task difficulty and task sequencing, that is, task input factors, learner factors, and procedural factors. Task input factors include grammatical complexity of the input (for example, a text made up of simple sentences is likely to be simpler than one consisting of non-finite verb constructions and subordination), length of a text, propositional density, the amount of low-frequency vocabulary, the speed of spoken texts and the number of speakers involved, the explicitness of the information, the discourse structure and the clarity with which this is signaled. Learner factors include factors that the learner brings to the task of processing and producing language such as background knowledge, linguistic knowledge, confidence, interest, motivation, observed ability in language skills. Procedural factors concern procedures, that is, the operations that the learners are required to perform on input data, for example, task relevance, task complexity such as the number of steps involved, the cognitive demands the task makes on the learner, and the amount of information the learner is expected to process.

III. CURRENT PSYCHOLINGUISTIC MODELS OF TASK DIFFICULTY

In the discussion of variability in learner task performance in the SLA research, two views figure prominently in the current task-based studies. One view is that a predictable cognitive or linguistic outcome from any classroom task is practically impossible due to learner individual variance. This is best reflected in the conceptions of task-as-work plan and task-in-process (Seedhouse, 2005). The other view is that the nature of learners' performance is broadly predictable, individual variation in L2 task performance is not significant, and that what is found for one group of learners will *ceteris paribus* be found in others (Foster, 2009). Researchers holding this view are thus confident that a particular type of cognitive and linguistic outcome can be obtained for all learners, regardless of individual differences, through manipulation of certain task characteristics like task design along dimensions of cognitive complexity (Robinson, 2001a; Tavakoli and Skehan, 2005; Tavakoli and Foster 2008). Foster argues that such confidence in generalisable task performance is vital if we want to take the step of making pedagogical prescriptions or recommendations (Foster, 2009). A corollary of this argument is that information on how task characteristics may affect learners' L2 performance in a predictable way in spite of individual variation has potential for the definition of a variable of task difficulty in assessment of learner spoken performance.

Importantly, Skehan (1998, 2001) argues that different task characteristics, task type and performance conditions influence task difficulty as well as learner performance. If hierarchy of task difficulty is established, students with greater levels of underlying ability will then be able to successfully complete tasks which come higher on such a scale of difficulty (Skehan, 1998). In line with this argument, Skehan (2001) proposed the following psycholinguistic categories that will affect task difficulty:

- Interactivity: Interactive tasks produce greater L2 complexity and accuracy than non-interactive tasks.
- Familiar information: tasks that involve retrieval of information familiar to the participants result in higher L2 fluency and accuracy
- Degree of structure: tasks that contain a clear sequential structure produce greater L2 accuracy.
- Complex and numerous operations: The greater the number of online operations and transformation of material that are needed, the more difficult the task. This may impact upon greater complexity, but at the expense of accuracy and fluency (see Fulcher, 2003).
- Communicative stress: This refers to the logistics of task performance, e.g., time pressure, nature of the prompt and number of participants. As pressure increases, task difficulty increases. Similarly, tasks which involve many participants are more difficult than tasks which involve few participants. Difficult tasks direct learners' attention to context and divert attention away from form, thus generating less accuracy and fluency

To operationalize these psycholinguistic categories, Skehan further proposes a number of task dimensions and task conditions that he claims affect the difficulty of a task, which are presented in the table below (see Table 1). For

example, tasks which require the speaker to communicate few elements and which involve few participants are less difficult than tasks which require the speaker to communicate many elements and which involve many participants. Similarly, tasks which require the speaker to describe concrete objects are less difficult than tasks which require the speaker to communicate abstract notions. Skehan (1998) thinks that such research work on task difficulty has the potential to inform task selection and assist test developers in structuring test tasks and the conditions under which these tasks are performed in appropriate ways, which will allow us to predict the relative difficulty of a particular task so that the full range of candidates' ability will be tapped. A number of empirical studies that replicated Skehan's psycholinguistic framework on task difficulty in language

TABLE 1
TASK DIMENSIONS AND PERFORMANCE CONDITIONS INFLUENCING TASK DIFFICULTY

<i>Task dimension</i>		<i>Performance conditions</i>	
Less difficult	More difficult		
Number of participants or elements		few	many
Abstractness of information or task		concrete	abstract
Type of task information		immediate	remote
here-and-now	there-and-then		
Nature of operation required on task information		retrieval	transformation
Familiarity of task information		familiar	unfamiliar

(Skehan, 1998; see also Iwashita et al., 2001)

testing settings, however, have so far offered little support for Skehan's difficulty hypotheses. For example, Norris et al. (2002) did not find that the cognitive dimensions operationalized in their study based on Skehan's psycholinguistic framework proved particularly useful for estimating actual performance difficulty differences among test tasks. Norris et al. also reported some difficulty in operationalizing Skehan's task dimensions in their actual task design. Iwashita et al. (2001) drew on the framework proposed by Skehan (1998) and investigated predictors of task difficulty in the measurement of speaking proficiency through two different kinds of analysis: 1) a quantitative analysis of candidates' test discourse; 2) a Rasch analysis of candidates' test scores assigned by trained raters using analytical rating scales. Their statistical analysis revealed no systematic discourse variation associated with the other various task dimensions for performance conditions. Also drawing on Skehan's psycholinguistic framework, Elder, Iwashita and McNamara (2002) investigated the impact of performance conditions on perceptions of task difficulty in a test of spoken language. The participants in their study performed a series of narrative tasks whose characteristics, and the conditions under which they were performed, were manipulated and operationalized in terms of Skehan's framework. Elder et al.'s results provided little support for Skehan's framework in the context of oral proficiency assessment, but showed that differences in ability were found to be associated primarily with steps on the rating scale, rather than with differences in task's cognitive demand. Also worthy to mention is Weir, O'Sullivan, and Horai's (2004) study that examined how the difficulty of the Part 2 task in the IELTS speaking test can be manipulated using a framework on effects of planning time, response time, and amount of scaffolding on performance proposed by Skehan and Foster (1997). One of the four equivalent tasks in Weir et al.'s study was left unaltered, the other three were manipulated along three variables: planning time, response time, and scaffolded support. Weir et al.'s results suggest that there were significant differences in the responses of the three ability groups to the four tasks indicating task difficulty may well be affected differently for test candidates of different ability. Weir et al. thus caution us that simply altering a task along a particular dimension may not result in a version that is equally more or less difficult for all test candidates. "Instead, there is likely to be a variety of effects as a result of the alteration" (Weir et al. 2004:143). This resonates with Brindley and Slatyer's (2002) argument that "simply adjusting one task-level variable will not automatically make the task easier or more difficult" (p. 390).

A partially different model has been proposed by Robinson (Robinson 2001a; Robinson, 2001b; Robinson, Cadierno, and Shirai, 2009) who makes a distinction between task complexity, task difficulty, and task conditions (see Table 2). *Task complexity* concerns the cognitive demands of a task which contribute to between task variation in learner performance; *task difficulty* concerns those learner factors which contribute to differences between learners in their performance on a task; *task conditions* concern the interactive factors involved in pair and group work, such as the differences in participant background and role, and the nature of interactive participation a task requires. Within the dimension of task complexity, Robinson further makes a distinction between the resource-directing dimensions of task complexity (e.g., number of task elements, reasoning demands of the task, immediacy of information provided) that make communicative/conceptual demands, and resource-dispersing dimensions of task complexity (e.g., planning time, number of tasks, prior knowledge) that make increased performative/procedural demands on participants' attentional and memory resources. Robinson claims that these cognitively defined complexity factors can be manipulated, and as a result of such manipulation, the cognitive demand (e.g., attentional, memory, and reasoning demand) required for task

performance will vary, which leads to variation in the quality of learner language output.

TABLE 2
ROBINSON'S MODEL OF THE THREE SUPERORDINATE CATEGORIES OF TASK CHARACTERISTICS

<i>Task Complexity</i> (Cognitive factors)	<i>Task Condition</i> (Interactive factors)	<i>Task Difficulty</i> (Learner factors)
(a) Resource-directing variables cognitive/conceptual demands	(a) Participation variables interactional demands	(a) Ability variables and task-relevant resource differentials
+/- Here and now	+/- Open solution	h/l Working memory
+/- Few elements	+/- One-way flow	h/l Reasoning
-/+ Spatial reasoning	+/- Convergent solution	h/l Task-switching
-/+ Causal reasoning	+/- Few participants	h/l Aptitude
-/+ Intentional reasoning	+/- Few contributions needed	h/l Field independence
-/+ Perspective-taking	+/- Negotiation not needed	h/l Mind/intention-reading
(b) Resource-dispersing variables performative/procedural demands	(b) Participant variables interactant demands	(b) Affective variables and task-relevant state-trait differentials
+/- Planning time	+/- Same proficiency	h/l Openness to experience
+/- Single task	+/- Same gender	h/l Control of emotion
+/- Task structure	+/- Familiar	h/l Task motivation
+/- Few steps	+/- Shared content knowledge	h/l Processing anxiety
+/- Independency of steps	+/- Equal status and role	h/l Willingness to communicate
+/- Prior knowledge	+/- Shared cultural knowledge	h/l Self-efficacy

(Based on Robinson, 2001a; Robinson, 2001b; Robinson et al., 2009)

It can be seen that Robinson's approach has emphasized the independence of the dimensions of complexity and difficulty with complexity being a feature of the task, and difficulty operationalized in terms of two sets of learner factors categorized respectively as ability and affective variables. In other words, Robinson distinguishes task complexity which is task-dependent from task difficulty which is learner-dependent. Note that the cognitively defined task complexity factors in Robinson's framework have actually been hypothesized or empirically researched in some way as factors that determine task difficulty by other applied cognitive psychologists, SLA researchers, and L2 testing and assessment researchers (e.g., Brown, Anderson, Shilcock, and Yule, 1984; Brindley, 1987; Waks and Barak, 1988; Nunan and Keobke, 1995; Skehan and Foster, 1997; Skehan 1998, Norris et al 1998; Iwashita et al, 2001; Weir, O'Sullivan, and Horai 2004; Elder et al., 2002; Elder and Iwashita, 2005). It is also clear that Robinson's conceptualization of 'difficulty' is different from the way the term is defined in the measurement community where task difficulty is seen as the product of an interaction between many facets of the assessment including both features of the task and test-taker abilities and affect (Bachman, 2002; see also Elder and Iwashita, 2005), whereas Robinson operationalizes 'difficulty' in terms of learners' perceptions of task difficulty (2001a). For example, Robinson (2001) identified students' ratings of task difficulty through use of the following items (each item was rated on 9-point Likert scale):

- 1 I thought this task was easy/ I thought this task was hard
- 2 I felt relaxed doing this task/ I felt frustrated doing this task
- 3 I didn't do well on this task/ I did well on this task
- 4 This task was not interesting/ This task was interesting
- 5 I don't want to do more tasks like this/ I want to do more tasks like this (p. 41)

As mentioned above, Robinson suggests a direct relationship between *manipulation* of his cognitively defined task complexity features and *variation* in the quality of learner language output. In light of Bachman's (2002) comments, however, we may think that Robinson's task complexity features are not features purely inherent in tasks, as during the actual task performance, they involve the interaction of learner attributes (e.g., the actual cognitive operations that a learner applies in completing the task, which vary from one individual to another) with task characteristics (e.g., the hypothesized cognitive demand required for the task performance). Another caveat to the claim about a direct relationship between manipulation of task complexity and learner language production could be that it is possible for learners to successfully complete acts of communication within a task with little use of the target language (Fulcher, 2000).

It can also be seen that Robinson's conceptualization of task complexity exclusively focuses on the cognitive demands tasks make on learners, ignoring the role of linguistic features that may impact upon learners' task performance, as he argues that cognitively defined task complexity factors should be the sole basis of prospective task selecting and sequencing decisions (Robinson, 2001b). For Robinson, task complexity is thus essentially a cognitive issue. A danger of too exclusive a focus on defining task complexity in cognitive terms is that task performance is seen as in some way a simple projection of the learner's cognitive problem-solving abilities needed in achieving the task outcome. Vygotsky (1978), however, argues that language is a powerful mediating mechanism in cognitive activity and/or processes in the sense that language provides the tools for reflection and reasoning. For many second or foreign

language learners, linguistic demand of a language task tends to pose the greatest challenge or difficulty in completing the task, and constitute the major source of anxiety and frustration in target language learning and communication (Tsui, 1996; Li, 1998; Widdowson 1998, 2003; Carless, 2004; Littlewood, 2007). Discourse and text researchers (e.g., Brown and Yule, 1983) believe that differences in the 'task content' (e.g., ideational and interpersonal content carried by the design of different tasks) may set different linguistic demands and affect learners' comprehension and production of language. Other researchers (e.g., Brindley, 1987; Nunan, 1989, 2004; Tavakoli, 2009) also agree that a range of linguistic factors concerning the complexity of textual language input to tasks are likely to cause difficulty to learners, especially beginner- and intermediate-level learners. These include grammar, vocabulary, discourse structure, length of the text, the propositional density, amount of low frequency vocabulary, and the delivery speed and accents used in spoken texts. For example, it can be predicted that many second or foreign language learners will generally find a task that involves use of only simple present tense less linguistically challenging than a task that involves use of the present perfect passive, or subjunctives and conditionals. Note that Oxford (1996), and Dörnyei and Csizer (2002) remind us that in foreign language environments learners rarely surpass intermediate language proficiency. It thus follows that the nature and types of linguistic demands inherent in a task can be an important source of 'objective task complexity', which eventually contributes to perceived task difficulty (Nunan and Keobke, 1995).

Clearly, L2 pedagogy researchers tend to focus on the pure intrinsic cognitive demands of a task which is believed to contribute to between task variation in spoken and other kinds of performance (Robinson, 2009). Consequently, less attention has been paid by these researchers to the fact that the cognitive processing of certain kinds of information is likely to be socially or environmentally driven (O'Sullivan, 2000; Dörnyei, 2009). A number of studies have explored the role of social and interpersonal dimensions of a given task in the prediction of task difficulty. For example, Fulcher and Marquez-Reiter (2003) studied how learner L1 cultural background and pragmatic task features have an impact upon task difficulty. Their study showed that the pragmatic approach "may be preferable to the use of abstract psycholinguistic categories in the prediction of task difficulty, because the pragmatic categories appear to be more sensitive to how difficult a task may be for students from certain L1 cultural backgrounds" (Fulcher et al., 2003, p. 339). Also relying on a pragmatic approach, Taguchi (2007) examined how different types of pragmatic tasks may create different demands on performance. Taguchi found that in the case of pragmatic tasks that were more conventionalized, immediate, and were also less face-threatening, participants were able to perform quickly and easily, whereas tasks that were less common and more face-threatening were more difficult and took longer to produce.

According to Bachman (1990), communicative language ability can be described as consisting of both knowledge, or competence, and the capacity for implementing, or executing that competence in appropriate, contextualized communicative language use. Bachman and Palmer (1996) make the distinction between language competence and strategic competence. Language competence consists of organizational competence (e.g., grammatical and textual competence) and pragmatic competence (e.g., illocutionary competence and sociolinguistic competence) (see Littlemore and Low, 2006). Strategic competence is a general ability that enables an individual to use available resources by regulating online cognitive processes in accomplishing a communicative goal (Phakiti, 2008). It can be seen that there is a clear split between knowledge and processing action in Bachman's model of communicative language ability. In light of this, Skehan and Robinson's psycholinguistic models of task difficulty or complexity focus mainly on the cognitive demands tasks make on learners, and ignores how tasks actually engage learners' language knowledge, topical knowledge and affective schemata. Also, what is lacking in the literature is how interpersonal dynamics and affective dimensions of a given task may influence the cognitive processing of the task. Future research in this area will be useful given the fact that the goal of L2 curriculum is to produce learners who are able to communicate effectively in the target language of a particular speech community.

IV. CONCLUSION

Despite apparent differences in assumptions about task difficulty among L2 testing researchers and L2 pedagogy researchers discussed above, Skehan and Robinson's work has certainly been intellectually stimulating, and provided important insights into how task characteristics may affect learner task performance. Admittedly, determining task difficulty is an extremely complicated and difficult business. As Nunan (2004) emphasizes, "in addition to the number of factors to be taken into consideration, there is also the issue that the factors themselves are interrelated" (p.136). In other words, task difficulty involves complex clusters of factors which are interactive rather than discrete. As such, accurate a priori identification of a separate set of theoretically motivated and empirically sustainable 'task difficulty features' may be extremely hard to achieve. Future research thus needs to examine how different psycholinguistic factors, linguistic demands, learner attributes and social conditions interact to make a given task more or less 'difficult' for different learners through use of different research methods (e.g., sophisticated statistical procedures and qualitative procedures such as discourse analysis). It is also desirable to conduct longitudinal studies that enable researchers to connect spoken performance on a task at one time with stability or instability of performance over time, and to establish whether and how the interactive effects of task features and learner factors may be variably manifested.

REFERENCES

- [1] Bachman, L.F. (1990). *Fundamental considerations in language testing*. Oxford University Press.
- [2] Bachman L.F., Palmer A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- [3] Bachman, L.F. (2002). Some reflections on task-based language performance assessment. *Language Testing*; 19; 453-476.
- [4] Breen, M. (1987). Learner contributions to task design. In *Language learning tasks*, ed. C. Candlin, and D.F. Murphy. Englewood Cliffs, NJ: Prentice-Hall International and Lancaster University.
- [5] Brindley, G. (1987). Factors implicated in task difficulty. In Nunan(ed.), *Guidelines for the development of curriculum resources*. Adelaide: NCRC.
- [6] Brindley, G. and Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Language Testing*. 19, 369–394.
- [7] Brown, G., Anderson, A., Shilcock, R. and Yule, G. (1984). *Teaching talk: strategies for production and assessment*. Cambridge: Cambridge University Press.
- [8] Brown, G. and Yule, G. (1983). *Teaching the spoken language*. Cambridge: Cambridge University Press.
- [9] Brown, J.D., Hudson, T., & Norris, J. (1999). Validation of test-dependent and task-independent ratings of performance assessment. Paper presented at the 21st Language Testing Research Colloquium, Tsukuba, Japan.
- [10] Brumfit, C. J. (1984). *Communicative Methodology in Language Teaching*. Cambridge University Press.
- [11] Bygate, M., Skehan, P., & Swain, M. (eds). (2001). *Researching pedagogic tasks, second language learning, teaching and testing*. Longman.
- [12] Candlin, C. (1987). Towards task-based language learning. In C. Candlin and D. Murphy (eds.), *Language learning tasks*. London: Prentice-Hall. 5–22.
- [13] Carless, D. (2004). Issues in teachers' reinterpretation of a task-based innovation in primary schools. *TESOL Quarterly* 38.4, 639–662.
- [14] Crookes, G.V. (1986). Task classification: A cross-disciplinary review. Technical Report No.4. Honolulu, Center for Second Language Research, Social Science Research Institute, University of Hawai'i at Manoa.
- [15] Dörnyei, Z. (2009). Individual differences: Interplay of learner characteristics and learning environment. *Language Learning* 59, 230-248.
- [16] Dörnyei, Z., and Csizer, K. (2002). Some dynamics of language attitudes and motivation: Results of a longitudinal nationwide survey. *Applied Linguistics*, 23, 421-462.
- [17] Elder, C. and N. Iwashita. (2005). Planning for test performance: Does it make a difference? In Ellis, R. (ed.): *Planning and Task Performance in a Second Language*. John Benjamins.
- [18] Elder, C., Iwashita, N., and McNamara, T. (2002). Estimating the difficulty of oral proficiency tasks: what does the test-taker have to offer? *Language Testing*; 19, 347-368.
- [19] Ellis, R. (2003). *Task-based learning and teaching*. Oxford: Oxford University Press.
- [20] Foster, P. (2009). Task-based language learning research: expecting too much or too little? *International Journal of Applied Linguistics* 19, 247-263.
- [21] Fulcher, G. (1996). Testing tasks: issues in task design and the group oral. *Language Testing* 13, 23–51.
- [22] Fulcher, G. (2000). The 'communicative' legacy in language testing. *System* 28, 483-497.
- [23] Fulcher, G. and R. Marquez-Reiter. (2003). Task difficulty in speaking tests. *Language Testing* 20, 321–44.
- [24] Hu, G. W. (2005). English language education in China: Policies, progress, and problems. *Language Policy* 4.1, 5–24.
- [25] Iwashita, N., McNamara, T. and Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information-processing approach to task design. *Language Learning* 51: 401–36.
- [26] Li, D. F. (1998). 'It's always more difficult than you plan and imagine': Teachers' perceived difficulties in introducing the communicative approach in South Korea. *TESOL Quarterly* 32.4, 677–703.
- [27] Littlemore, J. and Low, G. (2006). Metaphoric competence, second language learning and communicative language ability. *Applied Linguistics* 27(2), p.268-294.
- [28] Littlewood, W. (2004). The task-based approach: Some questions and suggestions. *ELT Journal* 58.4, 319–326.
- [29] Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching*. 40, 243–249.
- [30] Long, M.H. (1985). A role for instruction in second language acquisition. In K. Hyltenstam and M. Pienemann (eds.), *Modelling and Assessing Second Language Acquisition*. Clevedon Avon: Multilingual Matters.
- [31] Long, M.H. and Crookes, G.V. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly* 26, 27–56.
- [32] Long, M., and Norris, J. (2000). Task-based language teaching and assessment. In M. Byram (Ed.), *Encyclopedia of language teaching*. London: Routledge.
- [33] Mislevy, R.J., Steinberg, L.S. and Almond, R.G. (2002). Design and analysis in task-based language assessment. *Language Testing*, 19, 477–496.
- [34] Norris, J.M. (2009). Task-based teaching and testing. In K. Van den Branden, M. Bygate, J.M. Norris (eds.), *Task-based language teaching, a reader*. Amsterdam: John Benjamins Publishing Company.
- [35] Norris, J.M., Brown, J.D., Hudson, T. and Yoshioka, J. (1998). Designing second language performance assessments. (Vol. SLTCC Technical Report #18). Honolulu: Second Language Teaching and Curriculum Center, University of Hawaii at Manoa.
- [36] Nunan, D. (1989). *Designing Tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.
- [37] Nunan, D. (1993). Task-based syllabus design: Selecting grading and sequencing tasks. In *Tasks in a pedagogical context*, ed. S.M. Gass and G. Crookes, 58-59. Clevedon: Multilingual Matters.
- [38] Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly* 37.4, 589–613.
- [39] Nunan, D. (2004). *Task-based Language Teaching*. Cambridge: Cambridge University Press.
- [40] Nunan, D., and K. Keobke. (1995). Task difficulty from the learner's perspective: perceptions and reality. *Hong Kong Papers in Linguistics and Language Teaching* 18: 1–12.
- [41] O'Sullivan, B. (2000). *Towards a model of performance in oral language testing*. Unpublished PhD dissertation, University of Reading.

- [42] Oxford, R.L. (1996). New pathways of language learning motivation. In R. Oxford (Ed.), *Language learning motivation: Pathways to the new century* (pp. 1-8). Hawaii: University of Hawaii at Manoa.
- [43] Phakiti, A. (2008). Construct validation of Bachman and Palmer's (1996) strategic competence model over time in EFL reading tests. *Language Testing* 25. 2, 237-272
- [44] Pica, T., Kanagy, R. and Falodun, J. (1993). Choosing and using communication tasks for second language instruction and research. In Crookes, G. and Gass, S., editors, *Tasks and language learning: integrating theory and practice*. Clevedon: Multilingual Matters, 9-34.
- [45] Robinson, P. (2001a). Task complexity, task difficulty and task production: Exploring Interactions in a componential framework. *Applied Linguistics* 21, 27-57
- [46] Robinson, P. (2001b). Task complexity, cognitive resources, and syllabus design. In Robinson, P. (Ed.), *Cognition and Second Language Instruction* (pp. 287-318). Cambridge: Cambridge University Press.
- [47] Robinson, P., T. Cadierno, and Y. Shirai. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics* 30, 533-544.
- [48] Savignon, S. (1983). Communicative competence: Theory and classroom practice. Reading, MA: Addison-Wesley.
- [49] Seedhouse, P. (2005). 'Task' as Research Construct. *Language Learning* 55, 533-570.
- [50] Shohamy, E. (1994). The validity of direct versus semi-direct oral tests. *Language Testing*. 11, 99-124
- [51] Skehan, P. (1998). *A Cognitive Approach to Language Learning*. Oxford University Press.
- [52] Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P.Skehan, and M. Swain (eds.) *Researching pedagogic tasks, second language learning, teaching and testing*. Harlow: Longman.
- [53] Skehan, P. and P. Foster. (1997). 'Task type and task processing conditions as influences on foreign language performance,' *Language Teaching Research* 1, 185-211.
- [54] Taguchi, N. (2007). Task difficulty in oral speech act production. *Applied Linguistics* 28/1: 113-135
- [55] Tavakoli, P. (2009). Investigating task difficulty: learners' and teachers' perceptions. *International Journal of Applied Linguistics* 19, 1-21.
- [56] Tavakoli, P., and P. Foster. (2008). 'Task design and second language performance: The effect of narrative type on learner output,' *Language Learning* 58: 439-73
- [57] Tavakoli, P., and P. Skehan. (2005). 'Strategic planning, task structure, and performance testing' in R. Ellis (ed.): *Planning and Task Performance in a Second Language*. John Benjamins.
- [58] Tsui, A.B.M. (1996). Reticence and anxiety in second language learning. In K.M. Bailey & D. Nunan (Eds.), *Voices from the language classroom* (pp.145-167). Cambridge: Cambridge University Press.
- [59] Vygotsky LS (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University.
- [60] Waks and Barak, (1988). Characterisation of cognitive difficulty level of test items. *Research in Science & Technological Education*, 6, 181-191.
- [61] Weir, C., O'Sullivan, B., and Horai, T. (2004). Exploring difficulty in speaking tasks: An intra-task perspective. *IELTS Research Reports* Volume 6.
- [62] Widdowson, H. G. (1998). Context, community, and authentic language. *TESOL Quarterly* 32.4, 705-716.

Zhengdong Gan is currently an Assistant Professor in the Department of English at the Hong Kong Institute of Education. He obtained his PhD in applied linguistics from the Hong Kong Polytechnic University in 2003. From 2006 to 2007, he worked as a postdoctoral research fellow in the School-Based Assessment Consultancy Team of the University of Hong Kong. His research interests include second language performance assessment, conversation analysis, and self-directed language learning. He has published in *Applied Linguistics*, *International Journal of Applied Linguistics*, *Australian Review of Applied Linguistics*, *Modern Language Journal*, and *Language Testing*.