Task-Based Language Learning: A Review of Issues

Peter Robinson
Aoyama Gakuin University

Theoretically motivated, empirical research into task-based language learning has been prompted by proposals for task-based language teaching. In this review I describe early and more recent proposals for how task-based learning can stimulate acquisition processes and the theoretical rationales that have guided research into them. I also describe taxonomies of task characteristics that have been proposed and claims made about the effects of task characteristics on interaction, attention to input, and speech production. I then relate the issues raised to findings described in the five empirical studies in this issue concerning the effects of pedagogic task design on the accuracy, fluency, and complexity of learner language; the influence of individual differences in cognitive and motivational variables on task performance; the extent to which tasks, and teacher interventions, promote the quantity and quality of interaction that facilitate L2 learning; and the generalizability of task-based learning research in laboratory contexts to instructed classroom settings.

Keywords task characteristics; task complexity; task sequencing; taxonomic description; theoretical rationales; abilities; attention; conceptualization; interaction; speech production

Over the past 30 years, proposals for task-based language teaching (TBLT) have drawn on a variety of claims about—and prompted further research into—processes thought to promote successful second language acquisition (SLA). Many important contributions to task-based learning research addressing these claims have appeared in Language Learning throughout this period (see, e.g., Gass, Mackey, Alvarez-Torres, & Fernandez-Garcia, 1999; Platt & Brooks, 2002; Seedhouse, 2005; Skehan & Foster, 1999; Yule, Powers, & Macdonald, 1992). Certain of these claims for SLA processes that task-work can facilitate feature throughout the present review article and are the focus of the five recent empirical studies published in Language Learning that follow it:

Correspondence concerning this article should be addressed to Peter Robinson, Aoyama Gakuin University, 4-4-25 Shibuya, Tokyo 150-8366, Japan. Internet: peterr@cl.aoyama.ac.jp
• Tasks provide a context for negotiating and comprehending the meaning of language provided in task input, or used by a partner performing the same task.
• Tasks provide opportunities for uptake of (implicit or explicit) corrective feedback on a participant’s production, by a partner, or by a teacher.
• Tasks provide opportunities for incorporation of premodified input, containing “positive evidence” of forms likely to be important to communicative success and that may previously have been unknown or poorly controlled.
• Tasks provide opportunities for noticing the gap between a participant’s production and input provided and for metalinguistic reflection on the form of output.
• Task demands can focus attention on specific concepts required for expression in the second language (L2) and prompt effort to grammaticize them in ways that the L2 formally encodes them, with consequences for improvements in accuracy of production.
• Simple task demands can promote access to and automatization of the currently emerged interlanguage means for meeting these demands, with consequences for improved fluency of production.
• Task demands can also promote effort at reconceptualizing and rethinking about events, in ways that match the formal means for encoding conceptualization that L2s make available.
• Sequences of tasks can consolidate memories for previous efforts at successfully resolving problems arising in communication, on previous versions, thereby strengthening memory for them.
• Following attempts to perform simpler versions, complex tasks can prompt learners to attempt more ambitious, complex language to resolve the demands they make on communicative success, thereby stretching interlanguage and promoting syntacticization, with consequences for improved complexity of production.
• Additionally, all of the above happen within a situated communication context that can foster form-function-meaning mapping and can do so in ways that motivate learners to learn.

Research into task-based learning has followed a trajectory, with the first four of the above-listed SLA processes being explored by early research into the effects of the interactive demands of tasks on learning (e.g., Brown, 1991; Crookes & Gass 1993a, 1993b; Day, 1986; Doughty & Pica, 1986; Gass & Varonis, 1994; Long, 1983; Long & Porter, 1985; Pica & Doughty, 1985; Pica, Young, & Doughty, 1987; Swain & Lapkin, 1995). Research into
task-based interaction has continued into the present (e.g., Alcon-Soler & García Mayo, 2009; Gass, Mackey, & Ross-Feldman, this issue; Mackey, 1999, 2007; Mackey & Gass, 2006; Shehadeh, 2001), with the result that broad findings for the effects of interaction and corrective feedback have accumulated to the extent that meta-analyses showing the positive contributions of each to SLA are now available (Keck, Iberri-Shea, Tracy-Ventura, & Wa-Mbaneleka, 2006; Mackey & Goo, 2007; Russell & Spada, 2006).

In contrast, the latter six above-listed SLA processes are the focus of more recent research into the cognitive demands and motivational impact of variously classified task characteristics and their effects on speech production, uptake, and longer term memory for input provided during task performance (e.g., Baralt, 2010; Bygate, Skehan, & Swain, 2001; Cadierno & Robinson, 2009; Dornyei & Kormos, 2000; Ellis, 2005; Gilabert, 2005, 2007; Gilabert, Baron, & Llanes, 2009; Ishikawa, 2007, 2008a, 2008b; Kim, 2008, 2009a, 2009b; Kuiken & Vedder, 2007a, 2007b; Michel, 2011; Nuevo, 2006; Revesz, 2009, 2011; Robinson, 2001b, 2007c; Robinson, Cadierno, & Shirai, 2009; Skehan & Foster, 1999, 2001; Tavakoli & Foster, this issue; Tavakoli & Skehan, 2005). No comprehensive meta-analyses of the effects of task characteristics in these areas of L2 production, uptake, and memory for input are available as yet, although syntheses of the accumulating findings about the effects of task characteristics contributing to their “complexity” on the accuracy, fluency, and complexity of L2 speech production are beginning to appear (Jackson & Suethanaporᓀkul, 2010) and promise to go some way toward resolving competing claims made by Robinson (2001a, 2001b, 2005) and Skehan (1998, 2009b; Skehan & Foster, 2001) about the effects of simple versus complex task demands on each (these claims and rationales for them are described in more detail in a subsequent section of this article). Taken together, then, these early, and more recently researched, SLA processes constitute a large part of what has been called the “cognitive-interactionist” rationale for the effects of instruction on SLA (see Ortega, 2007) and, in particular, for the positive effects of TBLT on SLA. Although TBLT clearly calls upon much more than the SLA processes described earlier, it is these processes and their contribution to task-based language learning (TBLL) that are the focus of the empirical studies in the present issue.

Researchers exploring sociocultural rationales for language pedagogy (e.g., Lantolf & Thorne, 2006; Negueruela & Lantolf, 2006; Swain, 2000; Swain, Kinnear, & Steinman, 2010; Swain & Lapkin, 1995) also address many of the issues raised above about how TBLT can facilitate SLA processes. From a Vygotskian perspective, learning is a social, collaborative endeavor in which
both “expert” and “apprentice” take part in a shared, goal-oriented activity. Task-work provides a context for such activity and for “interactional scaffolding” (Gibbons, 2009) of individual learners’ attempts to use the L2, by another learner or by a teacher (see Toth, this issue). Task demands may also prompt effort at reconceptualization of events requiring linguistic expression in the L2, third language (L3), or other language being learned, leading learners to “think-for-speaking” in ways characteristic of native speakers of the language being learned (Han & Cadierno, 2010; Jarvis & Pavlenko, 2008; Odlin, 2008; Robinson, 2007d; Robinson & Ellis, 2008; Pavlenko, 2011; von Stutterheim & Nuese, 2003). This may be particularly so along those dimensions of L2 task complexity that have been proposed to direct learners’ attention to the ways in which the conceptual and communicative demands of tasks are linguistically encoded (Cadierno & Robinson, 2009; Robinson, 2003b, 2005, 2007b; Robinson et al., 2009). Vygotsky argued that concepts are only internalized following a period of mediational support provided by experience of the concept—in conjunction with attempts to verbalize activities related to it with an interlocutor. Lantolf and Thorne (2006) have argued, drawing on Gal’perin’s ideas for putting Vygotsky’s ideas on concept development into L2 pedagogic practice, that “it is necessary to represent in a concrete material way the concepts to be internalized” (p. 305). Consequently, tasks (and material input to tasks) making conceptual demands on learners provide opportunities for verbalizing and internalizing concepts and for becoming aware of how their L2 scope and form of expression may differ from their L1 scope and linguistic encoding.

From TBLT to Task-Based Learning Research

Although SLA research has subsequently informed it, TBLT was, initially, a proposal for improving pedagogy with only a slight foundation in empirical research into the SLA processes listed earlier. Arising out of pedagogic proposals for a greater emphasis on communicative activities in language teaching (see, e.g., Brumfit & Johnson, 1979; Skehan, 2003; Valdman, 1978, 1980; Widdowson, 1978, for reviews), TBLT places the construct of “task” at the center of curricular planning. As Cook (2010) recently noted, TBLT “sees second language learning as arising from particular tasks that students do in the classroom. . . . In a sense it reconceptualizes communicative language teaching as tasks rather than the language or cognition-based syllabuses of communicative language teaching,” and TBLT is the approach to language teaching “that has attracted most attention in the past decade” (p. 512). Although attention to proposals for TBLT still primarily comes from teachers and educational
authorities charged with designing, implementing, and coordinating effective programs of language instruction at local, national, and international levels (e.g., Council of Europe, 2001; Leaver & Willis, 2004; Van den Branden, 2006; Willis & Willis, 2007), utilizing an increasing range of available instructional technologies (e.g., Thomas & Reinders, 2010), attention to TBLT has also come from SLA researchers concerned with explaining the effects of design features of tasks, and their implementation, on learning. SLA researchers often draw implications from their findings for classroom learning, testing, and program design. However, because much empirical SLA research into the effects of tasks on learning has been—and continues to be—done in experimental settings, there is a clear need to examine the generalizability of implications drawn from this research to actual classroom language performance on tasks, and the use of tasks for testing and syllabus design purposes. Kim (2009b) is an example of such needed bridging research, examining the generalizability of laboratory findings for the effects of task complexity on interaction-driven learning and development to EFL classroom contexts, with learners at different proficiency levels (see also Ellis, 1997; Gass, Mackey, & Ross-Feldman, this issue).

This important caveat aside, many of the current lines of inquiry that SLA research into task-based learning has pursued were originally prompted by the earliest pedagogic rationales for the use of tasks in language teaching. These were exploratory suggestions for connections between pedagogic practice and the acquisition processes they may stimulate. For example, in an article entitled “Towards Task-Based Language Learning” recapitulating points he had made earlier (Candlin, 1984), Candlin (1987) argued for:

the introduction of tasks as the basis for classroom action . . . They serve as a compelling and appropriate means for realizing certain characteristic principles of communicative language teaching and learning, as well as serving as a testing-ground for hypotheses in pragmatics and SLA . . . task-based language learning is not only a means to enhancing classroom communication and acquisition but also the means to the development of classroom syllabuses. (p. 5)

What Candlin (1987) was arguing for was the adoption of “tasks” as the “units” of syllabus design rather than linguistic units such as grammatical structures, functional phrases, or vocabulary lists (see Long & Crookes, 1992; Long & Robinson, 1998; Robinson, 2009; White, 1988, for reviews of units of analysis for the purposes of syllabus design). Yet Candlin was also begging the question of whether classroom tasks (designed and operationally delivered
in various ways) could also be shown to serve as constructs that confirmed theoretically motivated hypotheses about SLA processes and whether these hypothesized processes were prompted or inhibited by performing any one pedagogic task, in contrast to another, that differed along some dimension of its design features and/or features regarding their implementation. This argument is clearly still an issue at the heart of TBLL research (as the articles in this issue testify). Candlin then went on to say:

Tasks must . . . be defined and their means of operationalisation explained. It will be necessary to offer ideas for their classification and their targeting. Above all their centrality to the syllabus cannot be taken for granted without evaluating how they can be selected and sequenced in a principled fashion. (p. 5)

This issue raised by Candlin (1987)—of selecting tasks to be performed in sequences—is also at the heart of much current SLA research (see Robinson, 2007b, 2010; Robinson & Gilabert, 2007), which promises both implications for program design and insights into the acquisition processes that task-based learning can promote, across the differing timescales that institutions and other authorities set for language instruction. The sequencing issue involves consideration not only of how differently designed tasks might affect opportunities for learning in different ways but also of how sequences of tasks—in the different combinations that sequencing decisions afford, and across the different timescales performing them requires—impact upon these learning opportunities.

Features of Tasks and Their Implementation

A necessary starting point for studying the influence of sequences of tasks on learning, however, is to study their isolated impact on learning: in order to analyze how these effects (if they are found) are subsequently multiplied (or not) by sequences in which different tasks, or versions of the same task, are performed. This agenda requires both experimental research into task effects on learning and parallel studies of these effects in language learning programs. With these issues in mind, Candlin (1987) offered one pedagogically-operational definition of “task” (for a summary of many other definitions that have subsequently been proposed, see Samuda & Bygate, 2008, pp. 62–70). A task, Candlin (1987, p. 19) wrote, is “one of a set of differentiated, sequencable, problem-posing activities involving learners and teachers.” Then Candlin went on to describe
certain key features of tasks that task designers should be able to accommodate and that teachers could provide, to optimally promote classroom learning:

- **Input.** This is the written, visual, or aural information that learners performing a task work on to achieve the goal of the task.
- **Roles.** These are the roles that learners have in performing a task, such as information-giver and information-receiver.
- **Settings.** These are the grouping arrangements in and outside of classrooms for which pedagogy prepares learners to communicate.
- **Actions.** These are the procedures to follow in performing the task or the various steps that learners must take along the road to task completion.
- **Monitoring.** This is the supervisory process of ensuring that the task performance remains on track.
- **Outcomes.** These are the oral, written, and/or behavioral outcomes in which the task is intended to result.
- **Feedback.** This includes evaluation of the whole or parts of a task performance by the teacher or other learners, including corrective feedback on language use as well as other helpful feedback.

Many of these features of tasks and their implementation—in the years since Candlin (1987) first described them—have been operationalized and studied with respect to their influence on task-based learning and performance. These include, for example, the following

- the facilitating effects of linguistically versus elaboratively modified input on comprehension (Yano, Long & Ross, 1994);
- the effects of task role (Yule & MacDonald, 1990) and grouping arrangements (Brown, 1991) on the amount of interaction;
- the effectiveness of different types of corrective feedback on uptake and development during task-based interaction (Mackey & Goo, 2007);
- the effects of task different task characteristics on spoken (Foster & Skehan, 1999) and written (Kuiken & Vedder, 2007a, 2007b) outcomes.

**Tasks and the Designed Delivery of Instruction**

However, what of the larger implications of the study of task effects on learning for decisions about instructional design and program development, implementation, and evaluation? Samuda and Bygate (2008, pp. 58–60) made a useful distinction between task-supported, task-referenced, and task-based approaches to the role of tasks in instructional design. On the one hand, tasks have been
Robinson

Task-Based Language Learning

proposed as a useful means to support delivery of programs that follow a structural (Ellis, 1993) or lexical (Willis, 1996) syllabus. In this approach, tasks can be designed to focus on variously determined and sequenced grammar structures or lexical items (see Ellis, 2009, pp. 231–232; Toth, this issue), and “tasks are not necessarily used for assessment purposes and the syllabus itself may be defined by categories other than tasks” (Samuda & Bygate, 2008, p. 59).

In contrast, task-referenced approaches use tasks principally as a way of setting achievement targets and assessing the desired outcomes of a program of instruction, as in the case of the Australian Adult Migrant Language Program (AMEP; see Brindley & Slatyer, 2002) or the Common European Framework of Reference for Languages (Council of Europe, 2001). In most task-referenced approaches, decisions about the units of classroom instruction (however conceived) are made prospectively, bearing in mind progress toward outcomes of task achievement. However, Candlin (1984, 1987) thought tasks could be used as units of analysis for referencing, “retrospectively,” what happened in classrooms over a course of instruction, with regard to what language was learned as a consequence of tasks performed, what content the tasks involved, and what activities the tasks led to for teachers and learners.

Writing at around the same time as Candlin (1984), Long (1985) described a different perspective on how the information about tasks and their influence on learning could be drawn on and integrated into “prospective” decisions about instructional program design. This is what Samuda and Bygate (2008) referred to as an example of the task-based approach to the use of tasks to organize program delivery and assessment. Essential to his proposals, Long distinguished between “target tasks”—what learners are expected to do on exit from instructional programs (identified via a needs analysis)—and “pedagogical tasks”—the tasks teachers and students work on in classrooms, which can be gradually increased in complexity so as to approximate target task demands. In contrast to a number of others who have subsequently argued for the important contribution that tasks can make to language pedagogy (e.g., Ellis, 2003; Skehan, 1998; Willis, 1996), Long (1985, 1998, 2005; Long & Crookes, 1992; Long & Norris, 2004) proposed that needs analysis was the necessary first stage in course design, allowing target tasks to be identified. Following this first stage, Long argued, target tasks should then be classified into task types, and following this, pedagogic versions of each task type should be designed. Subsequently, pedagogic tasks should be sequenced to form the syllabus, then performed in classrooms, and the extent to which they enable target task performance to be achieved should be evaluated with criterion-referenced, performance tests. Long’s proposal for integrating research on task
effects with other components of L2 program design, such as needs analysis, task sequencing, and evaluation outcomes, closely mirrored models proposed to guide educational decision making in English for specific purposes (ESP) programs in the 1970s and 1980s (e.g., Harper, 1984; Jupp & Hodlin, 1975; Munby, 1978) and in other domains of instruction following the generic ADDIE (Analysis, Design, Development, Implementation, Evaluation) model of instructional design (see Dick & Carey, 1996; Gagne, Wager, Golas, & Keller, 2005).

Recent Taxonomic Frameworks and Theoretical Rationales for Task-based Learning

Central to much recent task-based learning research are issues of the influence of task characteristics on learning and performance, the relative complexity of tasks having one or another of them, and the development of taxonomic models to facilitate prospective pedagogic decision making about sequencing tasks. These concerns aim to make instructional design decisions relevant for learners in L2 learning programs (Garcia-Mayo, 2007; Robinson, 2007b; Robinson & Gilabert, 2007). Essential to issues of program design to facilitate exit-program evaluations of success in learning are taxonomies of task characteristics (in any area of instruction) and rationales for how they can be used to implement classroom decisions about what task is presented to learners, at what point in an instructional program, and how such decisions can be validated (or not, and so changed) on the basis of exit-program evaluation procedures (see Clark & Elen, 2007; Merrill, 2007; Reigeluth, 1999; Reigeluth & Carr-Chellman, 2009).

Arising in part out of early speculations about what task characteristics may be influential on interaction and SLA (e.g., Candlin, 1984, 1987; Long, 1983; Prabhu, 1987) a number of increasingly elaborate taxonomies of task characteristics have been proposed as a basis for pedagogic task design (e.g., Pica, Kanagy, & Falodun, 1993; Prabhu, 1987; Robinson, 2007b; Skehan, 1998). Moreover, an increasing number of studies have operationalized characteristics of pedagogic tasks based on one or another of these taxonomic frameworks and have studied their effects (at different levels of complexity) on the amount of interaction they promote (Kim, 2009a, 2009b; Nuevo, 2006) and on the accuracy, fluency, and complexity of task outcomes (Gilabert, 2007; Tavakoli & Foster, this issue); language development (Collentine, 2010); and uptake of corrective feedback (Baralt, 2010; Revesz, 2009).
As with taxonomies of task characteristics, theoretical rationales motivating empirical research into task-based learning have also become more elaborate over this 30-year period, reflecting a development from the early emphasis on how task-based learning can facilitate comprehension of input through interaction and negotiation of meaning to more recent emphases on how task-based learning can facilitate attention to output and the development of increasingly targetlike speech production. Certain theoretical rationales have been offered for the predicted effects of tasks on learning, and specific taxonomic frameworks have been proposed for classifying task characteristics.

**Theoretical Rationales for Task-Based Learning Research**

**The Procedural Syllabus**
The first large-scale attempt to implement TBLT and to develop a theoretical rationale for it took place in India, between 1979 and 1984, and is described in Prabhu (1987). In his account of the theoretical motivation for the Bangalore Project and the task-based “procedural syllabus” it implemented, Prabhu argued that:

> task-based teaching operates with the concept that, while the conscious mind is working out some of the meaning-content, a subconscious part of the mind perceives, abstracts, or acquires (or re-creates as a cognitive structure) some of the linguistic structuring embodied in those entities, as a step in the development of an internal system of rules. The intensive exposure caused by the effort to work out meaning-content is thus a condition which is favorable to the subconscious abstraction—or cognitive formation—of language structure. (pp. 70–71)

Explicit instruction in grammar in the project was avoided because “teaching a descriptive grammar is likely—as has been pointed out at various times in the history of language pedagogy—to promote in learners an explicit knowledge of that grammar, rather than a deployable internal system” (pp. 72–73). Prabhu’s cognitive rationale for TBLT is thus compatible with Krashen’s (1982) claim that *comprehensible input* is necessary for learning and that “the effort to work out meaning-content” promotes incidental learning of tacit or implicit knowledge. Prabhu explained that the Bangalore Project “did not use group work in the classroom” because learner-learner interaction did not promote development of interlanguage (although see Adams, 2007, for evidence that it does), and may lead to fossilization:
The effect of learner-learner interaction will largely be a firming up of learners’ [interlanguage] systems: each learner output will reinforce the internal systems of others without their being a corresponding process of revision. . . The principle that interaction between the teacher and the learner, or between a text/task on paper and the learner, is more beneficial than interaction between one learner and another is thus part of the concept of learning which lies behind task-based teaching. (pp. 81–82)

The Interaction Hypothesis and Focus on Form
In contrast to Prabhu (1987), Long (1983, 1989) argued that the interaction task-work promotes is important because it not only provides one way in which input can be made comprehensible but additionally serves as a context for attending to problematic forms in the input and output during task-work. Such learner-driven attention to form, contingent on negotiation of meaning, can speed mapping of form-meaning relations and prompt interlanguage change in ways that respected each learner’s own developmental trajectory (see Keck et al., 2006; Mackey, 2007). Consequently Pica et al. (1993) described a taxonomy of task characteristics in order to promote further research into which of these characteristics optimally promoted interaction work. Drawing on Schmidt’s (1990; Schmidt & Frota, 1986) claim that attention to aspects of the surface structure of utterances, accompanied by the subjective experience of awareness or “noticing” them, was essential to learning, Long (1991, pp. 45–46) also argued that pedagogy could facilitate this process of attending to and noticing communicatively redundant, nonsalient features of the L2. This could be achieved by interventions that prompt a “focus on form” that “overtly draws students’ attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication.” Doughty and Williams (1998) described a series of techniques, and proposals for research into them, for providing such additional “cognitive processing support” (p. 3) to learners. Decisions about focus on form could be made offline, proactively, leading, for example, to textual enhancement of problematic forms in the input to task performance, or could take place reactively, online, as in the case of recasts of selected forms in learners’ output during task performance.

The Output Hypothesis
Swain (1995, pp. 125–126) argued that attention to output, whether this occurs in interaction between learners or not, has a facilitating role, because “in producing the target language. . . learners may notice a gap between what they want to say and what they can say, leading them to recognize what they do not know,
or know only partially.” Producing language also offers learners opportunities for testing hypotheses about well-formedness and for metalinguistic reflection on L2 form. Izumi (2003) and Kormos (2006) have described the stages of L2 speech production at which attention can operate to promote the three effects that Swain described, and current research is concerned with the extent to which the attentional demands of pedagogic tasks can be manipulated to lead learners to “push” their output.

A theoretical construct of interest here is the notion of attention as “capacity” and the related issues of how increasing the attentional demands of tasks affect the fluency, accuracy, and complexity of speech production (see Housen & Kuiken, 2009a, 2009b). Clearly, the human information processing system is limited in its ability to process and respond to information in the environment; but are breakdowns in performance caused by limits on attentional capacity? Skehan (1998, 2009) argued for this position, claiming that capacity limits on a single pool of attentional resources leads to decrements in the fluency, accuracy, and complexity of L2 speech when tasks are high in their attentional, memory, and other cognitive demands. In this view, capacity limits prevent learners attending to both accuracy and complexity of production on cognitively demanding tasks, leading learners to trade-off attention to one at the expense of the other. A contrasting position has been proposed by Robinson (2003a, 2007b), who argued that attentional capacity limits are an unsatisfactory, post hoc explanation for breakdowns in attention to speech. Following, in part, arguments made by Allport (1987), Neumann (1987), and Sanders (1998), Robinson (2003a, 2007d) suggested that breakdowns in “action-control,” not capacity limits, lead to decrements in speech production and learners’ failure to benefit from the learning opportunities attention directing provides. Consequently, increasing complexity along various dimensions of tasks, such as increasing the amount of reasoning a task requires, promotes greater effort at controlling production and more vigilant monitoring of output. This increased complexity leads to greater accuracy and complexity of L2 production when compared to performance on simpler task versions that require little or no reasoning.

**The Limited Capacity Hypothesis**

Schmidt (1990, p. 143) noted that “Task demands are a powerful determinant of what is noticed” in experimental settings “and provide one of the basic arguments that what is learned is what is noticed. . . . The information committed to memory is essentially the information that must be heeded in order to carry out a task.” The extent to which this is true of L2 learning tasks in classrooms is an important issue for the design of materials and instruction. Because much
TBLL research has taken place in experimental settings, the relevance of findings from these studies to instructional decision making in diverse classroom settings, and for diverse populations of learners, require generalizability studies, as Gass, Mackey and Ross-Feldman (this issue) have described. Skehan (1998) provided the first extended psycholinguistic rationale for the effects of certain aspects of task demands on attention, “noticing,” and speech production, focusing in particular on the extent to which having time to plan a task led to increases in the accuracy, fluency, and complexity of speech produced when compared to performance on tasks for which planning time was not available. Skehan’s limited capacity hypothesis (p. 97) proposed that more demanding tasks “consume more attentional resources... with the result that less attention is available for focus on form”: Therefore, sequencing tasks from less cognitively demanding to more demanding optimizes opportunities for attentional allocation to language forms. In Skehan’s view, task design is a means to promote “balanced language development” in the areas of accuracy, fluency, and complexity of production. This process can be facilitated because certain task characteristics “predispose learners to channel their attention in predictable ways, such as clear macrostructure towards accuracy, the need to impose order on ideas towards complexity, and so on” (Skehan, 1998, p. 112; cf. Tavakoli & Foster, this issue). However, due to limitations in attentional resources, tasks can lead either to increased complexity or accuracy but not to both—so learners must “trade-off” attention to one aspect of production to the detriment of the other. Tasks should therefore be sequenced by choosing those tasks with characteristics that lead to each, at an appropriate level of difficulty, as determined by three factors: (a) Code complexity is described in “fairly traditional ways,” as in descriptions of structural syllabuses, or developmental sequences (Skehan, 1998, p. 99); (b) cognitive complexity is the result of the familiarity of the task, topic, or genre, and the processing requirements; information type, clarity, and organization; and amount of computation required; and (c) communicative stress involves six characteristics, including time pressure, number of participants, and opportunities to control interaction.

The Cognition Hypothesis

The cognition hypothesis provides a theoretical rationale for the effects of task demands on language learning that differs from Skehan’s proposal. This rationale is rooted in claims by Cromer (1973, 1991) that conceptual development creates the conditions for first-language (L1) development, and in claims by Slobin (1993) that parallels in L1 and L2 development (as revealed in data reported in Perdue, 1993a, 1993b) are evident because adult L2 learners
initially attempt to linguistically encode concepts that emerge earliest during child development. In this view, learners adopt, by default, the ontogenetically determined order of meaning (concept)-form (language) mapping in attempts to produce the L2 (see Robinson, 2005). For instructed SLA, the fundamental pedagogic claim of the cognition hypothesis is that in task-based syllabi, pedagogic tasks should be sequenced solely on the basis of increases in their cognitive complexity, which mirror the sequences in which children are able to meet the cognitive demands of tasks during L1 acquisition, and that such sequences provide optimal support for L2 learners in their attempts to use accurate and complex language at the level needed to meet real-world target task demands (Robinson, 2001a, 2005). In this account, learners do not trade-off attention to accuracy against attention to complexity of production: Rather, on some dimensions of task demands (described below), increasing complexity is argued to promote more accurate, grammaticized production and more complex, syntacticized utterances. Referring to Givon’s (1985) distinction between pragmatic and syntactic modes of production, Robinson (1995, 2003b) has argued that simple task demands elicit the pragmatic mode (characterized by loose coordination of clauses and little use of grammatical morphology) in contrast to complex task demands. Complex task demands lead to greater effort at conceptualization and elicit the morphologically richer and structurally more complex syntactic mode, in line with Givon’s (2009, p. 12) claim that “More complex mentally-represented events are coded by more complex linguistic/syntactic structures.”

To guide research into these claims, and also pedagogy, Robinson (2007b) proposed an operational taxonomy of task characteristics. This taxonomic, Triadic Componential Framework (TCF) distinguishes three categories of task demands implicated in real-world task performance. Task condition refers to interactive demands of tasks, including participation variables (e.g., open vs. closed tasks) and participant variables (e.g., same vs. different gender). Characteristics of tasks distinguishing the demands made by task conditions thus include many of the factors described in Pica et al.’s (1993) taxonomy. A second category of task difficulty concerns individual differences in learner factors, such as working memory capacity, which can affect the extent to which learners perceive task demands to be difficult to meet. These factors explain why two learners may find the same task to be more or less difficult than each other, and, broadly speaking, these individual difference factors may combine to result in “aptitudes” for certain kinds of task performance and task-based learning. Task complexity refers to the intrinsic cognitive complexity of tasks, such as their reasoning demands. Robinson distinguished three different kinds of reasoning.
demands: (a) spatial reasoning, as involved in navigating through, and giving directions about, places like cities while driving, or giving instructions on how to back a car into a small parking space; (b) causal reasoning, involved in understanding and explaining why a natural or mechanical event occurred (why a bridge fell down in a storm, why exchange rates fluctuate); and (c) intentional reasoning, as involved when explaining behavior with reference to the intentions, beliefs, and desires of others (why Tom suddenly left the party, why Jill has stopped speaking to Mary, etc.).

The TCF distinguishes task features affecting the cognitive complexity of tasks along two dimensions. Resource-directing dimensions of complexity affect allocation of cognitive resources to specific aspects of L2 code. For instance, tasks that increase in their intentional-reasoning demands require linguistic reference to the mental states of others. These demands should therefore direct learners’ attention to forms needed to meet them during communication, such as psychological state terms in English (e.g., believe, wonder). These forms may be currently known but not well controlled, or if they are unknown, then attempts to complete the task may make them salient and “noticeable.” By increasing complexity along these dimensions, initially implicit knowledge of the L1 concept-structuring function of language (see Talmy, 2000) becomes gradually explicit and available for change during L2 production. In contrast, resource-dispersing dimensions do not do this: Making a task complex by removing planning time does not direct the learner’s attention to specific aspects of L2 code but rather disperses attention over many linguistic and other features. Increasing task demands along these dimensions has the effect of gradually removing processing support (such as planning time) for access to current interlanguage; thus, practice along them requires, and should encourage, faster and more automatic L2 access and use. Based on this resource-directing/dispersing distinction and the TCF, Robinson (2010) proposed two operational principles for sequencing tasks in a task-based syllabus: (a) Sequencing should be based only on increases in cognitive complexity, (b) increase resource-dispersing dimensions of task complexity first (to promote access to current interlanguage), then increase resource-directing dimensions of complexity (to promote development of new form-function mappings, and destabilize the current interlanguage system).

Task Demands and Stages of Speech Production
Theoretical rationales for the influence of task demands on writing and both reading and listening comprehension currently lag behind rationales for their effects on speech production in articulating linkages between rationales proposed
and explanatory psycholinguistic mechanisms. With respect to the latter, both Skehan (2009a, 2009b; cf. Tavakoli & Foster, this issue; Tavakoli & Skehan, 2005) and Robinson (1995, 2005) have drawn on Levelt’s model of speech production in their psycholinguistic rationales for how task demands should affect L2 speech performance, as have others who have been more generally concerned with identifying the mechanisms involved in producing L2 speech and responding to negative feedback on it (e.g., Bygate, 1999; Doughty, 2001; Izumi, 2003; Kormos, 2006, 2011; see also de Bot, 1996, 1998). Levelt’s model of speech production identifies stages in which speech is assembled for production, beginning with a conceptualization stage, leading to preparation of the preverbal message, followed by stages of lexical and grammatical encoding, articulation, and (optionally, possibly individually initiated or coconstructed) monitoring of utterances following production (which can lead to self-repair, see Gilabert, 2005, 2007; Kormos, 1999; Swain & Lapkin, 1995). Levelt’s model is a stage model (for arguments against stage models in general, see Larson-Freeman & Cameron, 2007; and see Dell, 1986, and Dell, Juliano, & Govindjee, 1993, for alternative spreading activation models of speech production), but preparation of speech at the stages Levelt described is proposed to be performed in parallel, and processing is incremental, so all stages of speech production are simultaneously active, with feed-forward and feedback operations connecting these stages.

Drawing on Levelt’s model (1989; Levelt, Roelofs, & Meyer, 1999) of speech production, Robinson (1995) argued that increasing the conceptual demands of tasks (naturally) leads to greater effort at conceptualization and “macroplanning” at the stage of message preparation, thus “creating the conditions for development and re-mapping of conceptual and linguistic categories” (Robinson et al., 2009, p. 537), during subsequent “microplanning” and the lexicogrammatical encoding stage into which macroplanning feeds. In Levelt’s model, the conceptualization stage generates a “preverbal message”: “the message should contain the features that are necessary and sufficient for the next stage of processing—in particular for grammatical encoding” (Levelt, 1989, p. 70). Therefore, greater effort at conceptualization during message preparation, induced by conceptually demanding tasks, should lead to what Dipper, Black, and Bryan (2005, p. 422) called “paring down” of conceptual information into a “linguistically relevant representation” for subsequent encoding, at the microplanning stage, with positive consequences for accurate and complex performance.

Skehan (2009a, 2009b) has recently proposed a fine-grained analysis of how some aspects of task demands that make them more difficult (what
Skehan, 2009a, calls “complexifying/pressuring” influences) or less difficult (what Skehan calls “easing/focusing” influences) are related to the stages of speech production that Levelt (1989) described. Skehan (2009a) argued that it is the connections he proposes between task demands and their influence on stages of speech production that cause the effects of task demands on the accuracy, fluency, and complexity of spoken performance. For example, where a task requires dynamic relations between task elements to be described (as when describing how cars were moving prior to the occurrence of a traffic accident), then this will lead to more complex language performance, compared to describing concrete, static information (e.g., the arrangement of furniture in a room). In this case, the locus of differences in performance on the two tasks is attributed to differences at the conceptualization stage of message preparation. In contrast, monologic tasks lead to more complex spoken performance than dialogic tasks, but this is attributed to differences at the formulation stage and subsequent lexisosyntactic encoding procedures. Skehan, then, basically argued that different task characteristics (either intrinsic to their design or characteristics of their implementation) have effects on spoken performance that are caused by more or less effort at the conceptualization stage versus task characteristics that result in effects caused by processing for speech production at the lexical and syntactic encoding and formulation stages. What is valuable about Skehan’s proposal is its attempt to ground discussion of task characteristics, and their effects on learning and performance, in a model of speech production mechanisms, thereby providing a psycholinguistic rationale for the effects of what has been metaphorically referred to in the SLA literature as “pushed” output. Much more theory and research in this area is needed and is likely to be forthcoming.

**Taxonomies of Task Characteristics**

For TBLL research to produce cumulative findings, with application to pedagogy, a taxonomy of task characteristics is needed. A major aim of task-based learning research is to establish such a taxonomy. The benefits of an agreed taxonomic framework for research into task effects are clear. On the one hand, a taxonomy describes a finite list of task characteristics and categories of them. This can serve as a focus for concerted research into the effects of those characteristics on learning, when used to deliver different kinds of task content, and with learners at different levels of proficiency, or with different task aptitude or other profiles, thus enabling the generalizability of findings for the characteristics and their effects to be charted. On the other hand, a taxonomy
with instructional relevance identifies characteristics of pedagogic tasks that can be used to approximate the demands made by a wide variety of real-world target-task performance objectives. There are thus constraints that an operational taxonomy of task characteristics should satisfy if it is to be both useful for coordinating research and with application to pedagogy. Two pedagogic constraints that a taxonomy must meet are (a) that it is detailed enough in its listing of characteristics to allow a wide variety of target-task performances to be approximated and (b) that the characteristics it identifies must be feasible, allowing task designers in different programs using the taxonomy to make similar decisions about how to operationalize the characteristic. An additional theoretical constraint is (c) that the taxonomy is motivated by what is currently known about SLA processes and development. A final hybrid theoretical-pedagogic constraint is that (d) it should be possible to use the taxonomy to classify and sequence tasks, following some metric for combining task characteristics in sequences and in ways that are hypothesized to lead to language learning.

For extensive reviews of the many task characteristics, and categories of them, that have been proposed, see Ellis (2003), Nunan (1993, 2004), Robinson (2001a, 2007b), and Samuda and Bygate (2008). Three taxonomies in particular continue to predominate in task-based language research. The earliest of these was proposed by Pica et al. (1993), who described a typology of “communication tasks,” each with different configurations of activity (the role relationships between participants, and direction of information flow between them) and goal (the outcomes the task was expected to result in). They further analyzed the goals and activities tasks that could result in terms of the opportunities they provided for learning. For example, where task goals are the same or convergent, then they claimed it was “expected” that the task would lead to (a) comprehension of input, (b) feedback on production, and (c) interlanguage modification. In contrast, where interactants have divergent goals, then each of these was only “possible,” therefore providing fewer opportunities for learning. The task types that Pica et al. identified were Jigsaw, Information-gap, Problem-solving, Decision-making, and Opinion exchange. Pica et al.’s typology therefore expands on the three types of task that the Bangalore Project used and that Prabhu (1987, pp. 46–47) had ordered from Information gap, to Reasoning gap, to Opinion gap in terms of the demands these task types made on learners. This taxonomy satisfies the feasible constraint described previously and also the theoretical constraint (with respect to interactionist rationales for SLA), but it fails to meet the first constraint because it lacks sufficient detail (consisting of only five task types) and it also fails to meet the fourth constraint
because it provides no information about how task types can be sequenced optimally for learners.

Skehan’s (1998, 2003; Skehan & Foster, 2001) taxonomy, described earlier, consists of categories of task characteristics that contribute to code complexity, cognitive complexity, and communicative stress and builds on Candlin’s (1987) proposals for factors that may influence task performance and learning. For example, the vocabulary load a task involves affects its code complexity, the clarity and sufficiency of information provided affects its cognitive complexity, and the number of participants and time limits affect its degree of communicative stress (see Kuiken & Vedder, 2007b, for discussion). In more recent work, Skehan (2009a, 2009b) has maintained a focus on how task characteristics can complexify (e.g., by requiring greater quantities of information to be expressed) and pressure task performance (by imposing tight time constraints), and as observed above, Skehan has linked complexifying/pressuring task characteristics to Levelt’s (1989) model of stages in speech production, which cause their effects on the accuracy, fluency, and complexity of learners’ language. Skehan’s taxonomy is theoretically motivated, not with respect to interactional influences on learning (as is the case with Pica et al.’s [1993] taxonomy) but with respect to the effects of tasks on the psycholinguistic processes causing variation in speech production. The characteristics that Skehan described are also feasible because task designers should consistently be able to manipulate time pressure, quantities of needed information, and so forth. However, as with Pica et al.’s taxonomy, the characteristics described lack sufficient detail to link pedagogic versions to a wide variety of real-world task performances (although Skehan has been clear that his model is not intended to address how pedagogic tasks can be used to do this), and no metric is offered for sequencing the characteristics described.

Robinson’s (2001a, 2007b) TCF, described earlier (and see the Appendix), combines elements from both Pica et al.’s (1993) and Skehan’s (1998, 2003) proposals. Following the rationale described previously, the category of task characteristics contributing to their cognitive complexity are proposed to affect speech production in different ways along resource-directing and dispersing dimensions. The category of task characteristics describing the different participation, and participant factors involved in task performance, are proposed to affect the extent of interaction in different ways. The third category of task difficulty in the TCF involves learner factors, which may influence performance and learning on tasks having characteristics distinguishing their interactive or cognitive demands (see Albert & Kormos, this issue). For example, along the ± single task dimension of task complexity, individuals high in working memory
capacity and in the ability to switch attention between task demands may find dual tasks (requiring two things to be done simultaneously, such as answering a phone call while monitoring a TV screen in the office) to be less difficult than those lower in these abilities. Similarly, when the solution to a task learners are performing is indeterminate and not fixed (+ open) as opposed to determinant and fixed (+ closed), then individual differences in measures of emotional control, such as openness to experience and tolerance of ambiguity (Costa & McCrae, 1985; Furnham & Ribchester, 1995), may predict more, or less, successful engagement in task participation to meet these goals (with those more open to experience and more tolerant of ambiguity adapting better to participation in open tasks, and vice versa). It is not yet clear what the ability and affective factors are that contribute to perceptions of task difficulty, and so both promote and mitigate successful performance on the simple and complex task characteristics listed under the category of Task Complexity in the appendix or affect performance under different interactional Task Conditions listed there.

Research into individual differences in affective and ability factors and the extent to which they affect task performance is much needed (see Albert & Kormos, this issue, for one example) because if these links can be established through research, they could be used to operationalize batteries of individual difference measures that can be used to profile “task-aptitudes”—with the twin aims of matching learners to tasks that optimize their opportunities for successful L2 learning and performance and of supporting them when their ability and affective profiles are not well matched to the demands tasks make on them (see Robinson, 2007a, 2007a; Snow, 1994).

In summary, with regard to the criteria that task taxonomies should meet in order to be pedagogically useful and acquisitionally optimal, the TCF is more detailed than Pica et al.’s (1993) or Skehan’s (1998) taxonomies while being equally feasible and theoretically motivated. It has the advantage too, of an associated sequencing metric described earlier, particularly, that (a) sequencing pedagogic versions of target tasks should be based only on increases in cognitive complexity and that (b) resource-dispersing dimensions of task complexity should first be increased (to promote access to current interlanguage) and then resource-directing dimensions of complexity should be increased (to promote development of new form-function mappings and destabilize the current interlanguage system). Whether this sequencing procedure (intrinsically linked to the TCF taxonomy) is optimal for promoting successful task performance and language learning is, as yet, empirically unresolved, because research into it has only recently begun (e.g., Romanko & Nakatsugawa, 2010).
The Articles in This Issue

The articles in this issue all describe careful empirical studies, selected from recent issues of *Language Learning* because they address one or another of the issues about TBLL raised earlier. The first two articles, by Tavakoli and Foster and by Albert and Kormos, dealt centrally with the effects of tasks on outcome measures of speech production. Tavakoli and Foster began by observing that research into the effects of tasks on production is important because it can “illuminate the proposition that task performance in itself drives inter-language change by causing learners to attend to and retain information about the target language as they use it” (p. 38). Additionally, research identifying “features of tasks that impact on a learner’s processing” may help provide “empirically sound principles for classroom materials design” (p. 38) and inform decisions about which tasks to choose to “guide a learners focus of attention to particular aspects of the language being learned” (p. 38). Drawing on Skehan’s (1998, 2009a, 2009b) limited capacity hypothesis—that given finite attentional resources, learners will “prioritize one aspect of performance, such as being accurate, over another, such as being suitably fluent or complex” (p. 41)—they hypothesized that two characteristics of narrative tasks (loose vs. tight *structure*, and storyline *complexity*) will have these effects. Their hypotheses are that (a) narratives with tight structure will lead to more accurate language, whereas those with loose structure will lead to greater fluency; and (b) compared to less cognitively complex narratives that only foreground events, more complex narratives that both foreground and background events will lead to more syntactically complex and lexically diverse language but have no effects on accuracy. The research found that in narrative performance on the task that has tight structure (and so is simple on what Robinson (2003a) called this resource-dispersing dimension of cognitive demand) and that has two storylines (and so is simultaneously complex on this resource-directing dimension of task demands in Robinson’s framework), there is greater accuracy, fluency, and complexity, as the cognition hypothesis predicts. However, other findings, such as performance on a narrative with loose structure and two storylines, appear to confirm Skehan’s predictions, because it appears that the higher fluency and complexity this narrative elicits is traded off against lower accuracy. Tavakoli and Foster also argued that their results for greater accuracy on narrative tasks with tight structure can be interpreted as support for Skehan’s claim that some effects task characteristics have on production are caused (as in this case) by the extent to which they free up attention at the conceptualization stage of speech production. Tight structure frees up attention
because it reduces the need for effort at macroplanning of coordinated narrative event descriptions during conceptualization, and attention freed up from effort at conceptualization is, in turn, allocated to the formulation stage, resulting in greater accuracy.

Albert and Kormos also studied the effects of narrative task performance on the accuracy, fluency, and complexity of production. However, they are principally interested in the extent to which individual differences in creativity can influence task performance. Building on Carroll’s (1993) three-stratum theory of cognitive abilities and his observation that one common higher order ability factor is idea production and drawing also on Guilford’s (1967) distinction between divergent and convergent thinking, Albert and Kormos proposed that divergent thinking contributes to three facets of creativity that may affect idea production on L2 tasks: creative fluency, flexibility, and originality. They argued that these facets of creativity should be particularly influential on certain tasks, “especially open-ended ones like narrative tasks, for which there is no correct solution, but a large number of solutions are possible” (p. 82). In other words, they argued, creativity thus measured should contribute positively to performance on open versus closed tasks. They hypothesized that creativity (operationalized and measured using a test developed by Barkoczi & Zetenyi) should be particularly influential on the quantity of talk produced (as evident in the number of idea units) as well as lexical variety and the variety of narrative structure. They found that creative fluency and originality were related to the amount of talk produced, and the complexity of narrative structure attempted but had no effect on accuracy of production.

The articles by Tavakoli and Foster and by Albert and Kormos are complementary with respect to their focus on narrative task production. However, the attempt by Albert and Kormos to identify individual differences contributing to the difficulty learners experience in meeting narrative task demands is particularly important. Tasks, and characteristics of them that designers and researchers manipulate, are unlikely to have effects on production and learning independently of the abilities different learners bring to the context of task performance. All learning is the result of complex interactions between task demands contributing to their intrinsic cognitive complexity (so a less cognitively complex task should always meet with more success for any one learner than its more intrinsically complex counterpart) and the cognitive abilities and affective dispositions learners have, which affect their experience of how difficult tasks are and so contribute to variation in levels of success reached by any two learners on the same task (see Robinson, 2001a, 2001b, 2003a, 2007a; Shuell, 1980; Skehan, 1989; Snow, 1989, 1994; Sternberg, 2002). To what
extent, therefore, might Tavakoli and Foster’s findings for the effects of the design characteristics of narrative tasks on speech production been different if they had used measures of participants’ creative fluency and originality as covariates in their analyses—as Albert and Kormos demonstrate that both of these affect the nature of learners responses to narrative task demands?

The following two articles each explore the extent to which tasks can be used to support vocabulary or grammar instruction. In Long’s terms (1991; Doughty & Williams, 1998; Long & Robinson, 1998) the researchers address how tasks can be used to focus attention on forms selected and sequenced for instruction following a lexical or grammatical syllabus. In her article, Kim operationalizes Hulstijn and Laufer’s (2001) motivational-cognitive construct of task-induced involvement in order to examine whether certain tasks are more effective than others in promoting L2 vocabulary acquisition. After reporting previous findings concerning the effects of task characteristics on vocabulary acquisition, such as de la Fuente’s (2002) finding that tasks involving negotiation plus output led to greater receptive and productive learning than exposure to words provided in premodified input to tasks, Kim cited Hulstijn and Laufer’s (p. 542) claim that “the more effective task required a deeper level of processing of the new words than the other task.” Hulstijn and Laufer’s involvement load hypothesis is an attempt to operationalize differences in the depth of processing that tasks can encourage, whereby greater involvement in task demands causes greater depth of processing, which, in turns, leads to better retention of vocabulary than does lower involvement in task demands. The motivational need component of involvement load is driven by the desire to comply with task requirements, whereas search and evaluation are cognitive components that affect the extent to which attention is paid to form-meaning relationships when encountering vocabulary during the task. In two experiments, Kim found, first, that a writing task, with a higher involvement load index compared to tasks involving reading with comprehension activities or gap fill activities, led to more effective initial learning and better retention of new words. Second, she found that different tasks (writing a composition vs. writing sentences containing words) that had the same level of involvement load were effectively equivalent in promoting initial learning and retention of words.

In his article, Toth presented a rich, quantitative and qualitative assessment of the extent to which learner led discourse (LLD) and teacher led discourse (TLD) during tasks designed to maximize meaningful language use of, and attention to, the target of an (one semester) earlier instructional episode (in which learners were explicitly taught metalinguistic information about the Spanish anticausative clitic se) facilitated their subsequent learning of it. Toth’s study is
most clearly an example of task-supported approaches to instruction, following a predetermined (in this case, following course book-prescription) order for grammar teaching. Additionally, his study used closed, two-way information gap tasks to deliver the LLD treatment but open topic discussion for eliciting the effects of TLD on learning, in order to promote “scaffolded feedback given to facilitate form-meaning mapping” (p. 148). Although the tasks used to deliver LLD and TLD treatments were very different in their participation and participant structure this was the difference Toth sought to investigate. Toth found that TLD, and the participant/participation (see the Appendix) task conditions it was delivered through, was more effective at promoting development of the grammatically targeted, and previously practiced, forms than was LLD in two-way information exchange tasks.

The final article in this issue begins by summarizing the earliest theoretical rationales for the benefits of task-based interaction—and the opportunity for metalinguistic reflection on output it provides, that is, “the twofold potential of negotiation—to assist L2 comprehension and draw attention to L2 form—which affords it a more powerful role in L2 learning than has been claimed so far” (Pica, 1994, p. 508), and that “under certain task conditions, learners will not only reveal their hypotheses, but reflect on them, using language to do so” (Swain, 1995, p. 132). Gass, Mackey, and Ross-Feldman then reviewed a number of claims that have been made (e.g., by Foster, 1998, and Nunan, 1991) that experimental conditions operationalized in studying these issues may well not reflect the interactional patterns typical of classroom language learning contexts. They pointed out that although “there is not a prototypical classroom, any more than there is a prototypical laboratory” (p. 193), researched findings in the latter settings must demonstrate their generalizability to classrooms. Gass, Mackey, and Ross-Feldman went on to show that findings for interactional patterns in the experimental and classroom contexts they studied suggest few differences between them while revealing—in both settings—differences in interactional patterns that depended on the type of task the learners carried out. This issue of demonstrating the generalizability of findings from empirical research into the influence of task demands on learning to classrooms, in diverse institutional contexts and with diverse populations, is one that must continue to mediate between all of the theoretical claims and empirical findings described in this issue, and claims about their instructional relevance, as Bygate, Norris, and Van den Branden (2009) have pointed out:

Underpinning this empirical impetus is the imperative for research to focus on how TBLT works within the context of ongoing programs. Hence
the TBLT enterprise will not be able to rely on individual case studies of learners conducted outside the context of programs of instruction, or on laboratory studies, nor on studies carried out in host classrooms in which the use of tasks is investigated without relating their use to the teaching of the ongoing program. Such work provides a valuable contribution—in a sense it might be seen as a form of piloting—for the empirical grounding of TBLT. However, more widespread pedagogically contextualized research is clearly needed. (p. 497)

Revised version accepted 10 January 2011

References


Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91–132). Amsterdam: Benjamins.


language research in cross-cultural perspective (pp. 39–52). Amsterdam: Benjamins.


Appendix

The Triadic Componential Framework for Task Classification—Categories, Criteria, Analytic Procedures, and Design Characteristics

<table>
<thead>
<tr>
<th>Task complexity (cognitive factors)</th>
<th>Task condition (interactive factors)</th>
<th>Task difficulty (learner factors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Classification criteria: cognitive demands)</td>
<td>(Classification criteria: interactional demands)</td>
<td>(Classification criteria: ability requirements)</td>
</tr>
<tr>
<td>(Classification procedure: information-theoretic analyses)</td>
<td>(Classification procedure: behavior-descriptive analyses)</td>
<td>(Classification procedure: ability assessment analyses)</td>
</tr>
<tr>
<td>(a) Resource-directing variables making cognitive/conceptual demands</td>
<td>(a) Participation variables making interactional demands</td>
<td>(a) Ability variables and task-relevant resource differentials</td>
</tr>
<tr>
<td>± Here and now</td>
<td>± Open solution</td>
<td>h/l Working memory</td>
</tr>
<tr>
<td>± Few elements</td>
<td>± One-way flow</td>
<td>h/l Reasoning</td>
</tr>
<tr>
<td>± Spatial reasoning</td>
<td>± Convergent solution</td>
<td>h/l Task-switching</td>
</tr>
<tr>
<td>± Causal reasoning</td>
<td>± Few participants</td>
<td>h/l Aptitude</td>
</tr>
<tr>
<td>± Intentional reasoning</td>
<td>± Few contributions needed</td>
<td>h/l Field independence</td>
</tr>
<tr>
<td>± Perspective-taking</td>
<td>± Negotiation not needed</td>
<td>h/l Mind/intention-reading</td>
</tr>
<tr>
<td>(b) Resource-dispersing variables making performative/procedural demands</td>
<td>(b) Participant variables making interactant demands</td>
<td>(b) affective variables and task-relevant state-trait differentials</td>
</tr>
<tr>
<td>± Planning time</td>
<td>± Same proficiency</td>
<td>h/l Openness to experience</td>
</tr>
<tr>
<td>± Single task</td>
<td>± Same gender</td>
<td>h/l Control of emotion</td>
</tr>
<tr>
<td>± Task structure</td>
<td>± Familiar</td>
<td>h/l Task motivation</td>
</tr>
<tr>
<td>± Few steps</td>
<td>± Shared content knowledge</td>
<td>h/l Processing anxiety</td>
</tr>
<tr>
<td>± Independency of steps</td>
<td>± Equal status and role</td>
<td>h/l Willingness to communicate</td>
</tr>
<tr>
<td>± Prior knowledge</td>
<td>± Shared cultural knowledge</td>
<td>h/l Self-efficacy</td>
</tr>
</tbody>
</table>

Adapted from Robinson 2007b, by permission of Multilingual Matters. Reproduced with permission of the publisher, Multilingual Matters.