Task complexity, the Cognition Hypothesis and second language learning and performance

PETER ROBINSON AND ROGER GILABERT

Abstract

In this paper we describe a taxonomy of task demands which distinguishes between Task Complexity, Task Condition and Task Difficulty. We then describe three theoretical claims and predictions of the Cognition Hypothesis (Robinson 2001, 2003b, 2005a) concerning the effects of task complexity on: (a) language production; (b) interaction and uptake of information available in the input to tasks; and (c) individual differences-task interactions. Finally we summarize the findings of the empirical studies in this special issue which all address one or more of these predictions and point to some directions for continuing, future research into the effects of task complexity on learning and performance.

1. The Cognition Hypothesis: Task complexity, task design and task sequencing

This special issue consists of a theoretical overview of two models of attention that have prompted extensive research into the effects of task demands on selective attention and co-ordination of attentional resources during dual and multi-task performance, followed by four empirical studies examining the effects of manipulating dimensions of task complexity on; (i) the accuracy, fluency and complexity of second language (L2) speech production; (ii) the extent of interaction and uptake of premodified input occurring during task performance, and iii) learner perceptions of task difficulty. A great number of previous studies have examined the effects of one or another aspect of L2 task demands, such as the availability of planning time (see Ellis 2005), or the nature and extent of participation on tasks (see Pica, Kanagy and Falodun 1993), individually. Drawing on this, and other previous research the studies in this special issue all address the issue of task complexity in the Triadic Componential Framework (Robinson 2001, 2005a, 2007a) which specifies component dimensions of task complexity in terms of three superordinate *cognitive*, *interactive* and *learner* factors in order that each dimension can be studied separately, and also that

IRAL 45 (2007), 161–176
DOI 10.1515/IRAL.2007.007
0019042X/2007/045-0161
©Walter de Gruyter

complex interactions among these dimensions and factors can be studied for their multiplicative effects on L2 learning and performance (see Figure 1). The empirical studies in this special issue each motivate their research questions with respect to this framework, and their research hypotheses with regard to the theoretical claims of the Cognition Hypothesis (Robinson 2001, 2003b, 2005a, 2007a) which motivates the dimensional and factorial structure of the Triadic Componential Framework. The Cognition Hypothesis claims that increasing the cognitive demands of tasks contributing to their relative complexity along certain dimensions will; (a) push learners to greater accuracy and complexity of L2 production in order to meet the greater functional and conceptual communicative demands they place on the learner; (b) promote interaction, and heightened attention to and memory for input, so increasing learning from the input, and incorporation of forms made salient in the input; as well as (c) longer term retention of input; and that (d) performing simple to complex sequences will also lead to automaticity and efficient scheduling of the components of complex L2 task performance. These concomitant theoretically motivated claims of the Cognition Hypothesis support one fundamental pedagogic claim concerning second language instruction.

The fundamental pedagogic claim of the Cognition Hypothesis is that pedagogic tasks should be *designed*, and then *sequenced* for learners on the basis of increases in their cognitive complexity (Robinson 1996, 2001, 2003b, 2005a, 2007a, in press a; Robinson, Ting and Urwin 1995). The aim of such *pedagogic task* sequences is to gradually approximate, in classroom settings, the full complexity of real-world *target task* demands, so prompting learners to shift from program-entry levels of L2 use and task performance to program-exit targeted levels along a manageable, but continuously extending, developmental and performative route. These increasingly complex sequences of pedagogic tasks should constitute the task-based syllabus and provide the basis for program-exit task-based assessment (see Harper 1986; Prabhu 1987; Long and Norris 2000; Norris, Brown, Hudson and Yoshioka 1998; Schank, Berman and MacPherson 1999 for similar pedagogic rationales). Gradually approximating target-task demands on increasingly complex pedagogic task versions requires both (i) an operational *taxonomy* for *classifying* features of target tasks which can be simulated by task designers, and performed and practiced by L2 learners in pedagogic settings (at first separately, and then increasingly in combination) and also (ii) *principles* for *sequencing* these features, and combinations of them, in an order which approximates target-task demands.

A taxonomic system (see Robinson 2007a; Sokal 1974) for pedagogic task classification which can accommodate the fundamental pedagogic claim of the Cognition Hypothesis should therefore be both (i) descriptively *adequate*, "listing" in its taxonomy important features of target-task performance which are susceptible to pedagogic design and simulation in classroom settings, and also

(ii) *theoretically motivated*, capturing in its taxonomic "system-structure" categories of the design features of pedagogic tasks which can be simulated and sequenced to promote both *automatization* of, and fluent *access* to the current L2 interlanguage knowledge needed to accomplish task demands, and also categories of the design features of tasks that can be simulated and sequenced to promote further *analysis* and *development* of existing interlanguage knowledge in line with the target L2. Figure 1 describes a taxonomic system for task classification that aims to meet these *descriptive adequacy* and *theoretically motivated* charges.

1.1. *The Triadic Componential Framework for pedagogic task classification and task design*

The taxonomic "listing" of pedagogic task features in the Triadic Componential Framework, illustrated in Figure 1, falls under three broad classificatory categories: features of tasks contributing to their intrinsic cognitive *complexity*; features of tasks determined by the situational setting, and *conditions* in which they take place; and learner factors which contribute to the extent of the *difficulty* faced in attempting to successfully accomplish cognitively complex tasks, in a variety of situational settings. Task conditions concern the *participation* (e.g., 1-way versus 2-way) and *participant* (e.g., familiar versus unfamiliar) factors that can be identified in any target-task context of use. Task difficulty concerns the learners' perceptions of task demands, as these are variably affected by the *affective* and ability *factors* listed in Figure 1. Task difficulty will contribute to between learner variation in performing any one (simple or more complex) task, in the same way differences in aptitude for Math will distinguish the speed and success of those solving calculus or geometry problems. Task complexity (the proposed basis of sequencing tasks for learners) contributes to intra-learner variation in performing any two tasks, such as doing simple addition versus calculus, or the various simple and complex tasks studied in this volume.

In the Triadic Componential Framework an important theoretical distinction is made between resource-directing and resource dispersing dimensions of complexity (Robinson 2003a). The first subcategory distinguishes task characteristics on the basis of the concepts that the task requires to be expressed and understood (e.g., relative time, spatial location, causal relationships, and intentionality). Clearly, conceptualization is more, or less demanding of cognitive resources. This is evidenced, for example, by the staged emergence of conceptual abilities – and their linguistic expression – in childhood (see e.g., Bartsch and Wellman 1995; Cromer 1974; Mandler 2004), and by similar stages in the ability to mark and code them linguistically in the L2 during adult naturalistic second language acquisition (e.g., Becker and Carroll 1997; Dietrich, Klien

Task Complexity (Cognitive factors)	*Task Condition* (Interactive factors)	*Task Difficulty* (Learner factors
(Classification criteria: cognitive demands) (Classification procedure: information-theoretic analyses)	(Classification criteria: interactional demands) (Classification procedure: behavior-descriptive analyses)	(Classification criteria: ability requirements) (Classification procedure: ability assessment analyses)
(a) Resource-directing variables making cognitive/conceptual demands	*(a) Participation variables* making 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 making performative/procedural demands	*(b) Participant variables* making interactant demands	*(b) Affective variables* and task-relevant state-trait differentials
+/− planning time	+/− same proficiency	h/l opennes 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

Figure 1. *The Triadic Componential Framework for task classification – categories, criteria, analytic procedures, and design characteristics (from Robinson 2007a)*

and Noyau 1995; Perdue 1993). Expending the mental effort (see Wickens, this volume) needed to make more demanding *cognitive/conceptual* distinctions in language should therefore prime learners – and direct their attentional and memory resources – to aspects of the L2 system required to accurately understand and convey them, thereby facilitating selective attention to, and "noticing" of these (Robinson 1995, in press b; Schmidt 2001; Wickens this volume) and so speeding up L2 grammaticization in conceptual domains. Grammaticization of the L2 means to mark conceptual distinctions in language follows, in many cases, a similar trajectory in adulthood to the one apparent in child L1 development and recapitulating these sequences of conceptual/linguistic development by presenting L2 tasks in an order of increasing conceptual complexity has been proposed as a 'natural order' for scheduling such task demands with likely beneficial effects for learners (see Gilabert this volume; Robinson 2003b,

2005a, this volume; Robinson and Ellis in press a, in press b). Discussing the parallels between child and adult language development in the emergence of prepositions for marking first topological relations of neighborhood, and containment, and later, axis-based projective relations of between, front/backness, in the European Science Foundation (ESF) project data Slobin (1993) comments as follows:

> The parallels, though, cannot be attributed to the same underlying factors. In the case of FLA (first language acquisition) one appeals to cognitive development: the projective notions simply are not available to very young children. But in the case of ALA (adult language acquisition) all of the relevant cognitive machinery is in place. Why, then, should learners have difficulty in discovering the necessary prepositions for spatial relations that they already command in the L1? There are at least two possibilities: (1) adult learners retain a scale of conceptual complexity, based on their own cognitive development, and at first search the TL (target language) for the grammatical marking of those notions which represent some primordial core of basicness or simplicity; and/or (2) these most basic notions are also used with relatively greater frequency in the TL … but learners "gate them out" due to their complexity. In this case cognitive factors play a role in both FLA and ALA, but for different reasons: the complex notions are not available to very young children, while they are available but not accessed in early stages of ALA. (Slobin 1993: 243)

With both of the possibilities Slobin describes in mind, then sequencing tasks making conceptual/linguistic demands in the order these are met in child and adult L2 development would be complementary to adult learners' own initial dispositions to make form-meaning mappings, and also helpful in prompting them to move beyond them and grammaticize L2 speech in increasingly target-like ways. Proposed resource-directing variables (see Figure 1) distinguishing task characteristics in terms of these conceptual/linguistic demands include: (1) whether the task requires reference to events happening now, in a mutually shared context (Here-and-Now) versus to events that occurred in the past, elsewhere (There-and-Then); (2) reference to few, easily distinguished, versus many similar elements; (3) reference to spatial location, where easily identifiable and mutually known landmarks can be used, versus reference to location without this support; (4) simple information transmission, versus reasoning about causal events and relationships between them; (5) simple information transmission, versus reasoning about other people intentions, beliefs and desires and relationships between them; (6) and whether the task requires the speaker/listener to take just one first-person perspective on an event, or multiple second, and third person perspectives.

In contrast to resource-directing variables are those that make increased *performative/procedural* demands on participants' attentional and memory resources, but do not direct them to any aspect of the linguistic system which can

be of communicative value in performing a task. Meeting these demands during pedagogic task performance therefore should facilitate not *analysis*, and development of new L2 form-concept mappings but rather automatic *access* to, and control of, an already established interlanguage system (cf. Bialystok 1994). These resource-dispersing variables include those that distinguish task characteristics on the basis of: (1) giving planning time (and so increasing resource availability) versus not giving it; (2) providing background knowledge needed for task performance, versus not giving it; (3) tasks requiring only one thing to be done, versus those requiring two (dual) or many (multiple) things to be done simultaneously; (4) tasks where there is a clear structure available to help on deciding which steps are needed to complete it, versus those without one; (5) tasks where one or few steps are needed to complete it, versus those requiring many steps; (6) and tasks where there is no necessary sequence or 'chain' in which steps are followed, versus those which require participants to follow a strictly chained sequence, in which one step *must* be performed before another.

1.2. *Predictions for the effects of task complexity on learning and performance*

1.2.1. *Effects of task complexity on language production.* Most studies of the effects of task demands on speech production have employed general measures of accuracy fluency and complexity, such as percentage of error free C-units, or clauses per C-unit. All studies in this volume employ such *general* measures (see Ortega 2000; Wolfe-Quintero, Inagaki and Kim 1998 for review), enabling comparability of findings across a wide variety of task demands. Following arguments by Givon (1985, 1995; cf. Sato 1988, 1990) that structural complexity tends to accompany functional complexity in discourse, and that demanding, formal communicative tasks and contexts elicit a syntactic mode of production (characterized by greater use of morphology, greater syntactic subordination, and a higher noun to verb ratio) in contrast to a simpler pragmatic mode, the Cognition Hypothesis predicts greater accuracy and complexity using such general measures of production, on complex versus simpler tasks along all *resource-directing* dimensions of tasks. These general measures, however, will need to be supplemented by *specific* measures of the accuracy and complexity of production, as these are relevant to particular resource-directing measures making conceptual/linguistic demands. For example, tasks requiring complex spatial reasoning, event construal, and reference to motion, can be expected to lead learners to attempt to use developmentally later acquired lexicalization patterns for describing motion events (Berman and Slobin 1994; Cadierno 2004, in press; Robinson and Ellis in press a, in press b). Similarly (see the papers by Gilabert this volume, and Robinson this volume) tasks

requiring increasingly complex reasoning about, and reference to the intentional states of others causing them to perform actions can be expected to involve greater use of psychological and cognitive state terms such as 'think', 'expect', 'know', and the complex syntactic predication use of these terms requires (see Astington and Baird 2005; Lohman and Tomasello 2003).

In contrast to these predictions, along *resource-dispersing* dimensions of tasks which divide (see Wickens this volume), but do not direct attention to features of linguistic code, such as taking away planning time, or making dual or multiple simultaneous task demands, then accuracy and complexity of production can be expected to decrease on complex tasks. Skehan's (1998) Limited Capacity Hypothesis makes the same predictions for the effects of planning time, and other resource-dispersing dimensions, such as removing supporting task structure. Where the Cognition Hypothesis differs from the Limited Capacity Hypothesis is over the claims described above for the beneficial effects on accuracy and complexity of increasing the resource-directing dimensions of tasks. The resource-directing/dispersing distinction is one that Skehan does not make, leading him to claim complex task performance, along any dimensions, degrades accuracy, fluency and complexity simultaneously.

The Cognition Hypothesis, importantly, also claims that there are likely to be *synergetic* effects on speech production when tasks are made complex along both resource-directing and resource-dispersing dimensions simultaneously (as is often the case in real world task performance, such as impromptu reasoning about and explanations of the causes of a multi-party social conflict without the benefit of planning time). In such cases the beneficial effects on speech production of increasing complexity along the resource-directing dimension are likely to be weakened or negated by increasing complexity along the resource-dispersing dimension, when compared to the same task made simpler along a resource-dispersing dimension (e.g., where planning time is available).

1.2.2. *Effects of task complexity on interaction and learning opportunities.* The Cognition Hypothesis also *connects* input and interaction to the cognitive and conceptual *demands* of tasks that lead to differential amounts of interaction, or uptake of forms made salient in the input to tasks (see Mackey 1999). It predicts that along resource-directing dimensions, and in general too along resource-dispersing dimensions, that more interactive complex tasks will result in greater amounts of interaction, and negotiation for meaning. It also claims, following Long (1996), that such negotiation provides a context for attending to problematic forms in the input and output, and additionally that on complex versions of tasks there will be greater attention to, and uptake of forms made salient during provision of *reactive* Focus on Form techniques such as recasts (see Doughty 2001; Long 2007; Long and Robinson 1998). Alternatively, where *proactive* Focus on Form is provided, for example in the form of

premodified input to the task, then it similarly claims there will be greater use of this on complex, versus simpler task versions.

1.2.3. *Effects of task complexity on individual difference-task performance and learning interactions* Finally, the Cognition Hypothesis acknowledges that learner factors (contributing to perceived difficulty) interact with task factors (contributing to their complexity) in determining the extent of the above predicted effects. When the ability and affective factors drawn on in meeting complex task demands are high in any group of learners, then the effects will be found most clearly, in contrast to learners low in the ability and affective variables implicated in successful complex task performance. An example of this interaction of task difficulty and task complexity with language production was found by Robinson (this volume) where only those learners low in output anxiety responded to complex reasoning task demands by producing the predicted increasingly complex speech. Learners high in output anxiety were not induced by task demands to 'push' or 'stretch' production in this way. Individual differences in task-relevant abilities and affective factors can be expected to increasingly differentiate task-based learning and performance as tasks increase in complexity – an interaction effect well established outside the field of SLA (see e.g., Snow, Kyllonen and Marshalek 1984), and in much need of empirical, task-based SLA research (see Robinson 2002, 2005b, 2007b).

2. The papers in the special issue

The theoretical and data-based articles that follow all address issues raised above concerning the effects of task complexity on L2 language performance and development as well as theoretical and empirical issues regarding research into task design. These include the role of attention during task performance, the effects of increasing task demands on both oral and written production, the impact of task complexity on interaction and uptake, the mediation of individual differences such as anxiety or proficiency, the convenience of using general and/or specific measure to capture such effects, and the impact of increasing task complexity across different task types. Their findings and conclusions can be extended beyond research interests to inform decisions regarding task-based design and testing.

In his article, Wickens describes both a model of *selective* and focused attention and a model of *divided* attention. Although originally designed to predict how visual and sensory attention operates in human-system interaction (e.g., driving or flying), Wickens suggests how such models and the constructs they advance may be applied to the areas of L2 learning, competence, and performance. The two models have been devised to make predictions about: first, how perception mechanisms may allow the matching of sensory data with al-

ready existing knowledge; second, how information is selected and responses executed; third, how information is retained and rehearsed in working memory; fourth, how internally generated information is thought about in working memory; and fifth, how the passage of information from working memory to long term memory takes place. Such an information processing sequence, he suggests, is mediated by what he metaphorically describes as the *filter* and the *fuel* models. The first model of selective attention presented by Wickens is called The Salience, Effort, Expectancy, and Value (SEEV) model. Saliency has to do with how salient events are captured while non-salient events go unnoticed. As for effort, some tasks may require more of an effort than others, and prolonged, repeated, or demanding performance may lead to effort inhibition with negative effects on performance. With regard to expectancy, while in visual attention it is well-established that humans look at rapidly changing visual information, the role of expectancy is not so well defined when it comes to attention to language. Regardless of expectancy, we attach lower or higher value to certain sources of information either depending on the intrinsic properties of the information or on the task being performed. In combination with expectancy, research has shown that the maximum expected value, which appropriately balances the expectancy and value of a source of information for attention allocation, distinguishes good from not so good performers (i.e., pilots). These four dimensions of attention in Wickens' model compete with one another and the results of such a competition determine what gets to be selected from the sensory input and further processed. Attention allocation can therefore be predicted based on the "weighed sum" of the four components, which are finely operationalized and can therefore be precisely measured. Wickens' second model of multiple-resources of *divided* attention, has been devised to account for the different levels of interference of two tasks which are performed simultaneously. The level of interference will depend on the effort or resource demands of each task, the similarity of attentional resources (e.g., visual or verbal) they draw upon, and the allocation policy based on the strategic choice of the performer, who will at times 'protect' the performance of one task to the detriment of the other. Although the ability to allocate and switch attention is related to stable individual differences, especially interesting for the domain of language learning is that such a skill is also trainable. Regarding effort or resource demands, Wickens distinguishes between effort competition, effort demands, and effort investment. The first dimension of effort predicts how effort will be distributed along the two tasks being time shared. Effort demands have to do with the complexity (i.e., number of elements or interconnectedness of elements) of the two tasks at hand. The third dimension of effort is associated with how much effort is invested in performing each task, and it is independent from the two other dimensions of effort. Particularly interesting for language learning are effort demands and effort investment since task demands may re-

quire different levels of effort and task investment may for example determine whether learning takes place or not. As far as the multiplicity of resources is concerned, Wickens distinguishes between processing modalities (i.e., visual or auditory), processing codes (i.e., spatial or verbal), and processing stages (i.e., perception, cognition, and responding). This multiple-resource conception of attention helps explain why two simultaneous tasks can be performed poorly, when they draw on the same dimensions of attention, or effectively, when they draw on different dimensions. For example, driving and listening to the news are possible because they draw on different modalities (i.e., visual and auditory respectively), or driving and speaking, a perceptual and a response task which cause little interference. Finally, attention allocation policies are driven by what the performer believes to be important, either intuitively or because priorities are established in association with resource demands and multiple resource structure. And it is this 'priority' in the multiple-resource model that connects it to the SEEV model, because if SEEV predicts what task performers attend to, in combination with the multiple-resource model it also predicts which tasks are favoured and which are neglected. Wickens' attention models are relevant to cognitive approaches to research into tasks which are designed to encourage balanced performance in the areas of fluency, complexity, and accuracy, and to promote interlanguage development, since attention mediates the processes of comprehension, production, and learning of languages.

In their article, Michel, Kuiken and Vedder present a data-based study aimed at measuring the impact of increasing task complexity on oral production. They specifically set as their goal to test Robinson's claims in his Cognition Hypothesis regarding the effects of Task Complexity on both monologic and dialogic production. They do so by increasing task demands along the number of elements of the task and measuring its impact on learner's fluency, structural and lexical complexity, and accuracy. By means of a 2×2 design, 44 Moroccan and Turkish learners of Dutch are asked to perform a task in which they are either to recommend an MP3 or mobile phone to a friend or to decide in pairs on which device to acquire. The number of devices is two in the simple task and six in the complex version. By basing their predictions on the Cognition Hypothesis, they hypothesize that: (1) accuracy and linguistic complexity will be positively affected at the expense of fluency; (2) in the dialogic condition learners' production will be more accurate, but less fluent and linguistically complex; (3) that and interaction between task complexity and task condition will be found. As for measures, the authors use both general and specific measures to capture the effect of task complexity on accuracy. Two measures of structural complexity involving clause calculation and two of lexical complexity combining lexical and function words, and speech rate and filled pauses for fluency. MANOVA results show partial confirmation of their first hypothesis, with one of the measures of accuracy, one of lexical complexity, and one

of fluency being in line with their predictions. As for their second hypothesis, accuracy and complexity behave in the predicted direction while fluency runs against what they predicted. No robust interaction between task complexity and task condition is found, which leads them to challenge the claim advanced by the Cognition Hypothesis regarding the effects of task complexity on dialogic performance.

In their article, Kuiken and Vedder are concerned with the effects of increased task complexity on written production. They aim at testing the conflicting claims advanced by two models, the Limited Capacity Attentional Capacity model and the Cognition Hypothesis, regarding the role of attention during task performance and its effects on the three dimensions of production. As they argue, while the first model predicts a negative effect of increasing task complexity on all dimensions of production, the Cognition Hypothesis predicts some beneficial effects of increasing complexity along resource-directing variables simultaneously on accuracy and linguistic complexity. After analyzing in detail the differences between the two theoretical stances, Kuiken and Vedder revise the measures used to capture written performance in the areas of accuracy, structural complexity, and lexical variation. 84 Dutch learners of Italian and 75 Dutch learners of French with three and two different levels of proficiency respectively participated in their study. Two levels of task complexity were established. In a repeated-measures design, learners were instructed to make a recommendation out of five holiday destinations with three requirements in the simple task, while in the complex task the number of requirements amounted to six. The authors use specific measures of accuracy which include a classification of error types (e.g., grammar, lexical or orthographic) and lexical variation (e.g., lexical frequency profile analysis). Their study follows up a previous study by the same authors (Kuiken and Vedder 2007) and is exploratory in nature, and so no hypotheses are advanced. ANOVA results show that task complexity leads learners to make fewer errors and use more high frequency words under complex performance. In their discussion of results, they advocate the need for a classificatory system which may specify how different types of task performance may contribute to L2 learning. They also raise a series of theoretical questions regarding the two models of task complexity and attention, and suggest that their own results are better explained from a multiple-resource attentional perspective than from a single-resource, limited capacity perspective.

In his article, Robinson analyzes the impact of increasing task complexity on the three dimensions of production, on interaction and uptake, and as mediated by individual differences in anxiety. The resource-directing variable manipulated by Robinson is the intentional reasoning demands that the task impose on learners' processing during task performance. These have to do with reasoning and understanding people's motives, beliefs, and thoughts in carrying

out actions. Robinson reprises (see Robinson 2003b, 2005a) the claims of the Cognition Hypothesis (and the Triadic Componential Framework which articulates them) concerning the effects of task complexity on speech production, interaction and uptake, and the role of anxiety during complex performance. He hypothesizes that increases in intentional reasoning demands will trigger greater accuracy and complexity at the expense of fluency as measured by both general and specific measures, which in he claims will correlate to each other. Additionally, Robinson hypothesizes that more interaction and uptake will result from complex performance, that complex tasks will perceived as more difficult, stressful, and learners will show less confidence in performing them, and that input, processing, and output anxiety production will differentiate interaction and uptake more in complex tasks than in their simple counterparts. 42 Japanese students participated in the study. Speakers were instructed to organize a series of jumbled vignettes and describe the right sequence for their listeners to organize the same set of jumbled vignettes. Written input in the form of statements describing the event in the vignettes is carefully controlled. Three levels of task complexity were set as specified by the revised Wechsler Adult Intelligence Scale. As for coding of learner production, general measures of fluency, structural, and lexical complexity are used. These are complemented by more specific measures related to the use of psychological terms. Different measures of interaction and total and partial uptake are applied. Results of parametric and non-parametric tests show an impact of task complexity on accuracy and complexity of production. However, results do not confirm the predictions set by Robinson in his first hypothesis but are in line with the predictions in his second hypothesis, suggesting that specific measures which are more closely associated to the specific variable being manipulated are more sensitive than general measures. Also, general measures do not massively correlate with specific ones. Task complexity has the predicted impact on interaction and uptake, leaner perception, and is confirmed to be mediated by output anxiety. Robinson advocates the use of general measures across studies for a wide range of task demands but stresses the need to use specific measure that may capture the beneficial effect on L2 use and development. Robinson claims that syllabus and task design decisions can contribute to linking input and output and, as a consequence, create opportunities for learning. Finally, Robinson suggests the need to keep exploring the joint effects of learner factors which contribute to task difficulty with task complexity factors.

In his article, Gilabert measures the impact of increasing task complexity across three different task types on self-repair behaviour. Gilabert focuses on the prediction of the Cognition Hypothesis regarding accuracy and measure it by means of the calculation of self-repair behaviour. In a repeated-measures design, 42 Spanish learners of English participated in the study and were organized in two different groups according to their low or high proficiency. Learn-

ers performed three different task types (i.e., a narrative, an instruction-giving task, and a decision-making task) with two differences levels of complexity each. Learners were asked to narrative one story in the here-and-now and one in the there-and-then. As for the instruction-giving task, the number of elements is manipulated and increased in the complex version of a map task in which learners are also led to navigate through different axes (i.e., lateral, vertical, and sagittal). In the decision-making task, elements are kept stable from the simple to the complex task but the interconnectedness of the events is complexified. Gilabert hypothesizes that increased task complexity will result in a larger frequency and amount of self-repairs. Additionally, he predicts a larger impact of task complexity on the narrative and map task as opposed to the decision-making task. No differences between low and high proficiency groups are predicted. Nine different measures combining errors and self-repairs with AS-units and the number of tokens are used. Repeated measures ANOVA results show either a trend or a significant impact of task complexity on increased self-repair behaviour. This happens at different levels in the three task types. Learners also repair more frequently and in larger amounts when performing the narrative task as opposed to the two other tasks, only partially confirming what was hypothesized. As predicted, no differences are found between low and high proficiency groups. Gilabert explains his results in terms of the psycholinguistic mechanisms described by L1 and L2 models of language production and highlights the potential benefits of increased self-repair behaviour on learning and acquisition.

3. Conclusions and future directions for task complexity research

Research into the effects of task complexity, such as that presented in this special issue, aims both at pedagogic applications of findings regarding the effects of task design and sequencing decisions on learning and performance, and also at a deeper understanding of the second language processing and learning mechanisms that cause these effects. Not all the findings reported here support the predictions of the Cognition Hypothesis described in the first section of this paper, but many do. Over time, and accumulated studies, one might expect to see more consistent tendencies supporting, or not, these various predictions (Norris and Ortega 2006), as they are clear from groups of studies of a variety of learners at different levels of proficiency and with differing L1s, and differing in their affective and ability profiles (Robinson 2002). The focus, to date, has been on research into the effects of task complexity on speech production, but the other issue raised in the first section of this paper, i.e., of the effects of task complexity on interaction and uptake of input made salient during task performance, is also worthy of extensive research. And while at this stage of research, studies have largely reported results of single comparisons between

simple and more complex tasks, there is a clear need to look at longtitudinal effects of cycles of simple to complex task versions taking place over longer periods of time – as well as a need to examine to what extent performing cycles of simple to complex tasks, over periods of time, leads to successful real-world transfer of the performance and production abilities they help develop. Such is already the case in the many, and continuing studies of the optimal effects of sequencing training task demands in the applied psychology and ergonomics literature that Wickens (this volume) reports on. These, therefore, are issues of constructing – in ways the papers that follow demonstrate – an empirical, theoretically grounded research tradition which can provide an ever deepening and solid foundation from which to reason to preferred options in pedagogic task design and sequencing for learners across a range of instructional contexts.

Aoyama Gakuin University
<peterr@cl.aoyama.ac.jp>
Universitat Ramon Llull
<RogerGG@blanquerna.url.edu>

References

Astington, Janet and Jodie Baird (eds.) (2005). *Why Language Matters for Theory of Mind*. Oxford: Oxford University Press.

Bartsch, Karen and Henry Wellman (1995). *Children Talk About the Mind*. Oxford: Oxford University Press.

Becker, Angelika and Mary Carroll (1997). *The Acquisition of Spatial Relations in a Second Language*. Amsterdam/Philadelphia: John Benjamins.

Berman, Ruth and Dan Slobin (eds.) (1994). *Relating Events in Narrative: A Cross Linguistic Study.* Mahwah, NJ: Erlbaum.

Bialystok, Elaine (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition* 16: 157–168.

Cadierno, Teresa (2004). Expressing motion events in a second language: A cognitive typological perspective. In *Cognitive Linguistics, Second Language Acquisition and Foreign Language Teaching*, Michel Achard and Susanne Niemeier (eds.), 13–49. Berlin: Mouton de Gruyter.

Cromer, Richard (1974). The development of language and cognition: The cognition hypothesis. In *New Perspectives in Child Development,* Brian Foss (ed.), 184–252. Harmondsworth: Penguin.

Dietrich, Raier, Wolfgang Klein and Colette Noyau (1995). *The Acquisition of Temporality in a Second Language*. Amsterdam/Philadelphia: John Benjamins.

Doughty, Catherine (2001). Cognitive underpinnings of focus on form. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 206-257. Cambridge: Cambridge University Press.

Ellis, Rod (ed.) (2005). *Planning and Task Performance in a Second Language*. Amsterdam/Philadelphia, PA: John Benjamins.

Givon, Talmy (1985). Function, structure and language acquisition. In *The Crosslinguistic Study of Language Acquisition*, Vol. 1, Dan Slobin (ed.), 1008–1025. Hillsdale, NJ: Lawrence Erlbaum.

— (1995). *Functionalism and Grammar.* Amsterdam/Philadelphia, PA: John Benjamins.

Harper, David (ed.) (1986). *ESP for the University: ELT Documents 123*. Oxford: Pergamon Press.

Lohman, Heidemarie and Michael Tomasello (2003). The role of language in the development of false belief Understanding: A training study. *Child Development* 74: 1130–1144.

Long, Michael (1996). The role of the linguistic environment in second language acquisition. In *Handbook of Second Language Acquisition*, William Ritchie and Tej Bjatia (eds.), 413-468. San Diego, CA: Academic Press.

— (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum.

Long, Michael and John Norris (2000). Task-based language teaching and assessment. In *Encyclopedia of Language Teaching*, Michael Byram (ed.), 597–603. London: Routledge.

Long, Michael and Peter Robinson (1998). Focus on form: Theory, research and practice. In *Focus on Form in Classroom Second Language Acquisition*, Catherine Doughty and Jessica Williams (eds.), 1–32. New York: Cambridge University Press.

Mackey, Alison (1999). Input, interaction and second language development. *Studies in Second Language Acquisition* 21: 557–587.

Mandler, Jean (2004). *The Foundations of Mind: Origins of Conceptual Thought*. Oxford: Oxford University Press.

Norris, John, Jim Brown, Thom Hudson and Jim Yoshioka (1998). *Designing Second Language Performance Assessments.* Honolulu, HI: University of Hawai'i Press.

Norris, John and Lourdes Ortega (eds.) (2006). *Synthesizing Research on Language Learning and Teaching*. Amsterdam/Philadelphia, PA: John Benjamins.

Ortega, Lourdes (2000). Understanding syntactic complexity: The Measurement of change in the syntax of instructed Spanish L2 learners. Unpublished Ph.D. dissertation, University of Hawai'i, U.S.A.

Perdue, Clive (ed.) (1993). *Adult Language Acquisition: Crosslinguistic Perspectives*, Vol. 2: *The Results.* Cambridge: Cambridge University Press.

Pica, Teresa, Ruth Kanagy and Jean Falodun (1993). Choosing and using communication tasks for second language teaching and research. In *Tasks in Language Learning: Integrating Theory and Practice,* Graham Crookes and Susan Gass (eds.), 9–34. Clevedon: Multilingual Matters.

Prabhu, Nabindranath (1987). *Second Language Pedagogy.* Oxford: Oxford University Press.

Robinson, Peter (1995). Attention, memory and the 'noticing' hypothesis. *Language Learning* 45: 283–331.

— (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 287–318. Cambridge: Cambridge University Press.

— (2003a). Attention and memory during SLA. In *Handbook of Second Language Acquisition*, Catherine Doughty and Michael Long (eds.), 631–678. Oxford: Blackwell.

— (2003b). The Cognition Hypothesis of adult, task-based language learning. *Second Language Studies* 21: 45–107. www.hawaii.edu/sls/uhwpesl/21(2)/Robinson.pdf

— (2005a). Cognitive complexity and task sequencing: A review of studies in a Componential Framework for second language task design. *International Review of Applied Linguistics in Language Teaching* 43 (1): 1–33.

— (2005b). Aptitude and second language acquisition. *Annual Review of Applied Linguistics* 25: 45–73.

— (2007a). Criteria for grading and sequencing pedagogic tasks. In I*nvestigating Tasks in Formal Language Learning,* Maria del Pilar Garcia Mayo (ed.), 7–27. Clevedon:Multilingual Matters.

— (2007b). Aptitude, abilities, contexts, practice. In *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology,* Robert DeKeyser (ed.), 256–286. New York: Cambridge University Press.

— (in press a). Syllabus design. In *Handbook of Second and Foreign Language Teaching*. Michael Long and Catherine Doughty (eds.), Oxford: Blackwell.

— (in press b). Attention and awareness. In *Encylopedia of Language and Education, Vol. 6: Knowledge About Language*, Jasone Cenoz and Nancy Hornberger (eds.). New York/Berlin: Springer Scientific.

Robinson, Peter (ed.) (1996). *Task Complexity and Second Language Syllabus Design: Data-based Studies and Speculations.* University of Queensland Working Papers in Applied Linguistics, 1. Queensland: Center for Language Teaching and Research. http://www.cl.aoyama.ac.jp/~peterr/.

— (ed.) (2002). *Individual Differences and Instructed Language Learning.* Amsterdam/Philadelphia, PA: John Benjamins.

Robinson, Peter and Nick Ellis (in press a). Conclusions: Cognitive linguistics, second language acquisition and L2 instruction – issues for research. In *The Handbook of Cognitive Linguistics and Second Language Acquisition*, Peter Robinson and Nick Ellis (eds.). New York/London: Routledge.

Robinson, Peter and Nick Ellis (eds.) (in press b). *The Handbook of Cognitive Linguistics and Second Language Acquisition.* New York/London: Routledge.

Robinson, Peter, Chi-Chien Ting and Jianjun Urwin (1995). Investigating second language task complexity. *RELC Journal* 25: 62–79.

Sato, Charlene (1988). Origins of complex syntax in interlanguage development. *Studies in Second Language Acquisition* 10: 371–395.

— (1990). *The Syntax of Conversation in Interlanguage Development*. Tubingen: Gunter Narr.

Schank, Roger, Tamara Berman and Kimberli Macpherson (1999). Learning by doing. In *Instructional-design Theories and Models: A New Paradigm of Instructional Theory, Volume II,* Charles Reigeluth (ed.), 161–182. Mahwah, NJ: Erlbaum.

Schmidt, Richard (2001). Attention. In *Cognition and Second Language Instruction*, Peter Robinson (ed.), 3–30. Cambridge: Cambridge University Press.

Skehan, Peter (1998). *A Cognitive Approach to Language Learning.* Oxford: Oxford University Press.

Slobin, Dan (1993). Adult language acquisition: A view from child language study. In *Adult language Acquisition: Crosslinguistic Perspectives*, Vol. 2: *The Results*, Clive Perdue (ed.), 239–252. Cambridge: Cambridge University Press.

Snow, Richard, Pat Kyllonen and Brian Marshalek (1984). The topography of learning and ability correlations. In *Advances in the Psychology of Human Intelligence*, Vol. 2, Robert Sternberg (ed.), 47–104. Hillsdale, NJ: Erlbaum.

Sokal, Robert (1974). Classification: Purposes, principles, progress, prospects. *Science* 185: 1115–1123.

Wolfe-Quintero, Kate, Shunji Inagaki and Hae-Young Kim (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy and Complexity.* Honolulu, HI: University of Hawai'i Press.