CrossMark

Lang. Teach. (2017), **50.4**, 507–526 © Cambridge University Press 2017
doi:10.1017/S0261444817000179

# First Person Singular

## Position paper: Moving task-based language teaching forward

**Rod Ellis** Curtin University, Perth, Australia

*The advocacy of task-based language teaching (TBLT) has met with resistance. The critiques of TBLT and the misconceptions that underlie them have already been addressed in Ellis (2009) and Long (2016). The purpose of this article is to move forward by examining a number of real problems that TBLT faces – such as how a 'task' should be defined, the nature and timing of the 'focus-on-form' required, how to determine task complexity and sequence tasks, the role of explicit instruction, the timing of focus on form and the teacher training needed to ensure effective uptake of TBLT – and to also consider what solutions are possible. Disagreements exist regarding the relative merits of task-based and task-supported language teaching. I will propose that a hybrid/modular syllabus that allows for a balance between an object-oriented and a tool-oriented view of language teaching offers the most promising way forward.*

### Introduction

In some respects the advocacy of TBLT has the characteristics of a movement and, as is the case with such movements, it has aroused resistance from advocates of traditional, object-oriented teaching (Sheen 1994; Bruton 2002; Swan 2005). The critiques levelled at TBLT are based on 'misunderstandings' about TBLT and concern 'non-issues' and have already been adequately addressed (see Ellis 2009; Long 2016). There are, however, a number of issues in need of discussion not least because advocates of TBLT do not always agree on whether these are 'real issues', and, if they do agree, they differ in how they can be best addressed. In this article, then, I want to stake out my own position on these issues, pointing out where I differ from other advocates of TBLT (Michael Long, in particular) and suggest a way of addressing a fundamental disagreement about the merits of task-supported and task-based language teaching.

### Real issues

Table 1 lists the issues that I see as very 'real' for the design and implementation of TBLT programmes. In many cases, these issues concern different positions among advocates of TBLT, suggesting the need to debate them and, if possible, arrive at some consensus. Long

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

**Table 1** Some real issues in the design and implementation of TBLT courses

| | |
|---|---|
| 1. | What is a 'task'? |
| 2. | What types of tasks should figure in a given type of a task-based course? |
| 3. | What makes a task complex and how can tasks be sequenced effectively? |
| 4. | What is the role of explicit instruction? |
| 5. | What types of focus-on-form are compatible with task-based teaching? |
| 6. | What types of corrective feedback are compatible with task-based teaching? |
| 7. | Should feedback be immediate or delayed until a task has been completed? |
| 8. | What kinds of participatory structure – group/ pair work versus whole-class – are compatible with task-based teaching? |
| 9. | Are task-based abilities transferable? |
| 10. | How can teacher education programmes enable teachers to overcome the problems they face in task-based teaching? |

(2016), in an issue of the *Annual Review of Applied Linguistics* devoted to TBLT, produced his own list of 'real issues' (only #3, 9 and 10 in Table 1) based on his particular view of what constitutes TBLT. My list is longer, reflecting my view that alternative versions of TBLT are possible and indeed are necessary to accommodate different instructional contexts.

## Issue 1: What is a 'task'?

The key question here is whether 'task' should be defined narrowly in terms of real-world activities or more generically in terms of activities that are purely pedagogic in nature. Nunan (1989) first introduced the distinction between 'real-world tasks' and 'pedagogic tasks'. In a sense, any task carried out in a classroom ends up being 'pedagogic' but the distinction is useful because it encapsulates two kinds of authenticity – situational and interactional (Bachman & Palmer 1996). Real-world tasks aim at situational authenticity as they are based on the target tasks performed in the outside world. An example might be a task where two students take on the roles of hotel receptionist and prospective guest where one has to make a booking for a room based on information provided by the latter. A pedagogic task lacks situational authenticity but aims at interactional authenticity (i.e. the kind of natural language processing found in communication in the world outside the classroom). An example is Spot the Difference. It is very unlikely that two people would engage in talk aimed at identifying the differences in two pictures in real life but this task can result in patterns of turn taking and repairs of misunderstandings that are typical of everyday talk and thus achieve interactional authenticity.

Long (1985; 2015; 2016) has consistently argued that the tasks in a course should be needs-based. He proposes that the starting point for establishing the content of a task-based syllabus should be the identification of those target tasks that a specific group of learners need in order to 'function adequately in a particular target domain' (Long 1985: 91). There is an obvious advantage of such an approach as it helps to ensure content relevance. Long (2005)

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

claimed that 'every language course should be considered a course for specific purposes, varying only (and considerably, to be sure) in the precision with which learner needs can be specified' (p. 1).

There are, however, problems with such an approach. One problem is that learners have wants as well as needs and that good language teaching should go beyond addressing purely functional needs. Another problem concerns the procedures to be adopted in conducting a needs-analysis. Long proposed first grouping specific target tasks into task types as the basis for designing pedagogic tasks and outlined the procedures to be followed (see Long 2005; 2015) but this raises the thorny question as to how broad these task types should be. The main problem, however, is that it is difficult and perhaps impossible to identify target tasks for some groups of learners. In an instructional context where the target language plays no significant role in the wider society, identifying target tasks is clearly more problematic and can only be conducted on the basis of speculation about possible future uses of the target language. Cameron (2001), for example, argued that a needs-based syllabus is not feasible for young foreign language learners. Bachman (2002), in a discussion of task-based assessment, also questioned whether tasks should be defined as target tasks for all second language (L2) learners and argued that a needs-based approach is not always appropriate. He saw situational authenticity as not particularly relevant when the purpose is to assess the communicative abilities of general purpose learners. For such learners, he argued, the aim should just be to achieve interactional authenticity in the assessment tasks. Whether tasks are to be defined as real-life tasks that have both situational and interactional authenticity or as pedagogic tasks that have interactional authenticity but no situational authenticity must surely depend on the instructional context.

Finally, clarification is needed as to whether a 'task' should be defined solely in terms of its design features or also in terms of how it is to be implemented. Skehan (2016) makes a clear distinction between design and implementation options. Robinson (2011), however, conflates the two in his Cognition Hypothesis. For example, Robinson includes +/− planning time in his list of variables that determine task complexity whereas Skehan treats planning as an implementation variable. My proposal is that a clear distinction should be made between design and implementation variables and that 'task' is a workplan and so should be specified solely in terms of its design features. The reasons for this will become clearer when I discuss issue #3 (task complexity and sequencing).

## Issue 2: What types of tasks should figure in a given type of a task-based course?

Various typologies of task types have been proposed. These often consist of lists of the pedagogic tasks commonly found in task-based teaching (e.g. information/opinion gap, role-playing, personal, problem-solving, story-completion) – see, for example, Bruton (2002) and Willis (1996). What is lacking is a principled way of classifying such tasks. In Table 2, I distinguish two intersecting dimensions of tasks that can inform a more systematic way of classifying tasks and provide examples.

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

**Table 2** Types of tasks

| | Unfocused – tasks involving general samples of language. | Focused – tasks designed to elicit the processing of specific, pre-determined linguistic features. |
|---|---|---|
| Input-based – tasks that do not require but do not prohibit production. | Written instructions about how to make a model aeroplane. Learners are required to read the instructions and assemble the model. | Oral descriptions of the location of animals in a zoo. The instructions are designed so that learners have to distinguish between singular and plural nouns. Learners place pictures of the animals in the correct locations (Shintani 2016). |
| Output-based – tasks that require speaking and/or writing to achieve the outcome. | Learners act as judges to decide what punishment to give to a number of offenders when given information about the crimes they had committed (Foster & Skehan 1996). | Things-in-pocket task (Samuda 2001); learners are shown the contexts of a person's pocket and asked to speculate who the person might be (target = epistemic models). |

The distinction between input- and output-based tasks is an important one for a number of theoretical and practical reasons. Input-based tasks are needed for beginners, who lack the resources to produce in the L2; they provide learners with the comprehensible input that helps get them started in learning an L2. Input-based tasks are also useful for introducing task-based teaching to learners who are accustomed to a more traditional, structural approach and who may be resistant to a mode of teaching that requires them to treat language as a tool rather than as an object. Input-based tasks are easily conducted with the whole class making them well-suited to large classes and to teachers more used to teaching in lockstep. Output-based tasks are needed to develop higher levels of proficiency (Swain1985). They provide greater opportunities for the negotiation of meaning (Long 1996) and of form (Lyster 2001), which have been shown to promote L2 acquisition.

Advocates of TBLT acknowledge that both input-based and output-based tasks have a role to play but, in fact, TBLT is nearly always discussed and researched in terms of output-based tasks. Willis & Willis (2007), for example, give scant attention to input-based tasks. All the studies in Robinson's (2011) book investigating the claims of the Cognition Hypothesis involve speaking tasks. Many of the misconceptions about TBLT have arisen because it is seen as necessarily involving speaking tasks. Given the importance attached to input in all theories of L2 acquisition, there is an obvious need to give greater attention to the role of input-based tasks in TBLT and to some extent this is now happening (e.g. Shintani 2016).

The utility of focused – as opposed to unfocused tasks – is more controversial, as I have already noted. Both Long (2016) and Skehan (1998) see no need for them in TBLT. In Long's

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

case, this is because all tasks should be derived from target tasks, thus allowing no place for tasks designed to focus on specific linguistic features. In Skehan's case, tasks are seen as the means for ensuring balanced development in terms of complexity, accuracy, fluency. In Ellis (2003), however, I argued that focused tasks have a place in both a language programme and in research designed to inform pedagogy. It is well known that certain linguistic features (e.g. subject-verb agreement and complex structures such as hypothetical conditionals), constitute learning problems that continue even at advanced stages. However, I do not envisage a task-based programme composed entirely of focused tasks but rather their selective use when it becomes apparent that learners are attempting to use specific linguistic forms but cannot do so accurately. Focused tasks can be used to raise learners' awareness of the functional or semantic meanings of linguistic features as in Samuda's (2001) 'Things-in-pocket' task. They allow for intensive corrective feedback directed at a target structure and, as is now well established, such feedback is effective (Lyster, Saito & Sato 2013). In research, focused tasks make it possible to investigate whether the performance of tasks results in learning through pre- and post-testing, as for example in Mackey (1999) and the numerous corrective feedback studies. Measuring the learning that results from unfocused tasks is difficult and, in fact, has been rarely undertaken. A focused task, of course, must satisfy the same criteria as an unfocused task.

A particular type of focused task is a consciousness-raising task (Ellis 1991). This makes a linguistic feature (typically grammatical or pragmatic) the topic of the task and aims to help learners achieve a metalinguistic understanding of a rule or regularity. Long (2016) dismisses consciousness-raising tasks on the grounds that they are 'components in the delivery of a traditional linguistic syllabus' (p. 6). But this is to misunderstand the case I have made for them. Again, I do not envisage that such tasks should comprise a complete programme but see them as a means for developing explicit knowledge of specific features that are problematic to learners on the grounds that such knowledge might facilitate attention to these features in subsequent input and output. Furthermore, consciousness-raising tasks also provide opportunities for communicating if learners work together to 'solve' the linguistic problems they pose. It is for this reason that they can still be called 'tasks'. After all, one can make language the topic of talk just like any other topic. Consciousness-raising tasks, however, clearly do not have a place in Long's version of TBLT where all tasks must address functional needs although an exception might be in a task-based programme for linguists!

## Issue 3: What makes a task complex and how can tasks be sequenced effectively?

The issue of task complexity is important for syllabus design if one accepts the basic premise that there should be a progression from 'simple' to 'complex'. The problem is how to define 'task complexity'. Long (2016) saw this as a 'real issue'. He noted that while 'much good work has been published on task complexity … the overall yield has been disappointing' (p. 27). He sees the solution as more research to 'help make findings cumulative, encourage replication studies, increase productivity, and generally speed up progress on this issue'. To

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

my mind, however, little progress is likely until some fundamental issues have been addressed. These are considered below.

The problem is perhaps less 'real' where input is concerned. There are established ways of determining the complexity of input. In the case of written input there are standard measures of readability, for example, the Dale-Chall formula (1948). A limitation of these older measures is that they were based on vocabulary level and sentence length only. However, in more recent work (e.g. Pitler & Nenkova, 2008) measures that also take account of discourse factors such as cohesion and coherence have been developed. There is an enormous literature on readability and researchers interested in the design of task-based syllabuses would do well to refer to it. Readability measures may also be relevant for determining the listenability of oral texts as research has shown that the cognitive processes involved in reading and listening are not modality specific but rather draw on unimodal cortical regions of the brain (Jobard et al. 2006). There is also research on foreigner-talk that points to the kinds of modifications that speakers intuitively make to facilitate communication. Also relevant is research showing that modification consisting of elaboration that makes the comprehension of the input easier by restructuring the propositional content more clearly is more effective than simplification (Long & Ross 1993). Determining input complexity is relevant for all tasks that incorporate linguistic input in the task materials.

The main problem arises when it comes to determining the complexity of output tasks. The most developed framework for sequencing tasks from simple to complex – and the one that has attracted the most research – is Robinson's (2001; 2011) Cognition Hypothesis, in particular those factors in his Triadic Componential Framework relating to task complexity, which he claims provide 'a parsimonious way to sequence L2 tasks in a program of instruction' (2011: 10). Robinson is concerned with how a given task can be varied to approximate ever more closely to the demands of a target task. For example, he proposes a task sequence that manipulates two task complexity factors ('simple' and 'complex reasoning' and +/−'planning'). The same basic reporting task is performed under different conditions:

Pedagogic task 1: − reasoning/+ planning time
Pedagogic task 2: − reasoning/− planning time
Pedagogic task 3: + simple reasoning/− planning time
Pedagogic task 4: + complex reasoning/− planning time

Such a proposal provides a predictive model that can be tested empirically. But there are some obvious problems with it and with the research based on it.

The first is that it cannot be assumed that a task designated as 'complex' actually entails a greater cognitive load. Current research (e.g. Révész, Michel & Gilabert (2016)) is starting to address this by investigating how the specific design features of tasks affect the level of mental activity involved in performing the tasks. Sasayama (2016) used a dual task methodology, time-estimation and learners' self-ratings to distinguish the cognitive load imposed by four narrative tasks predicted to vary in complexity. Her study showed that these tasks did indeed differ in the cognitive load they imposed. However, clear differences were only evident

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

between the least and the most complex task, suggesting that fine gradations may not impact on the actual complexity of tasks.

The second – and to my mind the more serious problem – is that tasks are holistic and thus involve conglomerates of factors. It is difficult to see how a task can be designed based solely on two of the many factors[1] that comprise the resource-directing and resource-dispersing variables in Robinson's model. Sasayama's study found that there were task variables (such as code complexity) other than those built into the design of her tasks that affected the cognitive load of the tasks. The question is, then, how do we take account of the clusters of variables involved in any task in predicting the complexity of a given task? I can find no discussion of this is Robinson's or Sasayama's work.

The third problem is that the complexity of any single task will depend not just on the design of the task workplan but also on how the workplan is implemented. Robinson's task sequence does in fact take this into account as it builds in +/− planning time – clearly an implementation variable – into the task sequence. But there are numerous other implementation variables that can affect the difficulty of performing a task – for example, asking the learners to perform a similar task first (Prabhu 1987), asking learners to listen to a model performance of the task (Aston 1982), brainstorming ideas relevant to the topic of the task (Skehan 1996), and, +/− pressure to perform the task rapidly (Yuan & Ellis 2003). An inspection of Robinson's list of resource-directing and resource-dispersing variables reveals that apart from +/− planning there are no implementation variables. My point is this; if we want to investigate the effect of task complexity we must necessarily involve one or more implementation options. Sasayama's study, for example, incorporated two implementation variables – the learners did a 'practice' task before they started the sequence of four narrative tasks and they were given 3 seconds to look at the picture stories before they started to tell the stories. These implementation variables may have affected the cognitive load the learners experienced when performing the different tasks. Would the cognitive load of the different tasks have differed had they been given longer planning time, for example? In effect, the complexity of a task can never be considered separately from how the task is implemented. One might also ask what has the greatest impact on the complexity of a task and on cognitive load – the design features of the task or the conditions under which it is performed. Skehan (2016) is fairly clear on this. While he does not reject the impact that design features can have on performance, he concludes that 'they do not generate consistency or robust generalizations' (p. 37) and that 'putative manipulations of task complexity may not in reality produce different levels of complexity' (p. 40) and then goes on to argue that studies that have investigated the effect of task conditions – such as pre-task planning – have produced more consistent results.[2]

In articulating these problems, however, I do not want to suggest that we should abandon the search for a model of task complexity. Long is right as course designers and materials writers do need reliable criteria for determining task complexity. However, we do not currently have such criteria and we need not just a list of putative design variables but a theory of how the variables interact to determine complexity. Above all, we need to recognize that such a theory must also take account of the conditions under which a task is to be performed. Clearly, building such a theory is a major challenge.[3]

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

Finally, developing a task-based syllabus involves both the vertical and horizontal sequencing of tasks. By 'vertical sequencing' I mean the order in which specific tasks (or task types) will figure in the syllabus. For example, if the starting point is the target tasks for a specific group of learners (as Long proposes) then decisions need to be taken about how to order these functionally oriented tasks. By 'horizontal sequencing' I mean how different versions of the same task can be developed so that they lead incrementally to a simulation of the target task itself. Robinson's work on task complexity appears to be primarily concerned with horizontal sequencing but clearly we also need to consider vertical sequencing as well.

These are all very real issues. It is not easy to see how they will be resolved and difficult to see how research will resolve them, at least in the short term. In the meantime, perhaps, we need a practical way forward. In Ellis (2003; 2017) I suggested that designers of task-based courses will need to draw on both what research has shown about task complexity and on their own experience and intuitions about what constitutes the right type and level of task for a particular group of learners. From a practical standpoint intuition is needed.[4]

## Issue 4: What is the role of explicit instruction?

In 'pure' TBLT – the kind that Long (2015) promotes – there is no room for explicit instruction preceding the performance of a task. Long argues that the default processes involved in L2 acquisition are those relating to incidental and implicit learning. He acknowledges, however, that these have 'reduced power' (p. 40) in adults and that it is therefore necessary to facilitate intentional initial perception of new forms in the input. An obvious way to achieve this is through explicit instruction. However, Long rejects this on the grounds that it constitutes a return to 'focus on forms', which he sees as incompatible with research that makes it clear that 'learners, not teachers, have most control over their development' (p. 24). In Long's view explicit instruction followed by practice can only result in automatized declarative knowledge, not true implicit knowledge. Thus, Long argues for 'focus on form', where various pedagogic procedures are utilized to attract learners' attention to linguistic features WHILE they are performing a task. In contrast, in task-supported language teaching (TSLT), specific linguistic forms are first taught explicitly and then practised under 'real operating conditions' using tasks (DeKeyser 1998). TSLT is supported by skill-learning theory, which claims that practice enables declarative knowledge to be proceduralized and automatized.

In pitting 'focus on form' against 'focus on forms' some obvious questions need to be asked. One such question is whether focus on forms does or does not result in true implicit knowledge. This question is notoriously difficult to address given the difficulty in devising tests that afford separate measures of explicit and implicit knowledge. Some progress has been made (see, for example, Ellis 2005) although problems still remain (see Suzuki & DeKeyser 2015). What is clear, however, is that TSLT can, at least, result in the ability to deploy the target of instruction in unplanned language use. Long does not deny this but argues that it only occurs when massive practice of the target feature such as that found in Day & Shapson (1991) and Harley (1989) is provided, which he suggests is not feasible if teachers

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

need to cover all the linguistic features listed in the syllabus. This is a legitimate point but it is only a reason for rejecting TSLT as the basis for a complete course, not for the wholesale rejection of TSLT. TSLT might still have a place in helping learners acquire those linguistic features that are not learned 'naturally'. In fact, it was precisely this aim that motivated the studies by Day & Shapson and Harley, who investigated the effects of direct instruction on grammatical structures that previous research had shown immersion learners typically fail to acquire.

If we accept that both focus on form and focus on forms can result in the kind of L2 knowledge (implicit or automatized explicit knowledge) that enables learners to communicate fluently and accurately – and there is plenty of evidence to support this – then we still need to ask whether, in the long term, one approach is more effective or efficient than the other. Unfortunately, most of the studies that have sought to compare TBLT and TSLT (e.g. De la Fuente 2006; Laufer 2006; Sheen 2006) have been short term and/or methodologically flawed. This question cannot be answered until we have well-designed, longitudinal comparative studies.

There is, however, another question that needs to be considered. Does prior explicit instruction affect how a task is performed – for example, by causing learners to focus on practising the target structure rather than trying to communicate? In other words, if learners are focused on trying to produce the target structure correctly, does the task still afford 'real operating conditions'? Again, there has been surprisingly little research that has addressed this question. Two studies that have examined the effect explicit instruction has on the performance of a task have produced different results. Mochizuki & Ortega (2008) reported that a group that received a hand-out providing explicit information about English relative clauses produced more than twice as many relative clauses when they performed the focused task than a group that did not receive the hand-out but that there was no difference in the global complexity and fluency of the production of the two groups. In the second study, Li & Ellis (forthcoming) also found that learners who had received explicit instruction in the target structure (past passive) were more likely to try to use it when they performed the tasks than learners who did not receive this instruction. In this study, however, the explicit instruction did have an effect on global complexity, accuracy and fluency with the explicit instruction group producing language that was less complex and fluent and that also tended to be less accurate overall. Whether and to what extent explicit instruction 'interferes' in the way a task is performed is an important issues for TSLT. If it does interfere, then clearly tasks in TSLT are not functioning as they do in TBLT. Even if they result in better learning of the target structure, they will not provide the same opportunities for natural language use and thus, in the long term, may be less successful in developing all round proficiency in the L2. Clearly, though, this is an issue in need of further investigation.

## Issue 5: What types of focus-on-form are compatible with task-based teaching?

Ellis, Basturkmen & Loewen (2001) distinguished pre-emptive and reactive form-focused episodes (FFEs) that occurred in task-based lessons. In the former either a student or the teacher initiated a focus on a specific linguistic feature. In the latter a classroom participant

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

(normally the teacher) drew attention to a particular linguistic form that was the source of a problem either because the meaning of an utterance was not clear or because an utterance contained a linguistic error. In the 448 FFEs Ellis et al. identified in 12 hours of teaching, 165 were student-initiated pre-emptive FFEs, 41 pre-emptive teacher-initiated FFEs and 223 reactive FFEs. This study shows that both pre-emptive and reactive FFEs can occur frequently in task-based teaching.

Long (2015), however, insists that focus-on-form should be reactive. He defines it as involving 'reactive use of a wide variety of pedagogic procedures to draw learners' attention to linguistic problems in context, as they arise in communication'. Thus, for Long focus-on-form is necessarily 'a *response* to a difficulty' (p. 317; italics in original). He recognizes that there will be times when it is necessary for teachers to draw learners' attention to form explicitly but this too should occur as a response to what the learner has said or written. Long does not discuss pre-emptive focus on form but it would seem that he has little time for it.

Research has concerned itself almost entirely with reactive focus on form (see, for example, the corrective feedback studies). There has been almost no research that has examined pre-emptive focus on form. Ellis, Basturkmen & Loewen (2001) investigated the likelihood of successful uptake (defined as the correct repair of a linguistic feature or demonstrated understanding of an item) occurring in pre-emptive and reactive FFEs. They reported that successful uptake was more likely to occur in learner-initiated pre-emptive and in reactive FFEs than in teacher-initiated pre-emptive FFEs. Successful uptake, of course, cannot be equated with acquisition but in another study Loewen (2005) reported that successful uptake was the most significant variable predicting learning as measured in tailor-made tests of those linguistic forms that were addressed during the FFEs. Taken together, these two studies suggest that, as Long claims, reactive focus on form facilitates learning, but also that pre-emptive focus on form, especially when initiated by students, is beneficial.

The issue here is how narrowly we are to define focus on form – as purely reactive or as also pre-emptive. Given that students (adult ones at least) tend to naturally topicalize problematic linguistic points during the performance of a task, it would seem unwise to dismiss pre-emptive focus on form.

## Issue 6: What types of corrective feedback are compatible with task-based teaching?

Reactive focus on form mainly consists of corrective feedback. The six types of corrective feedback that Lyster & Ranta (1997) identified in their study of immersion classrooms has informed the considerable body of experimental research that followed. This research was directed at investigating the effect that different types of corrective feedback have on acquisition. Some researchers (e.g. Lyster 2004; Ammar & Spada 2006) investigated the relative effects of input-providing feedback (e.g. recasts) versus output-prompting feedback (e.g. clarification requests, metalinguistic feedback and elicitation). Other researchers (e.g. Ellis, Loewen & Erlam 2006) have investigated implicit feedback (i.e. recasts, clarification requests, and repetitions) versus explicit feedback (i.e. explicit correction and metalinguistic feedback). These two dimensions of feedback intersect so each feedback strategy can be classified in terms of whether it is input-providing or output-prompting and implicit or

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

explicit. For example, recasts (at least those that are 'conversational' in nature) can be classified as input-providing and implicit and metalinguistic feedback as output-prompting and explicit.

This typology of corrective feedback strategies has provided a basis for making and testing theory-based claims about the efficacy of different types of feedback. Lyster (2004) drew on skill-learning theory to argue that output-prompting feedback is more likely to be beneficial than input-providing feedback (recasts) on the grounds that learners more frequently repair their errors following the former. Long (2006), however, claimed that recasts are more likely to promote acquisition as they provide learners with positive input for 'new' features and disputed whether the uptake of feedback has any utility. Theoretical arguments can also be made for implicit and explicit feedback; implicit feedback caters to implicit learning whereas explicit feedback is more likely to result in the conscious attention to form that Schmidt (2001) saw as important for acquisition and explicit learning. From a theoretical perspective, there is an obvious need to investigate the effects of these different corrective strategies. Goo & Mackey (2013), for example, argue for carefully designed studies that investigate the effects of narrowly prescribed types of corrective feedback.

However, there are several reasons for doubting the practical relevance of the research. First, the relative effects of different types of feedback are contextually constrained; thus one type may be more effective in one context and another in a different context. Second, teachers are likely to quite naturally employ a variety of strategies rather than stick to a single type – although research does suggest that recasts tend to predominate in the meaning-focused interactions that tasks elicit. Third, it is possible to combine different strategies when addressing a particular error; in 'corrective recasts' (Doughty & Varela 1998) where errors are first corrected by means of a repetition (i.e. output-prompting) and then by means of a recast (i.e. input-providing) if the repetition did not result in successful uptake. Combining strategies in this way is likely to make the feedback explicit and therefore more salient to learners. Fourth, and perhaps most important, the research has typically investigated the effects of intensive feedback directed at a single linguistic feature (e.g. regular past tense) elicited by means of focused tasks. For obvious reasons, there has been very little experimental research that has investigated extensive feedback directed at a whole range of linguistic problems as these arise during the performance of unfocused tasks, which some advocates of TBLT (Long; Skehan) have argued should constitute the content of a task-based course. A more pedagogically relevant approach to investigating the effects of corrective feedback would involve longitudinal studies where the feedback is extensive rather than intensive but such research is currently lacking.

The theoretical controversies concerning the relative merits of input-providing versus output-prompting and implicit versus explicit feedback appear to be abating. Long's early commitment to recasts as an implicit form of feedback (see Long 1996; 2006) is less evident in his later publications (Long 2015) in part because he now sees feedback as catering to intentional learning as well as to incidental learning. He currently considers the provision of explicit grammar rules as legitimate so long as this occurs in response to problems that arise during communication. In other words, he advocates a variety of feedback strategies – a view echoed by Ellis (2009) and Lyster & Ranta (2013). Teachers do need to know, however,

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

the potential benefits of the different strategies and to guard against over-utilizing a single type.

## Issue 7: Should feedback be immediate or delayed until after the task has been completed?

A common misconception about TBLT is that 'acquisition only takes place online during communication' (Swan 2005: 379). Long (2016) pointed out that that this is not the case but he also questioned whether metalinguistic knowledge acquired at one time can be used by the learner at another time – as Swan claimed can happen. What Long emphasized is that focus on form must be reactive. At issue, however, is whether reactive focus on form needs to occur during the performance of a task or can delayed until the task has been completed. There is little discussion of this in TBLT circles – but see Samuda (2001). It would seem that immediate, online focus on form is preferred given the psycholinguistic importance attached to learning-through-interaction and the fact that the TBLT research has investigated this. It is worth noting, though, that delayed feedback is still 'reactive' as long as it addresses linguistic problems that arose during the performance of the task. Also, a theoretical case can be made for delaying feedback. Quinn (2014) drew on Preparatory Attention and Memory Theory, which proposes that learning is enhanced when complete attention is available rather than divided between multiple goals. This theory also claims that when learners perform a difficult task they will realize they lack crucial information and thus are more likely to pay attention to this information when it is subsequently made available to them.

The timing of feedback is an important issue because a common position in the pedagogical advice given to teachers is that feedback should be immediate in accuracy work but delayed in fluency work (Hedge 2000; Scrivener 2005). In TBLT, however, the distinction between accuracy and fluency work does not apply as the goal is to develop accuracy and fluency contiguously. Thus, there is an apparent conflict in the position taken by authors of teacher guides and in the position found in theoretical accounts of TBLT.

There is, though, very little research that has investigated delayed feedback in TBLT. McDonough & Mackey (2006) reported a study of learners' responses to recasts. They found that what they called 'primed repetition' of the recasts (i.e. repetition that did not take place immediately following the recasts but within six turns) was associated with L2 development. Immediate repetition was not. However, it could be argued that the 'primed repetition' in this study was closer to an immediate than a delayed response. In Quinn (2014) 90 intermediate-level adult English as a second language learners were randomly assigned to immediate-, delayed- and no-feedback conditions. Feedback took the form of prompts that pushed the learners to self-correct followed, if necessary, by a recast (i.e. corrective recasts). In this study the delayed condition involved feedback provided after the tasks had been completed. There were statistically significant improvements in the accurate use of the target structure (English passive constructions) in both the immediate- and delayed-feedback groups but no difference between them. However, in this study the no-feedback condition also resulted in similar levels of improvement suggesting that gains resulted from simply performing the task. This study was also laboratory-based. Li, Zhu & Ellis (2016) reported a classroom-based study. The target structure was also English passive and the feedback type was again corrective recasts. Both the

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

immediate- and delayed-feedback groups showed gains in accuracy whereas the no-feedback group did not. However, the gains were only evident in a grammaticality judgement test, not in an elicited oral imitation test, suggesting that the feedback only resulted in explicit knowledge. In this respect, immediate feedback showed some advantage over delayed feedback for those learners that had some prior knowledge of the target structure.

There is insufficient research to adjudicate on this issue. It is clear, however, that the pedagogic advice that teachers should desist from correcting learners in fluency work has no warrant. In fact, teachers do focus on form during communicative tasks even when they believe they should not be doing so (Basturkmen, Loewen & Ellis 2004).

## Issue 8: What kinds of participatory structure – group/pair work, whole-class – are compatible with task-based teaching?

A common misconception about TBLT is that it must involve small group work. This misconception may again have arisen because so much of the research has investigated how learners perform tasks in pairs or groups. But tasks can be performed in a variety of participatory structures, including teacher–class – as is necessary with input-based tasks – and individually when learners work by themselves – as in Prabhu's (1987) Communicational Teaching Project. Teachers have a choice of participatory structures in TBLT. Even speaking tasks can be conducted through teacher–class interaction, as illustrated in Ellis, Basturkmen & Loewen (2001) and Loewen (2005).

What is lacking is any discussion about what constitutes an appropriate participatory structure for different groups of students. The advantages of small group work are well documented – see, for example, Long & Porter (1985). But there are also potential disadvantages. As I mentioned above, it is not reasonable to expect beginner-level learners to perform speaking tasks in groups as they lack the necessary proficiency and thus will resort to their first language (L1) or else produce the kind of restricted, pidgin-like output that Seedhouse (1997) was critical of. I have also argued that using group work may not be the best way to introduce students to TBLT if their previous experience of classroom instruction consisted only of traditional, lockstep teaching. There is also the problem that focus on form may be limited in group work. Adams (2007) reviewed a number of studies that have examined whether and to what extent learners correct each other when performing tasks in groups, concluding that 'while learner–learner interactions may provide a site for feedback to occur, the restricted set of feedback types may not provide evidence appropriate to learner developmental needs' (p. 33). Of course, much depends on the nature of the tasks the learners perform, with some tasks more likely to lead to the negotiation of meaning and a focus on form than others. Adams' own study found that learners did provide each other with feedback when performing opinion-gap and information-gap tasks and that it resulted in both grammar and vocabulary learning. Swain's research (e.g. Swain 1998) also provides convincing evidence that language-related episodes arise as learners perform tasks and that learning results especially if the problems that these episodes address are resolved successfully. However, in general, studies such as these were carried out in a laboratory setting and with intermediate or advanced level learners. More evidence is needed to show how far focus on

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

form is common and facilitative of learning in different participatory structures and with beginner-level learners.

## Issue 9: Are task-based abilities transferable?

This is one of the 'real issues' that Long (2016) addressed. There are, in fact, two related issues here. One is whether the performance of one task (or one of a series of tasks) results in improved performance on subsequent new tasks? This is a fundamental issue for TBLT. There is a tension in TBLT circles between the emphasis on how tasks can promote effective learning and the actual demonstration of this in terms of the transferability of whatever is learned to a new task. The second issue concerns assessment. How do we determine whether performing tasks results in learning?

Experimental studies have frequently used tasks as pre- and post-tests to measure the learning of specific linguistic forms that results from performing the treatment tasks. In cases where the performance of the treatment task involves some kind of intervention (for example, corrective feedback) there is clear evidence of learning. However, in cases where there is no such intervention, transferability to a new task may not be found (see Ellis 2009). The clearest evidence of this can be found in task-repetition studies. In these studies, learners are asked to repeat the same task without any intervention between performances and then to perform a new task (typically of the same type). Complexity, accuracy and fluency improve when the task is repeated, but this improvement does not seem to transfer to the performance of a new task. It is possible, however, that while transferability is not evident in these global measures it may become evident in the use of specific linguistic features (see, for example, Shintani & Ellis 2014). Also, task-repetition studies that included some kind of intervention between performances of the task (e.g. explicit focus on specific linguistic forms) have shown that transfer of learning to a new task can occur (e.g. Baleghizadeh & Derakhshesh 2012; Hawkes 2012). Much of the TBLT research has been premised on the assumption that by manipulating task design features and implementation options (such as pre-task planning) it is possible to prioritize the attention that learners pay to complexity, accuracy and fluency, and over time this will help them to achieve a balanced proficiency. But research that focuses on the one-off performance of a task cannot show that development is taking place. Again, to establish transferability longitudinal studies are needed.

There is plenty of good work on assessing the outcomes of task-based instruction (see, for example, the task-based assessment in the special issue of *Language Testing* in 2002). However, from my own experience teachers often say they can't assess students in TBLT when there are no explicit linguistic targets. Norris (2016), in a review of work on task-based language assessment (TBLA), commented that 'designing, implementing and evaluating TBLAs are by no means simple tasks' and went on to list the issues involved:

> . . . the extent to which tasks and associated language use contexts can be adequately simulated within controlled testing conditions, the authenticity of criteria for evaluating task performances, the reliability of raters in judging task performance, the comparability of tasks in terms of construct representation as

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

> well as difficulty, the generalizability and extrapolation limitations as a basis for assessment design, and the practicality and expense involved in creating and using task-based tests. (pp. 239–240)

These problems are more acute with speaking assessment tasks. Input-based tasks can be assessed more easily in terms of whether learners achieve the outcome of the task (e.g. can correctly draw in the route described on a map). Whatever the difficulties, however, assessment of learning must be task-based. TBLT is directed at integrative, holistic learning and cannot be assessed through discrete-point tests, which if used would also have a negative wash-back effect on teaching.

## Issue 10: How can teacher education programmes enable teachers to overcome the problems they face in task-based teaching?

This is another real issue that Long acknowledges. In many respects, it is the most serious problem. TBLT, with its emphasis on holistic teaching and learning-through-doing constitutes a major innovation for many teachers. Prabhu (1987) recognized this:

> A new perception of pedagogy, implying a different pattern of classroom activity, is an intruder into teachers' mental frames – an unsettling one, because there is a conflict of mismatch between the old and new perceptions and, more seriously a threat to prevailing routines and to the sense of security dependent on them. (p. 105)

In Ellis (1997), I listed the key characteristics of successful innovation. These included the teacher's dissatisfaction with his/her current practice, the feasibility of the new approach in the particular context in which the teacher works, the extent to which the new approach is seen as matching the needs of the students, how easy the principles and procedures of the new approach are to grasp, the extent to which teachers are required to demonstrate a high degree of originality in order to implement the innovation, and, perhaps, crucially, the extent to which teachers feel that they 'own' the innovation. It is easy to see how TBLT might run into problems and, indeed, evaluations of TBLT show that in some contexts it was not successful. Teachers may lack confidence in their L2 proficiency and thus feel that they cannot use tasks (Butler 2011). Students may be unconvinced that the incidental learning that TBLT caters to is the most efficient way of learning an L2. A particular problem noted by Carless (2004) is that teachers do not always have a clear idea of what a 'task' is. He found that the tasks used by primary school teachers in Hong Kong often ended up as 'language practice' rather than affording opportunities for genuine communication. Erlam (2016) found that the tasks she asked teachers to design in an in-service programme did not fully match up to a definition of a 'task'.[5] In particular, the teachers tended to design tasks that involved the prior presentation of language reflecting their pre-existing ideas of what a task should do.

These are real problems. They can only be addressed through carefully designed initial and in-service teacher training/education programmes that take account of the characteristics of successful innovations. An excellent model of how TBLT can be successfully introduced can be found in Van den Branden's (2006) account of its

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

implementation in Flanders (Belgium). This involved the development of task-based syllabuses, extensive teacher training, and ongoing research into the implementation of task-based programmes. However, to date, there has been very little research investigating how effective training/education programmes are in enabling teachers to implement TBLT.

## Conclusion

Misconceptions about TBLT abound. In part these are the product of the kind of research that has investigated tasks – often laboratory-based and often focusing on the performance of single tasks. These misconceptions are likely to prevail until there are more accounts of fully task-based courses. Doubts will also continue to exist until it can be shown that TBLT is effective – not just in developing communicative ability but also in achieving linguistic accuracy. Critics of TBLT are unlikely to be convinced until there are more studies like Shintani (2016), who produced evidence that TBLT can be more effective than traditional, structural teaching.

But while it is clearly necessary to address the misconceptions, it is equally important to acknowledge and try to address the real problems of designing and implementing task-based courses. This has been the main focus of this article. I have discussed ten problems and suggested ways in which they might be addressed.

I do not see TBLT as constituting a single, monolithic approach and in this I differ from Long – and this affects how I see the kinds of developments that are needed to move TBLT forward. Whereas Long's version of TBLT emphasizes the importance of needs-analysis to identify the target tasks relevant to a specific group of learners, I have argued such an approach is not feasible in contexts such as the state educational systems in East Asia that I am familiar with. Neither do I see TBLT and TSLT involving explicit instruction and focused tasks as incompatible but rather as mutually supporting. There is plenty of evidence to show that explicit instruction involving focused tasks is effective. Not all linguistic features can be acquired incidentally and implicitly through task-based teaching. Redundant, non-salient and complex features, especially when these are 'blocked' by the learner's L1, are unlikely to be acquired even with the help of focus on form. For such features explicit instruction is necessary. It is for this reason that I have argued for a hybrid syllabus consisting primarily of a task-based component but supported by a task-supported component to address recalcitrant linguistic problems when these become evident. Task-based and task-supported teaching draw on different psycholinguistic rationales – as Long (2015) pointed out – but it does not follow that we should opt for just one of these. Successful L2 acquisition clearly does call for learning incidentally through task-based teaching but it can also benefit from the skill development that task-supported teaching can provide. To my mind, then, the final 'real issue' that has to be addressed is how to design courses that are task-based but that also build in, where needed, opportunities for working on specific linguistic problems through task-supported teaching.

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

## Endnotes

1 Robinson (2001; 2011) provides a taxonomy of resource-directing and resource-dispersing factors. Examples of the former are +/− few elements, +/− here and now, +/− reasoning demands. Examples of the latter are +/− planning, +/− single task and +/− prior knowledge. His taxonomy does not exhaust all the possible features of a task.
2 A reviewer of this article also noted why the design of a task cannot be expected to account for complexity – learners can use strategies to handle problems that they experience with tasks, for example, by breaking down propositions into sub-propositions or sharing out parts of the task dialogically.
3 Skehan (2016) suggests that the challenge of developing a theory that can inform how design and implementation variables impact on task performance might best be met by drawing on Levelt's (1989) model of speaking.
4 It is also possible that learners' engagement with a task will be more influenced by whether they find the tasks relevant or interesting than by task complexity.
5 The tasks used in research on TBLT also do not always satisfy the definition of a task.

## References

Adams, R. (2007). Do second language learners benefit from interacting with each other? In A. Mackey (ed.), *Conversational interaction in second language learning*. Oxford: Oxford University Press, 29–51.
Ammar, A. & N. Spada (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition* 28, 543–574.
Aston, G. (1982). Trouble-shooting in interaction with learners: The more the merrier. *Applied Linguistics* 7, 128–143.
Bachman, L. (2002). Some reflections on task-based language performance assessment. *Language Testing* 19, 453–476.
Bachman, L. & A. Palmer (1996). *Language testing in practice*. Oxford: Oxford University Press.
Baleghizadeh, S. & A. Derakhshesh (2012). The effect of task repetition and noticing on EFL learners' oral output. *International Journal of Instruction* 5, 141–152.
Basturkmen, H., S. Loewen & R. Ellis (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics* 25, 243–272.
Bruton, A. (2002). From tasking purposes to purposing tasks. *ELT Journal* 56, 280–288.
Butler, Y. (2011). The implementation of communicative and task-based language teaching in the Asia-Pacific region. *Annual Review of Applied Linguistics* 31, 36–57.
Cameron, L. (2001). *Teaching languages to young children*. Cambridge: Cambridge University Press.
Carless, D. (2004). Issues in teachers' reinterpretation of a task-based innovation in primary schools. *TESOL Quarterly* 38.4, 639–662.
Dale, E. & J. Chall (1948). A formula for predicting readability: Instructions. *Educational Research Bulletin* 27, 37–54.
Day, E. & S. Shapson (1991). Integrating formal and functional approaches to language teaching in French immersion: An experimental study. *Language Learning* 41, 25–58.
De la Fuente, M. J. (2006). Classroom L2 vocabulary acquisition: Investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research* 10. 3, 263–295.
DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (eds.), *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press, 42–63.
Doughty, C. & E. Varela (1998). Communicative focus on form. In C. Doughty & J. Williams (eds.), *Focus on form in classroom second language* acquisition. Cambridge: Cambridge University Press, 114–138.
Ellis, R. (1991). Grammar teaching – practice or consciousness-raising. In R. Ellis (ed.), *Second language acquisition and second language pedagogy*. Clevedon: Multilingual Matters, 232–241.
Ellis, R. (1997). *SLA research and language teaching*. Oxford: Oxford University Press.

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition* 27, 141–172.
Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics* 19, 221–246.
Ellis, R., H. Basturkmen & S. Loewen (2001). Learner uptake in communicative ESL lessons. *Language Learning* 51, 281–318.
Ellis, R., S. Loewen & R. Erlam (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition* 28, 339–368.
Erlam, R. (2016). I'm still not sure what a task is: Teachers designing language tasks. *Language Teaching Research* 20, 279–299.
Foster, P. & P. Skehan (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition* 18, 299–324.
Goo, J. M. & A. Mackey (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (ed.), *Input, interaction and corrective feedback in L2 learning*. Oxford: Oxford University Press, 379–452.
Goo, J. & A. Mackey (2013). The case against the case against recasts. *Studies in Second Language Acquisition* 35.1, 127–165.
Harley, B. (1989). Functional grammar in French immersion: A classroom experiment. *Applied Linguistics* 19, 331–359.
Hawkes, M. (2012). Using task repetition to direct learner attention and focus on form. *ELT Journal* 66.3, 327–336.
Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
Jobard, G., M. Vigneau B. Mazoyer & N. Tzourio-Mazoyer (2007). Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage* 34, 784–800.
Laufer, B. (2006). Comparing focus on form and focus on forms in second-language vocabulary learning. *The Canadian Modern Language Review* 63, 149–166.
Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge: Cambridge University Press.
Li, S. & R. Ellis (forthcoming). The influence of pre-task grammar instruction on task performance and L2 learning in task-supported language instruction: A process-product study.
Li, S., Y. Zhu & R. Ellis (2016). Task-based versus task-supported language instruction: An experimental study. *Annual Review of Applied Linguistics* 36, 205–229.
Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition* 27, 361–386.
Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (eds.), *Modelling and assessing second language acquisition*. Clevedon: Multilingual Matters, 77–100.
Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (eds.), *Handbook of second language acquisition*. San Diego: Academic Press, 413–168.
Long, M. (2005). Methodological issues in learner needs analysis. In M. Long (ed.), *Second language needs analysis*. Cambridge: Cambridge University Press, 19–76.
Long, M. H. (2006). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum Associates.
Long, M. (2015). *Second language acquisition and task-based language teaching*. Malden, MA: Wiley-Blackwell.
Long, M. H. (2016). In defence of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics* 36, 5–33.
Long, M. & S. Ross (1993). Modifications that preserve language and content. In M. Tickoo (ed.), *Simplification: Theory and application*. Singapore: SEAMEO Regional Language Centre, 29–52.
Lyster, R. (2001). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning* 51 (Suppl. 1), 265–301.
Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition* 26.3, 399–432.
Lyster, R. & L. Ranta (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition* 19.1, 37–66.
Lyster, R. & L. Ranta (2013). The case for variety in corrective feedback research. *Studies in Second Language Acquisition* 34, 591–626.

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

Lyster, R., K. Saito & M. Sato (2013). Oral corrective feedback in second language classrooms. *Language Teaching* 46.1, 1–40.
Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition* 21, 557–587.
McDonough, K. & A. Mackey (2006). Responses to recasts: Repetitions, primed production and linguistic development. *Language Learning* 56.4, 693–720.
Mochizuki, N. & L. Ortega (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research* 12, 11–37.
Norris, J. M. (2016). Current issues for task-based language assessment. *Annual Review of Applied Linguistics* 36, 230–244.
Nunan, D. 1989. *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.
Pitler, E. and A. Nenkova (2008). *Revisiting readability: A unified framework for predicting text quality*. Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing. Honolulu: Association for Computational Linguistics, 86–195.
Prabhu, N. S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.
Quinn, P. (2014). Delayed versus immediate corrective feedback on orally produced passive errors in English. Unpublished Ph.D. thesis, University of Toronto.
Révész, A., M. Michel & R. Gilabert (2016). Measuring cognitive task demands using dual task methodology, subjective self-ratings, and expert judgments: A validation study. *Studies in Second Language Acquisition* DOI: http://dx.doi.org/10.1017/S0272263115000339
Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 285–318.
Robinson, P. (2011). Second language task complexity, the cognition hypothesis, language learning, and performance. In P. Robinson (ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance*. Amsterdam: John Benjamins, 3–38.
Samuda, V. (2001). Guiding relationships between form and meaning during task performance: The role of the teacher. In M. Bygate, P. Skehan & M. Swain (eds.), *Researching pedagogic tasks, second language learning, teaching and testing*. Harlow: Longman, 119–140.
Sasayama, S. (2016). Is a 'complex' task really complex? Validating the assumption of task complexity. *The Modern Language Journal* 100.1, 231–254.
Schmidt, R. (2001). Attention. In P. Robinson (ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press, 3–32.
Scrivener, J. (2005). *Learning teaching: A guidebook for English language teachers*. Oxford: MacMillan Education.
Seedhouse, P. (1997). The case of the missing 'no': The relationship between pedagogy and interaction. *Language Learning* 47, 547–583.
Sheen, R. (1994). A critical analysis of the advocacy of the task-based syllabus. *TESOL Quarterly* 28, 127–157.
Sheen, R. (2006). Focus on forms as a means of improving accurate oral production. In A. Housen & M. Pierrard (eds.), *Investigations in instructed second language acquisition*. Berlin: Mouton de Gruyter, 271–310.
Shintani, N. (2016). *Input-based tasks in foreign language instruction for young learners*. Amsterdam: John Benjamins.
Shintani, N. & R. Ellis (2014). Tracking learning behaviours in the incidental acquisition of two dimensional adjectives by beginner learners of L2 English. *Language Teaching Research* 18, 521–542.
Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics* 17, 38–62.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for language pedagogy. *Annual Review of Applied Linguistics* 36, 34–49.
Suzuki, Y. & R. DeKeyser (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning* 65, 860–895.
Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (eds.), *Input in second language acquisition*. Rowley, MA: Newbury House, 235–253.

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (eds.), *Focus-on-form in classroom second language acquisition*. Cambridge: Cambridge University Press, 64–81.
Swan, M. (2005). Legislating by hypothesis: The case of task-based instruction. *Applied Linguistics* 26, 376–401.
Van den Branden, K. (ed.) (2006). *Task-based language education: From theory to practice*. Cambridge: Cambridge University Press.
Willis, J. (1996). *A framework for task-based learning*. Harlow: Longman.
Willis, D. & J. Willis (2007). *Doing task-based teaching*. Oxford: Oxford University Press.
Yuan, F. & R. Ellis (2003). The effects of pre-task and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics* 24, 1–27.

https://doi.org/10.1017/S0261444817000179 Published online by Cambridge University Press