CrossMark

Annual Review of Applied Linguistics, **36** (2016), pp. 34–49.
© Cambridge University Press, 2016
doi: 10.1017/S0267190515000100

Tasks Versus Conditions: Two Perspectives on Task Research and Their Implications for Pedagogy

PETER SKEHAN
St. Mary's University, Twickenham
peterskehan@gmail.com

ABSTRACT

This chapter explores the contrast between the effects on second language performance of tasks and task characteristics, on the one hand, and the conditions under which tasks are done, on the other. The first major section explores the evidence on this issue and proposes that the impact of conditions such as pretask planning, task repetition, and posttask activities is greater and more consistent than the impact of tasks and features such as time perspective or number of elements. The second major section explores the theoretical accounts that have been proposed regarding tasks and conditions. It is suggested that deductive accounts have, so far, only had limited success regarding the use of tasks, but that psycholinguistic models of speaking do provide a looser but more useful framework to account for the effects of conditions. It is also suggested that an important difference between tasks and conditions concerns the tension between constraint and flexibility in performance and that the flexibility provided by task conditions is an important component in the more dependable results they have generated. Finally, pedagogic implications are discussed linking task conditions to the methodological choices that are available to teachers.

INTRODUCTION

The title of this chapter is not entirely serious: The aim here is not to pit the two elements—task characteristics and the conditions under which tasks are done—in an artificial contest to see which is the winner. But it is the case that a comparison between a focus on tasks and task design, on the one hand, and a focus on performance conditions, on the other, is revealing about a number of things, including organizing the results from studies within the task literature and the challenges faced by theories of second language performance. The contrast is also illuminating as a basis for exploring pedagogic claims that might be made and how task research could be relevant to teachers.

There are three main sections to the chapter. The first provides an overview of findings about task characteristics and performance conditions, given the extensive literature that is now available. The second section explores possible theoretical accounts that might explain why there are differences in extent and strength of findings comparing tasks and conditions. The final section explores whether pedagogic implications can be drawn from the research, suggesting that what has been learned about the effects of task conditions provide a more accessible source of applications to pedagogy and teacher decision making.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

RESEARCH ON TASKS: RESEARCH ON CONDITIONS

This section will explore findings regarding tasks and conditions, and so some definitional clarity is appropriate at the outset. Tasks will be interpreted both as task types such as narrative or decision making, and task characteristics such as task structure, number of elements, and information organization. Task conditions concern choices that can be made about how tasks are implemented in task research. The particular focus taken here is to explore whether there have been planning opportunities, whether a posttask is involved, whether the task is repeated, what the time conditions are while the task is running, and whether the task is interactive or monologic.

Approaches to researching tasks and performance conditions vary in where they stand regarding the inductive and deductive approaches to theorizing (Jordan, 2004). Since these two approaches have influenced a wide range of studies, they need to be briefly outlined (and will be examined in greater depth in the next section). The inductive approach generally takes a case-by-case view, identifying task features from general theory, previous research, or classroom experience, and then explores whether each feature has a systematic relationship with performance, which is typically measured by complexity, accuracy, lexis, and fluency (CALF) (Housen, Kuiken, & Vedder, 2012; Skehan, 2009a, 2009b). As an exemplifying study, Foster and Skehan (1996) contrasted the effects of planning with personal, narrative, and decision-making tasks (with this tripartite distinction arising from an examination of coursebook materials and the task types they included). The study showed linkages between planning and all three performance areas and indicated specific connections between personal tasks and both increased accuracy and increased complexity, and between narrative and decision-making tasks and increased structural complexity. The clearest example of a deductive approach is the cognition hypothesis (Robinson, 2011), which makes central the construct of task complexity and proposes a distinction between resource-directing factors and resource-dispersing factors, with each of these overarching categories then leading to specific performance predictions. A typical study is Malicka and Levkina (2012), which manipulated the resource-directing factors of number of elements and reasoning demands to explore their impact on performance. These researchers report fewer words per clause and fewer errors with spatial expressions, but for a high proficiency group only.

Sasayama (2015) discussed task research in comparable terms to the inductive–deductive contrast, providing two tables that compare the effects on performance of two variables: planning (which I will interpret as more inductive in nature) and the number of elements in a task (which is more deductive, since it derives from the cognition hypothesis). These two tables are adapted here as Table 1, with a few additional studies in the number of elements category. Table 1 gives the study authors and the effects on the performance areas of CALF. The original tables (Sasayama, 2015) provide much more detail about participants, mode, participant organization, and the like.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

TABLE 1. *Studies Researching Performance Effects of Number of Elements and Planning*

No. of Elements	Comp.	Acc.	Lexis	Fluen.	Planning	Comp.	Acc.	Lexis	Fluen.
Robinson, Cadierno, & Shirai (2009)	=	=	+	−	Foster & Skehan (1996)	+	+	X	+
Gilabert (2007)	X	=	X	X	Foster & Skehan (1996)	+	=	X	+
Kuiken & Vedder (2007, 2008)	=	+	=	X	Ortega (1999)	+	+/−	X	+
Michel, Kuiken, & Vedder (2007)	=	+	+	−	Iwashita, McNamara, & Elder (2001)	=	=	X	=
Révész (2011)	−	+	+	X	Yuan & Ellis (2003)	+	=	=	+
Sasayama & Izumi (2012)	+	=	X	X	Ellis & Yuan (2004)	+	=	=	+
Sasayama (2011)	=		=		Mochizuki & Ortega (2008)	+	=	X	=
Michel (2011)	=	=	+	=	Mochizuki & Ortega (2008)	+	+	X	=
Levkina & Gilabert (2012)	=	=	+	−	Tavares (2009)	=	+	X	+
Michel (2013)	=	=	=	=	Farahani & Meraji (2011)	+	=	=	+
					Meraji (2011)	+	+	=	+
					Mohammadzadeh, Dabaghi, & Tavakoli (2013)	=	+	X	=

Notes: Comp. = Structural complexity: Acc.= Accuracy; Fluen: = Fluency. + indicates a positive influence; − indicates a negative influence; = indicates no difference; X indicates that the dependent variable was not studied.

Starting with the number of elements studies in Table 1, structural complexity typically is not affected by number of elements, and that is also the most typical outcome (in six studies) for accuracy, although there are three studies that report a positive effect (and one negative). The most common result for lexis is a positive effect (five studies), with two showing no effect and no studies showing a negative influence. Fluency does show some consistency with four negative effects. So, rather curiously, the clearest effects are with lexis (positive) and fluency (negative), with structural complexity and accuracy decidedly mixed. With planning, there is slightly greater clarity. Structural complexity shows effects in nine studies, no effect in three, and no negative effects. Accuracy has five positive effects and six no effects. Lexis, only represented in four studies, is unaffected in each. Finally, fluency is positively influenced in eight studies and not affected in four, with no negative effects. I conclude from these results that the (deductive in origin) task variable of number of elements has little overall effect on structural complexity and accuracy, some positive effect on lexis (lexical diversity, usually), and a negative effect on fluency. Planning has a very solid positive influence on structural complexity and fluency, a fairly positive effect on accuracy (and see review on this in Ellis, 2009), and no effect on lexis.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

The variables of number of elements and planning have been shown in some detail because they represent the inductive–deductive contrast well, and because they are based on a wide range of studies (by the standards of task-based research). They suggest that the generalizations that are possible in each case contrast: number of elements shows tendencies, although curiously not for what might be considered to be the most central variables for the theory in question. Planning is associated with more robust generalizations, implicating three of the four areas concerned, and with consistency in the literature, as the studies build upon one another in developing our understanding of the effects of planning.

I claim that the way number of elements and planning have been treated is broadly typical of many other inductive–deductive variables. Even so, brief mention of some of the others is worthwhile. Regarding deductive variables, reasoning demands have figured in some studies, such as Malicka and Levkina (2012), who reported fewer words per clause and fewer errors with spatial reasoning, but the latter only with higher proficiency participants. Kuiken and Vedder (2011) also reported an accuracy effect for reasoning demands, particularly accuracy related to lexis, rather than morphosyntax, but no other effects. Gilabert, Barón, & Llanes (2009), based on a study with three different tasks, reported effects for accuracy with a there-and-then narrative (but no other effects on that task), for lexical richness and accuracy with a more complex instruction-giving task (but no effect for structural complexity), and no effects at all for a more complex decision-making task. Révész (2011) reported higher accuracy and lexical diversity for a there-and-then task, but lower structural complexity. The story is similar when other deductive variables are explored—there are effects, but they do not generate consistency or robust generalizations, and almost never demonstrate that complexity and accuracy are jointly raised, as the cognition hypothesis predicts. Consistent with this is a meta-analysis of cognition hypothesis studies (Jackson & Suethanapornkul, 2013), which reported small effects for task complexity raising accuracy, but none for complexity, and only a small effect for fluency, but in the reverse direction to prediction.

There have been some successes with task characteristics, though, but I argue that these successes have had something of an inductive basis. Two will be mentioned here. Task structure emerged in early studies conducted within a limited attention capacity approach (Skehan, 2015a). It was first proposed, post hoc, after Foster and Skehan (1996) and Skehan and Foster (1997), where its presence was inferred after performances on certain tasks showed greater accuracy and fluency. Subsequently, a series of studies have confirmed and extended these effects, as in Skehan and Foster (1999), Tavakoli and Skehan (2005), Tavakoli and Foster (2008), Skehan and Shum (2014), and Wang and Skehan (2014). Interestingly, Tavakoli and Skehan (2005) also generated an unexpected finding in which one of the structured tasks that was used yielded an unpredicted structural complexity effect, and this task was seen to require the integration of foreground and background information. Subsequently, other studies have confirmed this effect (Skehan & Shum, 2014; Tavakoli & Foster, 2008) and broadened the claim to concern information organization and the relationship between elements in a task. So, it does seem that task effects can be found—factors such as familiarity (Bui,

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

2014) and personal relevance (Skehan & Foster, 2008) figure also—but the effects that have tended to be consistent have arisen from ongoing research programs, rather than been the result of theory-derived predictions.

There has also been considerable research effort in exploring the consequences of manipulating the conditions under which tasks are done, and I am going to characterize this as more inductive than deductive in nature. Broadly, this work can be categorized as exploring pretask, during-task, and posttask implementation conditions, and it is interesting that these three stages have clear connections to similar stages that figure when tasks are implemented by teachers—in some ways the practical origin of research into these stages attests to the inductive nature of the inquiry. Regarding the first, pretask work potentially could have explored many things that happen before a task proper is done (constructing splash diagrams, doing parallel tasks, and answering orienting questions). In practice, researchers have concentrated almost exclusively on pretask planning, and this area alone has generated a considerable subliterature, as indicated in Table 1. Planning consistently raises complexity and fluency, with impressive effect sizes. It is less consistent in raising accuracy, though this often happens (Ellis, 2009), although with smaller effect sizes when significance is achieved. In addition, more detailed generalizations are now possible, suggesting that pretask planning has greater effects on more complex tasks (Skehan & Foster, 1997); that during the available planning time, attention is first given to accuracy, then complexity, and that fluency benefits from additional time for planning up to 10 minutes (Mehnert, 1998); and that planning can be directed toward specific forms (Mochizuki & Ortega, 2008). In addition, qualitative studies of planning (Ortega, 2005; Pang & Skehan, 2014) suggest different learner reactions to planning opportunities and that some planning behaviors are more effective than others at raising performance (and some are even harmful).

Ellis (2005), reacting to the finding that pretask planning is not so strongly associated with accuracy effects, advocated research with *online* planning, that is, planning conducted while speaking is going on, possible when there are unpressured time conditions of speaking. In a series of studies (such as Ellis & Yuan, 2005; Yuan & Ellis, 2003), Ellis reported that providing online planning opportunities does have a selective influence on accuracy—less time pressure enables monitoring and the concern and space to avoid error. However, in a carefully controlled study, Wang (2014) showed that online planning works most effectively when it is allied with pretask planning opportunities. In Leveltian terms, the pretask planning supports and stimulates conceptualizer work, and then the online planning opportunities facilitate formulator operations, leading to the greater accuracy.

Online planning is an example of a during-task condition. There are others of this sort. For example, Skehan and Foster (2005) reported that giving learners additional (and conflicting) information midway through an interactive, decision-making task had surprisingly little impact on performance. In contrast, they report from this same study that participants found it difficult to sustain levels of performance, with statistically significant differences between the first 5 and second 5 minutes of performance. Interactivity is another during-task condition, relating clearly to the contrast between a monologic performance and one in which the presence of

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

an interlocutor and the potential for joint development of a task are important. Michel (2011) argued that interactivity, particularly if associated with alignment (Pickering & Garrod, 2004) leads to higher performance. Foster and Skehan (1996) also reported that higher performance results when tasks provide potential for more engaged interaction (Skehan & Foster, 2008), and that also when there is interaction, other effects (e.g., from a posttask) lead to bigger differences between experimental groups (Foster & Skehan, 2013).

The existence of a posttask condition has also yielded interesting results. The intention with this work is to induce an anticipation effect with participants so that the way a task is done is influenced by the expectation of what will happen later. Following Willis (1996), Skehan and Foster (1997) used the threat of a subsequent public performance (of the same task), hypothesizing that anticipation of this later condition (and its possibly higher stakes) would lead to greater accuracy in the actual performance as attention is selectively directed toward this area. It did, but any effect only reached statistical significance in one task out of three that were used. Subsequently, using the different posttask operationalization of participants knowing that they would have to transcribe their own performance subsequently (Lynch, 2001), a more personal and language-focused posttask activity, Foster and Skehan (2013) showed effects on accuracy for both tasks used, a narrative and a decision-making task, and also a complexity effect for the latter. These results were broadly confirmed by Li (2014), who also showed that different manipulations of how the transcribing was done could lead to stronger accuracy effects or complexity effects.

An interesting condition effect, and a growing literature, has focused on the effects of repeating a task. Earlier research (Bygate, 2001) has shown that repetition after a time interval of 10 weeks is associated with a higher level of performance, despite the time interval. The second performance is more condensed, and the complexity of language is greater. Interestingly, Wang (2014), with immediate repetition of a video-based narrative retelling, showed significant differences in performance with complexity, accuracy, and fluency—an unusual combination of improvements—and this with large effect sizes. Skehan (2014) suggested that the repeated performance benefits from traces of the partial lemma retrieval that has taken place during the first performance, since this retrieval leaves traces that can be activated more readily in the second performance, thus easing formulator operations. Priming may also be an influence. These accounts are consistent with the finding (e.g., Bygate, 2001) that task repetition does not generalize to other examples of the same task type—the gain is associated with the specific lexical elements within the particular task, although learners are likely to learn something about the advantages of depth of processing.

TASKS VERSUS CONDITIONS: THEORY

To summarize the previous section covering trends in research results, the claim is essentially that the more deductively motivated studies have tended to generate

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

results that focus on task characteristics, but the results tend to be less consistent in nature and do not present a clear picture in relation to the original hypotheses. In contrast, the more inductively oriented studies, whether of tasks or conditions (though the studies exploring conditions are far more numerous) have tended to produce more consistent results. The studies that contribute to this range of findings have tended to be relatively low-theory in nature and often contain developments from previous literature (e.g., planning, repetition) or are connected with pedagogic suggestions (posttask conditions, interactivity).

Of course, this claim is about the current state of affairs, rather than anything that is ideal. A functioning theory, which generates research studies and makes falsifiable predictions, would be desirable. This would hold out the prospect of greater understanding and explanation and, as a result, potentially more effective implications for pedagogy. But the main point is that we are where we are, and so one can ask whether the preconditions are right for the development of ambitious and wide-ranging theory, or whether the more inductive approach is all there is for now. It is the purpose of this section to consider these issues and to explore how much of a contribution theory can make, as well as what existing results can contribute to theory.

Clearly, the most influential deductive approach we have at present is the cognition hypothesis. The empirical results associated with this hypothesis have already been covered, and they have been seen to be only partially supportive. Here we will focus on the theoretical claims in their own terms, partly in an attempt to understand the results that have been reported. Three problem areas can be considered (and more extensive discussion can be found in Skehan, 2015a):

- There is the difficulty that the constructs that underlie the hypothesis are problematic. The distinction between resource-directing and resource-dispersing variables is not totally clear and often seems to have a slightly arbitrary quality. Alternative analyses of the different detailed characteristics are possible (Skehan, 2015a). Planning, for example, through what we have learned from qualitative studies (Ortega, 2005; Pang & Skehan, 2014) could just as easily be considered a resource-directing variable, as could task structure (Wang & Skehan, 2014). There is also the difficulty that the different subvariables are difficult to separate, as, for example, number of elements and reasoning demand, which can easily be confounded.
- The central construct of the hypothesis, task complexity, needs considerable clarification. The difficulty is that putative manipulations of task complexity may not in reality produce different levels of complexity. We are currently seeing impressive efforts to overcome this problem, particularly through the use of secondary measures, such as time estimation, dual-task performance, ratings of difficulty and mental effort (Sasayama, in press). For now, this essentially practical difficulty has yet to be overcome, although one can be optimistic about the future.
- There is the general problem of participant interpretation of tasks. Task designers and researchers try to use tasks that will generate interpretations and performances that are predicted. But the reactions of the participants in research studies are, of course, something potentially different. The desired outcome of a task might be interpreted differently by the participant and the task designer; the relevance of the

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

task to the individual completing it and its familiarity to the participant will vary; and participants may also subvert the task to play to their strengths and away from their weaknesses. Task studies that have qualitative components (Ortega, 2005; Pang & Skehan, 2014; Sasayama, 2015) consistently show that task interpretation is a major (and often unacknowledged) influence.

More broadly than these individual points, there is, with task performance, something of a tension between constraint and flexibility. Typically, the tasks that researchers use are chosen because of some expectation or even detailed prediction. They generally meet criteria for "taskness," and so typical tasks will give meaning primacy, require real-time communication, have a (researcher-chosen) purpose, and so on, and so can be considered to capture reasonably spontaneous spoken language. But they will contain elements that are difficult to avoid completely, such as the number of features a holiday might contain, the number of people to be rescued and the equipment available in a building fire, the route through a map, the characters and events within a picture-based narrative, or the years in prison that a particular crime is supposed to merit. The constraints that are involved (and that are probably integral to questions of task design in a particular study) may well prevent the participant from exploring alternative approaches to a problem. Speakers are not able to shape the discourse in ways to make it more personal or to capitalize on what they know and what they are good at, or to avoid areas in which they feel that they are weak (e.g., certain areas of vocabulary). In all these cases, constraint is a major influence on how the task is done, and the scope for the speaker to make personal interpretations and to create an effective conceptualizer–formulator balance is limited. Of course, one might say that such manipulations are precisely the point of experimental designs, and indeed, if consistent and predicted results were the norm, the previous analysis would be one associated with the success of the approach. But the point being made here is that these constraints seem to be associated with lack of findings, probably because they lead learners into a form of communication where the need to comply with task requirements washes out any opportunity for significant differences between manipulated task contrasts to emerge.

Turning to implementation conditions with tasks, the earlier empirical section presented the claim that in this more inductively oriented area, consistent and nontrivial results have been much more the norm. Given the above analysis of difficulties in demonstrating effects of task characteristic manipulation, why has a different pattern of findings emerged with conditions? Of course, one interpretation could be that an inductive approach is more appropriate, given the state of development of the research with task-based performance: It has the advantage, mentioned earlier, that "problems" emerge from practical settings, and then when some results are published, it becomes more possible to build upon them, and extend and consolidate them. But this interpretation can only be part of the story, and it needs to have some sort of theoretical underpinning, even if this is less tight than in more deductive approaches.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

Particularly interesting in this regard are psycholinguistic theories of first language speaking. The model that has been most influential for task condition research has been Levelt's model of first language speaking (Levelt, 1989, 1999). This model contains three general phases: conceptualization—in which ideas are assembled, a stance is taken, and a preverbal message is developed; formulation—in which first the lemmas relevant to the preverbal message are accessed, and then the process of syntax building is embarked upon; and articulation—in which phonetic plans are produced to clothe the output of the formulator in sound. These stages, in the first language case, operate in parallel (Kormos, 2006), hence the speed with which we can produce speech. In the second language case, the more limited state of the second language mental lexicon means that parallel processing is more difficult (Kormos, 2006), and the speaker often has to resort to a serial mode of functioning as problems at one of the three stages consume more attention than is feasible if parallel processing is to be sustained (Skehan, 2014, 2015b), and then the speaker has to try to regain a parallel mode.

What is interesting here is how readily the different condition effects, summarized above, can be related to this view of second language speaking. We can illustrate this through the pretask planning research. Pretask planning has clear connections with conceptualizer operations. From the large number of quantitative studies (Ellis, 2009), we have seen connections with greater complexity and greater fluency in performance, suggesting that often pretask planning is concerned with the ideas to be expressed. From the smaller number of qualitative studies (Ortega, 2005; Pang & Skehan, 2014), it has become clear that there are effective connections between participants who concentrate on ideas at the planning stage and greater complexity in actual performance. But this qualitative work also suggests that while the priority in pretask planning relates to the conceptualizer, some rehearsal (Ellis, 2005) also takes place, and this may then ease subsequent formulator operations, and possibly connect with raised accuracy. In addition, some of the fluency effects seem to come from the lemma preparation stage, which eases the syntax-building stage within the formulator, following the lemma retrieval that the planning has primed (Skehan, 2014). Online planning studies (Ellis & Yuan, 2005; Wang, 2014) explore how time pressure during speaking has selective influences on performance. Here the connection is with formulator operations, and claims are made (Ellis & Yuan, 2005) that accuracy is enhanced. Wang (2014), however, argued that online planning in itself is important but not sufficient: There also needs to be scope for the conceptualizer to be involved, and it is the synergy between effective the conceptualizer and formulator operations that leads to increased accuracy and complexity.

Similar arguments can be developed for other condition effects, although these will not be analyzed in the same detail. Posttask effects on accuracy seem to be the result of formulator operations and specifically monitoring (a central process in the Levelt model) where attention is selectively directed to avoiding error at this stage in performance (Foster & Skehan, 2013). Repetition is also largely formulator based, in that the first performance triggers partial lemma retrieval in that the less extensive and organized second language mental lexicon yields less rich

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

information about the lemma on fast, first access, but the traces of this activity persist and facilitate the second performance, supporting more effective syntax building and message creation (Skehan, 2014). Interactivity similarly can be related to both conceptualizer and formulator operations. The former involves the joint construction of ideas during the development of interaction, while the latter is implicated through the eased time pressure and opportunity to regroup while one's interlocutor is speaking, as well as the chance to exploit the language introduced by one's interlocutor (Foster & Skehan, 2013). Finally, Skehan (2014) offered a two-edged analysis of how support materials (or their absence) function. On the one hand, support materials reduce memory demands and so release attentional resources for general processing, both for the conceptualizer but especially the formulator. On the other hand, if the materials have to be integrated into the task performance, they may constrain what has to be said and, as such, present a hurdle for the speaker since there is less scope to play to one's strengths and avoid difficulties. In this case, conceptualizer demands may be made that the formulator finds it difficult to handle. In contrast, of course, absence of support materials may make memory demands greater but enable "packaging" of what is to be said that is more flexible, is less constraining, and enables the conceptualizer to shape and ease the demands on the formulator. A final point is that processing stages complement nicely the earlier point regarding flexibility. Such stages do not constrain—they generally provide flexibility for speakers to exploit in any way that they want, without any need to incorporate specific task features. The significant results that have been found support the contention that most learners approach these opportunities in a broadly similar manner and capitalize on the flexibility that exists to exploit speech production processes more effectively.

Thus far we have considered how the existing (inductively generated) findings can be interpreted, after the event, through the Levelt model. But as indicated earlier, theory-based approaches are ultimately more satisfying, and so exploring what can now be done to build upon existing findings, but motivated by theory, is a desirable path. Skehan (2015a) discussed these possibilities in relation to second language applications of the Levelt model, suggesting two exemplifying lines of development. The first concerns the central role of the second language mental lexicon as the basis for the formulator's response to the problems posed by the ideas in the preverbal message. Skehan suggested that if the second language mental lexicon grows as proficiency increases, then the separation between CALF measures should diminish, and that complexity and accuracy should come together more at higher proficiency levels. A second possibility concerns the way conceptualizer- and formulator-derived variables (such as information organization and structure, respectively) can be used to produce joint effects on performance. Other predictions, based on other combinations of variables (task or condition-oriented), have considerable potential for new studies. In each case here, the basis for the prediction is the structure of the Levelt model, adapted to second language spoken performance. The research has yet to be done. The central point is that theory can be applied to what has been a largely inductive tradition, and now interesting predictions are possible.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

At the beginning of this section on theory, then, the emphasis was on the more consistent results (particularly regarding task implementation conditions) of a more inductive approach to research, with the possibility that a more deductive approach might be a little premature. But we have seen that there are some robust results with task characteristics, and we have also seen that there are more general psycholinguistic processing accounts that, at a lower level, can be used to interpret existing findings and, at a more demanding level, can make predictions about second language performance when assessed through CALF measures. In many ways, therefore, the discussion has concerned tactics of research given the general state of development of understanding of second language performance. The central argument is that a first language speech production model is the best bet for second language spoken performance research. Inductive research relating to conditions has delivered a range of results. On this basis we are now more ready to develop second language processing theories.

PEDAGOGY

Broadly, then, the research literature appears to show contrasting fortunes for tasks and for conditions, and this generalization inevitably may have some implications for the usefulness of research findings for pedagogy, as we shall see next. First, though, there is an even more fundamental question to consider. In the past, it has been argued that the advent of communicative language teaching and a task-based approach to instruction somewhat eroded the distinction between syllabus and methodology. Syllabus is conventionally interpreted as concerning the units of instruction (e.g., grammatical structures and functions) and principles for their sequencing (White, 1987). Methodology is concerned with the implementation of whatever units compose the syllabus (e.g., through presentation, practice, and production), partly through classroom activities that follow different methodology traditions (Larsen-Freeman, 2000) and partly through teacher decision making. The erosion of the syllabus–methodology distinction is based on the idea that the task as the syllabus unit integrally shapes the nature of methodology and classroom activities, since the way the task is done depends on the nature of the task itself. This chapter has argued, in contrast, that the difference between tasks (syllabus) and implementation conditions (methodology) is still a valid one, and that research into the latter has delivered more consistent results than research into the former.

Looking at syllabus, then, the simpler challenge with tasks is to know what implications there are for performance by selecting any particular task. The more difficult challenge, however, would be to have solid reasons not simply for task selection but also for sequencing. To its considerable credit, the cognition hypothesis, through analysis of task complexity, addresses this challenge. Unfortunately, as we have seen, current research is not particularly supportive as to how this could be done, and so we have to conclude that the contributions of task research with regard to syllabus design thus far have been somewhat piecemeal. Some insights have been achieved about task qualities that have systematic connections with

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

performance, but not much more than that—certainly not enough to be the basis for practical intervention beyond identifying useful individual tasks.

In contrast, if we switch to methodology, the research we have covered is more encouraging. The claims that could be made are much more robustly based (and go some distance to justifying the syllabus–methodology distinction) and are worth exploring further. As a starting point, Skehan (1998) discussed a pedagogic sequence involving the performance dimensions generally explored in task research and suggested that there is a possible move from structural or lexical complexity (reflecting new, challenging language but used haltingly and with error) to the beginning of control (with accuracy achieved, but relatively slowly, and possibly with less challenging language) to a final stage of fluency (where not only is accuracy achieved, but this is done at reasonable production speed). (And, of course, this sequence is repeated with the many items and subsystems that are learned in the course of language development.) If we accept this sequence, then condition-oriented research can feed into pedagogic, methodological decision making and allow teachers to target areas of language performance that they have identified as desirable, and to do so in a way that strikes a balance between form and meaning. This, possibly, would work with whatever task is chosen, for whatever pedagogic reason. In this way conditions that promote complexity, or accuracy, or fluency could be identified, as well as conditions that produce combinations of these, such as a posttask condition with interactive tasks (Foster & Skehan, 2013). Then this information could be used by teachers wishing to achieve particular pedagogic goals, given the current position of their learners.

There is another point that can be made here regarding implementation conditions and methodology. The sorts of variables investigated in condition-oriented research are relatively simple to use in pedagogic situations. Pretask planning is a fairly straightforward option to apply, with perhaps the only complication being decisions on the details of planning, such as the source or focus of planning (Foster & Skehan, 1999) or whether instructions might be given to push the planning in any particular direction (Mochizuki & Ortega, 2008). In addition, learners might find that having planning time allows them to invest a more personal element to a task, as they shape it to reflect themselves as human beings. Online planning, repetition, and posttask conditions would similarly be easy to apply and could be used with a whole range of tasks. In addition, to the extent that research findings of performance improvements are replicated in the pedagogic situations, then learners' motivation levels may be enhanced if they realize higher performance levels have been achieved. There is even the possibility (unresearched as yet) that teachers have something of an ace they can play here. The opportunity to use particular conditions over extended time periods might themselves contribute to our understanding of how conditions could be trained to become even more effective as learners become accustomed to their use. Best of all, perhaps, exploiting condition-oriented research and its findings would not have cost implications. There are no materials or books to buy, or technology to obtain—giving planning time or requiring a task to be repeated are activities the teacher can use whenever they want, and without costs!

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

A word of warning is needed here. The use of tasks and conditions focuses on performance and using whatever interlanguage one has to maximum effect. It is possible to imagine new language being present in input, being noticed, and then progressively incorporated into a growing, underlying system, consistent with the complexity-accuracy-fluency sequence outlined earlier. But such learning through interaction is unlikely to be the whole story. Skehan (2007, 2013) argued that a second posttask stage (distinct from anticipation-oriented posttask activities such as public performance and transcription) is important here. Such a stage would capitalize on the emerging language or the language problems that had been noticed by learners in the earlier task input (Schmidt, 1990) or output (Swain, 1995) and then use such language as the starting point for further pedagogic work, which might be aimed at explicating, or restructuring, or consolidating, or practicing, as appropriate. Such an approach would then be more convincing in supporting the change in the interlanguage system, including the introduction of new language, that is needed since its starting point would be language whose salience is derived from learners themselves during the earlier task performance.

REFERENCES

Bui, H. Y. G. (2014). Task readiness: Theoretical framework and empirical evidence from topic familiarity, strategic planning, and proficiency levels. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 63–94). Amsterdam, The Netherlands: John Benjamins.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Harlow, UK: Longman.

Ellis, R. (2005). Planning and task-based performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 3–36). Amsterdam, The Netherlands: John Benjamins.

Ellis, R. (2009). The differential effects of three types of task planning on fluency, complexity, and accuracy in L2 oral performance. *Applied Linguistics*, *30*, 474–509.

Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167–192). Amsterdam, The Netherlands: John Benjamins.

Farahani, A. A. K., & Meraji, S. R. (2011). Cognitive task complexity and L2 narrative writing performance. *Journal of Language Teaching and Research*, *2*, 445–456.

Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition, 18*(3), 299–324.

Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, *3*, 215–247.

Foster, P., & Skehan, P. (2013). The effects of post-task activities on the accuracy of language during task performance. *Canadian Modern Language Review, 69*, 249–273.

Gilabert, R. (2007). The simultaneous manipulation of task complexity along planning time and [+/- here-and-now] effects on L2 performance. In M. P. Garcia-Mayo (Ed.), *Investigating tasks informal language settings* (pp. 44–68). Clevedon, UK: Multilingual Matters.

Gilabert, R., Barón, J., & Llanes, À. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics in Language Teaching*, *47*, 367–395.

Housen, F., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy, and fluency in SLA*. Amsterdam, The Netherlands: John Benjamins.

Iwashita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information-processing approach to task design. *Language Learning*, *51*, 401–436.

Jackson, D. O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning, 63*(2), 330–367.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

Jordan, G. (2004). *Theory construction in second language acquisition*. Amsterdam, The Netherlands: John Benjamins.
Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Erlbaum.
Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching*, *45*, 261–284.
Kuiken, F., & Vedder, I. (2008). Cognitive tasks complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing*, *17*, 48–60.
Kuiken, F., & Vedder, I. (2011). Task complexity and linguistic performance in L2 writing and speaking: The effect of mode. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 91–104). Amsterdam, The Netherlands: John Benjamins.
Larsen-Freeman, D. (2000). *Techniques and principles in language teaching* (2nd ed.). Oxford, UK: Oxford University Press.
Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, UK: Cambridge University Press.
Levelt, W. J. (1999). Producing spoken language: A blueprint of the speaker. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 83–122). Oxford, UK: Oxford University Press.
Levkina, M., & Gilabert, R. (2012). The effects of cognitive task complexity on L2 oral production. In Housen, A., Kuiken, F., & Vedder, I. (Eds.), *Dimensions of L2 performance and proficiency: investigating complexity, accuracy, and fluency in SLA* (pp. 171–198). Amsterdam, The Netherlands: John Benjamins.
Li, Q. (2014). Get it right in the end: The effects of post-task transcribing on learners' oral performance. In P. Skehan (Ed.), *Processing perspectives on task performance*. Amsterdam, (pp. 129–154). The Netherlands: John Benjamins
Lynch, T. (2001). Seeing what they meant: Transcribing as a route to noticing. *English Language Teaching Journal, 55*, 124–132.
Malicka, A., & Levkina, M. (2012). *Measuring task complexity: Does EFL proficiency matter?* In A. Shehadeh & C. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and implementation* (pp. 42–66). Amsterdam, The Netherlands: John Benjamins.
Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition, 20*, 83–108.
Meraji, S. R. (2011). Planning time, strategy use, and written task production in a pedagogic vs. a testing context. *Journal of Language Teaching and Research*, *2*, 338–352.
Michel, M. C. (2011). Effects of task complexity and interaction on L2 performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 141–174). Amsterdam, The Netherlands: John Benjamins.
Michel, M. C. (2013). The use of conjunctions in cognitively simple versus complex oral L2 tasks. *The Modern Language Journal, 97*, 178–195.
Michel, M., Kuiken, F., & Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching*, *45*, 241–259.
Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research, 12*, 11–37.
Mohammadzadeh, M. A., Dabaghi, A., & Tavakoli, M. (2013). The effects of simultaneous use of preplanning along +/-Here-and-Now dimension on fluency, complexity, and accuracy of Iranian EFL learners' written performance. *International Journal of Research Studies in Language Learning*, *2*, 49–65.
Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21*, 109–148.
Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 77–109). Amsterdam, The Netherlands: John Benjamins.
Pang, F., & Skehan, P. (2014). Self-reported planning behaviour and second language performance in narrative retelling. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 95–128). Amsterdam, The Netherlands: John Benjamins.
Pickering, M., & Garrod, S. (2004). Towards a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–226.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *The Modern Language Journal*, *95*, 162–181.
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* (pp. 3–38) Amsterdam, The Netherlands: John Benjamins.
Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language production. *Applied Linguistics, 28*, 533–554.
Sasayama, S. (2011). Cognition hypothesis and second language performance: Comparison of written and oral task performance. *Second Language Studies*, *29*, 107–129.
Sasayama, S. (2015). *Validating the assumed relationship between task design, cognitive complexity, and second language task performance* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.
Sasayama, S. (in press). Is a "complex" task really complex? Validating the assumption of cognitive task complexity. *Modern Language Journal*.
Sasayama, S., & Izumi, S. (2012). Effects of task complexity and pre-task planning on EFL learners' oral *production*. In A. Shehadeh & C. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and implementation* (pp. 23–42). Amsterdam, The Netherlands: John Benjamins.
Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 17–46.
Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, UK: Oxford University Press.
Skehan, P. (2007). Task research and language teaching: Reciprocal relationships. In S. Fotos (Ed.), *Form- meaning relationships in language pedagogy: Essays in honour of Rod Ellis*. Oxford, UK: Oxford University Press.
Skehan, P. (2009a). Models of speaking and the assessment of second language proficiency. In A. G. Benati (Ed.), *Issues in Second Language Proficiency*. (pp. 202–215). London: Continuum.
Skehan, P. (2009b). Lexical performance by native and non-native speakers on language-learning tasks. In B. Richards, H. Daller, D. D. Malvern, & P. Meara (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application* (107-124). London, UK: Palgrave Macmillan.
Skehan, P. (2013). Nurturing noticing. In J. M. Bergsleithner, S. N. Frota, & J. K. Yoshioka (Eds.), *Noticing and second language acquisition: Studies in honor of Richard Schmidt* (pp. 169–180). Honolulu: University of Hawai'i, National Foreign Language Resource Center.
Skehan, P. (2014). Synthesising and applying task research. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211–260). Amsterdam, The Netherlands: John Benjamins.
Skehan, P. (2015a). Limited attentional capacity and cognition: Two hypotheses regarding second language performance on tasks. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 123–155). Amsterdam, The Netherlands: John Benjamins.
Skehan, P. (2015b). Working memory and second language speech performance. In Z. Wen, M. Borges Mota, & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 189–201). Clevedon, Avon, UK: Multilingual Matters.
Skehan, P., & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task based learning. *Language Teaching Research*, *1*, 185–211.
Skehan, P., & Foster, P. (1999). Task structure and processing conditions in narrative retellings. *Language Learning, 49*(1), 93–120.
Skehan, P., & Foster, P. (2005). Strategic and online planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 193–216). Amsterdam, The Netherlands: John Benjamins.
Skehan, P., & Foster, P. (2008). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard, & I. Vedder (Eds.), *Complexity, accuracy, and fluency in second language use, learning, & teaching* (pp. 207–226). Brussels, Belgium: KVAB.
Skehan, P., & Shum, S. (2014). Structure and processing condition in video-based narrative retelling. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 197–210). Amsterdam, The Netherlands: John Benjamins.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in the study of language* (pp. 125–144). Oxford, UK: Oxford University Press.

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 58*(2), 439–473.

Tavakoli, P., & Skehan, P. (2005). Planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–276). Amsterdam, The Netherlands: John Benjamins.

Tavares, M. G. G. (2009). The relationship among pre-task planning, working memory capacity, and L2 speech performance: A pilot study. *Linguagem & Ensino*, *12*, 165–194.

Wang, Z. (2014). Online time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). Amsterdam, The Netherlands: John Benjamins.

Wang, Z., & Skehan, P. (2014). Task structure, time perspective and lexical demands during video-based narrative retellings. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 155–186). Amsterdam, The Netherlands: John Benjamins.

White, R. V. (1987). *The ELT Curriculum*. London, UK: Heinemann.

Willis, J. (1996). *A framework for task-based learning*. London, UK: Longman.

Yuan, F., & Ellis, R. (2003). The effects of pre-task and online planning on fluency, complexity, and accuracy in L2 monologic oral production. *Applied Linguistics*, *24*, 1–27.

https://doi.org/10.1017/S0267190515000100 Published online by Cambridge University Press