Language Learning & Language Teaching   **11**

Planning and Task Performance in a Second Language

Edited by
Rod Ellis

John Benjamins Publishing Company

Planning and Task Performance in a Second Language

Language Learning and Language Teaching

The *LL&LT* monograph series publishes monographs as well as edited volumes on applied and methodological issues in the field of language pedagogy. The focus of the series is on subjects such as classroom discourse and interaction; language diversity in educational settings; bilingual education; language testing and language assessment; teaching methods and teaching performance; learning trajectories in second language acquisition; and written language learning in educational settings.

Series editors

Nina Spada
Ontario Institute for Studies in Education, University of Toronto

Jan H. Hulstijn
Department of Second Language Acquisition, University of Amsterdam

Volume 11
Planning and Task Performance in a Second Language
Edited by Rod Ellis

Planning and Task Performance in a Second Language

Edited by

Rod Ellis

University of Auckland

John Benjamins Publishing Company

Amsterdam / Philadelphia

∞™ The paper used in this publication meets the minimum requirements
of American National Standard for Information Sciences – Permanence
of Paper for Printed Library Materials, ANSI Z39.48-1984.

Table of contents

Preface

The purpose of this book is to bring together a series of articles on the nature of planning and its effects on task-based performance in laboratory, classroom and testing contexts. The idea for the book originated in a colloquium on this topic given at AILA Conference in Singapore in December 2002. Papers given by Bygate and Samuda, Elder and Iwashita, Ellis and Fanguan, and Sanguran were subsequently developed into chapters for this book. A number of other researchers (Batstone, Foster, Ortega, Kawauchi, Skehan, and Tavakoli) were later invited to submit chapters and did so.

Planning and its role in task-based performance are of both theoretical interest to second language acquisition (SLA) researchers and of practical significance to language teachers. In the case of SLA researchers, planning is important because it links in with the current interest in the role of attention in language learning. Whether learners plan strategically before they perform a task or engage in careful within-task planning, opportunities arise for them to attend to language as form, or as Ortega (Chapter 3) puts it 'form-in-meaning'. Thus, investigating planning serves as one way of studying what learners attend to and what effect it has on the way they use language. Further, it is also hypothesized that the kind of language use that learners engage in will influence the process of acquisition itself. Its significance for language teachers lies in the fact that planning is a relatively straightforward way of influencing the kind of language that learners produce. It serves, therefore, as an effective device for intervening indirectly in interlanguage development.

The predominant methodological paradigm in planning studies is experimental. That is, the task performance of learners who engage in planning of one kind or another is compared with a task performance where there is no opportunity for planning. This paradigm continues to be reflected in several of the studies reported in this book (e.g. the chapters by Kawauchi, Ellis and Yuan, and Skehan and Foster). It has proved very fruitful in demonstrating that planning does indeed affect the way in which learners perform a task. Nevertheless, this paradigm also has its limitations. It tells us nothing about what learners actually do when they are planning; it does not show us whether learners actually

do what they planned to do; and, more crucially, perhaps, it fails to recognize that planning and task-performance constitute social as well as cognitive activities.

Clearly, then, there is a case for broadening the paradigm to incorporate both a process element and to acknowledge the social nature of tasks. A number of the chapters in this book address planning as a process. Ortega extends her earlier research on tasks to examine the strategies that learners use when engaged in pre-task planning. Sanguran (Chapter 4) discusses how the instructions learners are given can influence the way in which they plan. Several authors report the results of post-task questionnaires designed to investigate how learners responded to the opportunities to plan. Skehan and Foster (Chapter 7) undertake a detailed analysis of what they call 'breakdown fluency' with a view to identifying process features of task performance that will provide evidence of on-line planning. All of these studies extend the research on planning in significant ways.

There is less evidence of any attention to the social aspect of planning and task-performance. The prevailing tenor of this book is psycholinguistic. In the concluding chapter, however, Batstone (Chapter 10) develops a convincing argument for a social perspective. He points out that learners can approach tasks in two different ways – as requiring economical and efficient communication or as providing opportunities for them to engage in learning activities. The idea that tasks always have a context and that this context will help to shape how learners plan for and perform them is further supported in the two chapters that address the role of task planning in a testing situation (by Elder and Iwashita [Chapter 8] and Tavakoli and Skehan [Chapter 9]). The very different results of these two studies are perhaps best explained in terms of the differences in the specific testing contexts.

It is to be hoped, then, that this book both reflects mainstream research into the role of planning in task-based performance and also extends it.

Rod Ellis
Auckland, April 2004

Section I

Introduction

The last decade has seen a growing body of research investigating various aspects of L2 learners' performance of tasks (see, for example, Bygate et al. (2001) and Ellis (2003)). This research has focused broadly on a variety of design features of tasks and implementation procedures and how these impact on such aspects of language use as comprehension, input processing, meaning negotiation and the fluency, complexity and accuracy of L2 production (Skehan 1996, 1998a). While task-based research has been able to identify a number of variables that impact on performance (e.g. whether contextual support is available, whether the information is shared or split, whether the outcome is closed or open, whether there is inherent structure to the task's content), the results have not always been consistent. This has led some researchers (e.g. Coughlan & Duff 1994) to argue that the 'activity' that results from a 'task' is necessarily co-constructed by the participants on each occasion, making it impossible to predict accurately or usefully how a task will be performed.

However, one implementation variable that has attracted considerable attention and that has been shown to produce relatively consistent effects on L2 production is task planning. A number of studies (e.g. Foster & Skehan 1996) have shown that when learners have the opportunity to plan a task before they perform it, the language they produce is more fluent and more complex than when no planning is possible. Other studies (e.g. Yuan & Ellis 2003) have shown that unpressured on-line planning also has predictable effects, albeit somewhat different from those arising from pre-task planning.

The choice of planning as the variable for investigation in this book is motivated both by its importance for current theorizing about L2 acquisition (in particular with regard to cognitive theories that view acquisition in terms of information processing) and its value to language teachers, for unlike many other constructs in SLA, 'planning' lends itself to pedagogical manipulation. The study of task planning, then, provides a suitable forum for establishing the interconnectedness of theory, research and pedagogy in SLA (Pica 1997).

This introductory chapter has a number of purposes. It seeks to provide a framework for the subsequent chapters by identifying and defining different

types of planning. It examines the theoretical backgrounds that have informed the study of planning in task-based performance. It reviews earlier research that has investigated the effects of the different types of planning. It examines a number of key methodological issues related to the study of the effects of planning on task performance.

Planning and task-based performance
Theory and research

Rod Ellis
University of Auckland

Types of planning

All spoken and written language use, even that which appears effortless and automatic, involves planning. That is speakers and writers have to decide what to say/write and how to say/write it. Planning is essentially a problem solving activity; it involves deciding what linguistic devices need to be selected in order to affect the audience in the desired way. As Clark and Clark (1977) noted, planning takes place at a number of different levels, resulting in discourse plans, sentence plans and constituent plans, all of which have to be interwoven in the actual execution of a language act.

Principal types of task planning

Figure 1 distinguishes two principal types of task-based planning – *pre-task planning* and *within-task planning*. These are distinguished simply in terms of when the planning takes place – either before the task is performed or during its performance. Pre-task planning is further divided into *rehearsal* and *strategic planning*. Rehearsal entails providing learners with an opportunity to perform the task before the 'main performance'. In other words, it involves task repetition with the first performance of the task viewed as a preparation for a subsequent performance. Strategic planning entails learners preparing to perform the task by considering the content they will need to encode and how to express this content. In pre-task planning, the learners have access to the actual task materials. It is this that distinguishes strategic planning from other types of pre-task activity (e.g. brainstorming content; studying a model performance of

Rehearsal

Pre-task planning

Strategic planning

Planning

Pressured

Within-task planning
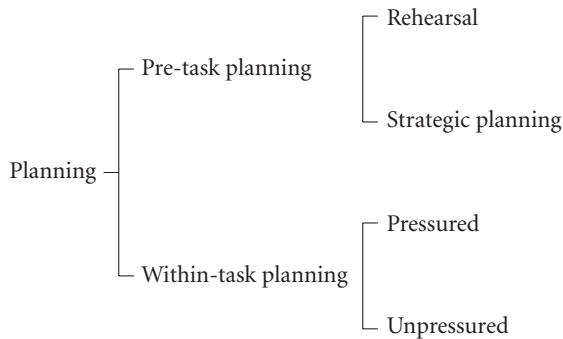
Unpressured

Figure 1. Types of task-based planning

the task; dictionary search). Within-task planning can be differentiated according to the extent to which the task performance is pressured or unpressured. This can be achieved most easily by manipulating the time made available to the learners for the on-line planning of what to say/write in a task performance. In an unpressured performance learners can engage in careful on-line planning resulting in what Ochs (1979) has called 'planned language use'. In pressured performance learners will need to engage in rapid planning resulting in what Ochs calls 'unplanned language use' (although, of course, all language use involves some level of planning). Ochs documents a number of linguistic differences between the two types of discourse. For example, unplanned discourse tends to manifest non-standard forms acquired early whereas planned discourse contains more complex, target-like forms.

While pre-task planning and within-task planning constitute distinctive types of planning they should not be seen as mutually exclusive. As shown in Figure 2, it is possible to envisage four basic combinations of the two planning conditions. In condition 1, learners have no opportunity for either pre-task planning or unpressured within-task planning. Given that learners (especially with low proficiency) have a limited processing capacity and are likely to experience difficulty in accessing and encoding their linguistic knowledge, this can be considered the most demanding condition. In condition 2, learners are given the opportunity to pre-plan their performance (either by means of task rehearsal or strategic planning) but are not allowed to plan their utterances carefully on-line. In condition 3, the reverse occurs; learners are required to start performing the task straight away but are given as much time as they wish to take. Both of these conditions may ease the processing burden of the learner.

Planning conditions	Pre-task planning	Unpressured within-task planning
1	No	No
2	Yes	No
3	No	Yes
4	Yes	Yes

Figure 2. Planning conditions

Condition 4, where the learner has the opportunity for both pre-task planning and unpressured within-task planning, can be expected to create the conditions that help learners maximize their competence in performance.

Sub-categories of task planning

Both pre-task and within-task planning can be categorized further in ways not shown in Figure 1 but which are of potential theoretical and practical significance. For example, learners can be left to their own devices when planning a task (unguided planning) or they can be given specific advice about what and how to plan (guided planning). In this case, they can be directed to attend to linguistic form, to meaning or to form and meaning. Chapter 4 by Sangarun, for example, explores how directing learners to focus on some specific aspect of language in their strategic planning of tasks influences subsequent performance. Earlier studies (e.g. Hulstijn & Hulstijn 1984) have explored the effects of directing attention to form or meaning on within-task planning and performance. Another option relevant only to strategic planning concerns participatory structure, i.e. whether the planning is undertaken by the learners working individually, collaboratively in small groups, or with the teacher (see Foster & Skehan 1999). As Batstone discusses in the concluding chapter to this volume this can potentially affect the way a task is performed.

Clearly, which types and combinations of types of planning are of relevance must ultimately be decided empirically. That is, each type/option needs to be systematically examined to establish if it has any effect on the language produced in a task performance. As we will see when we examine the previous research on planning and task-based performance this has been one of the major goals of enquiry to date.

Theoretical background to the study of planning in task-based research

I will consider three theoretical frameworks that have informed the study of task planning in second language acquisition (SLA) research. These are (1) Tarone's (1983) account of stylistic variation, (2) models of speech production and writing, and (3) cognitive models of L2 performance and language learning. These theories explicitly or implicitly draw on three central constructs involved in psycholinguistic accounts of language processing – attention and noticing, a limited working memory capacity, and focus-on-form – so I will begin by briefly outlining each of these constructs, as they have been applied in SLA research.

L2 production as information processing: Some key constructs

Information processing models constitute the dominant approach to theorizing about language comprehension and production in cognitive psychology today. While the current models differ in some major ways (see Robinson 1995 for a review of these), they all share a number of features; they all seek to account for how information is stored and retrieved; they all view information processing as involving input, temporary storage of material attended to, long-term storage of (some of) this material and mechanisms for accessing this material from long-term memory. Lantolf (1996) has referred to this general approach as the 'computational model' as it is based on an analogy between the human mind and a computer.

There are a number of general principles that inform this model (Huitt 2003). One is the assumption of a limited capacity. That is, there are limits on the amount of information that human beings can process from input or for output. These limits cause bottlenecks in working memory and can lead to language users prioritizing one aspect of language over another. A second principle is that there is a control mechanism that language users will need to access when they are confronted with a new task for which they do not possess proceduralized linguistic knowledge. This control mechanism draws on explicit stored knowledge. As such, it uses up processing power and thus taxes working memory. A third principle is that human beings process information by means of both top-down processes that draw on encyclopedic knowledge of the world and on situational context and bottom-up processes that involve close attention to the linguistic signals in the input. These general principles underlie the three central constructs discussed below.

1. *Attention and noticing*

In a number of seminal articles in the 90s, Schmidt (1990, 1994) advanced the hypothesis that conscious attention, or what he called 'noticing', is essential for language learning. He states 'although unattended stimuli may have subtle but undeniable effects on humans (as in subliminal perception experiments), it is widely argued in psychology that *learning* without attention to what is to be learned is impossible' (Schmidt 1994:17). He goes on to argue that in the case of learning attention must necessarily be conscious as 'all demonstrations of detection without conscious registration . . . demonstrate only the processing of what is already known, not learning'. This is a view that has not gone un-challenged, however. In particular, Tomlin and Villa (1994) have proposed that three components of attention can be distinguished; alertness (a general readi-ness to deal with incoming stimuli), orientation (the aligning of the attentional mechanisms to some specific aspect of language) and detection (the actual process by which a specific feature of language is attended to focally). They claim that none of these components necessarily involves consciousness and that even detection can occur without any conscious registration of the stim-uli attended to. More recently, Schmidt (2001) has been less dogmatic about whether (conscious) attention is required, writing 'the question of whether *all* learning from input requires attention to that input remains problematic, and conceptual issues and methodological problems have combined to make a definitive answer elusive' (p. 29). He continues to assert, however, that in-tentional, conscious attention is beneficial for learning as it can help learners process features of language that otherwise would not be noticed.

Much of the discussion of noticing (as conscious attention) in language learning has focussed on its role in input processing and, as such, might be seen as having little relevance to theorizing about how task planning aids acqui-sition. Task planning, whether of the pre-task or within-task type, may involve learners attending to the linguistic input provided in the task materials (e.g. in a text reformulation task), but in many tasks (e.g. those that involve a pic-torial rather than verbal input) it clearly does not. Planning primarily entails learners accessing their own implicit and explicit knowledge of the L2 for use in production, as suggested by Ochs' (1979) account of planned language use. The question arises, then, as to whether noticing has any role to play in output-processing. Swain (1985b, 1995) claims that it does. According to the Output Hypothesis, production requires learners to process syntactically, which in-volves bottom-up rather than top-down processing and requires attention to form. Similalarly, Robinson (2001b) suggests that output as well as input re-quires attention and that the degree of attention will depend on the complexity

of the task they asked to perform, with more complex tasks requiring more attention. Providing learners with the opportunity to plan a task, therefore, may aid performance. However, as we will see later, there is some disagreement as to how pre-task planning affects attention. One view is that it encourages greater attention to form during task performance, resulting in increased accuracy and complexity. An alternative view, promulgated by Robinson, is that pre-task planning simplifies the task and thus obviates the need to attend closely to form during performance but assists automatic access to stored language and so leads to greater fluency.

2. *Limited working memory capacity*

There are number of models of working memory (see Miyake & Shah 1999). One of the most commonly cited in the task planning literature is that of Baddeley (e.g. Baddeley & Hitch 1974; Baddeley & Logie 1999). This identifies three components of working (or short-term) memory; the central executive or supervisory attentional system, the phonological loop, and the visual spatial sketchpad. Two of these seem relevant to a role for task planning (i.e. not the visual spatial sketchpad).

The central executive system governs the relationship between working memory and long-term memory, allocating attention to specific long-term memory systems. This system is limited in capacity, and thus the extent to which language learners are able to attend to a specific system will depend on the extent to which other systems are automatized. For example, if learners use up available processing space in lexical searches the attention they can pay to grammar will be limited. Providing learners with the opportunity for pre-task planning or for unpressured within-task planning can ease the burden on working memory, allowing learners the opportunity to engage in controlled processing and to process multiple systems linearly.

The phonological loop is comprised of two sub-components – the phonological store, which affords a temporary representation of material drawn from the input or long term memory, and a mechanism that allows for articulatory rehearsal, which enables decaying material introduced into working memory to be sustained. Planning is likely to draw extensively on this component, allowing learners to maintain one set of material while drawing on another set to modify or refine it. For example, learners will be able to access linguistic material from their interlanguage grammars and maintain this in the phonological loop while they edit it through reference to their explicit knowledge of the L2. In other words, the phonological loop is likely to play a central role in monitor-

ing (discussed below). In short, planning is seen as a means of helping learners overcome the limitations in capacity of their working memory.

3. *Focus-on-form*

The term 'focus-on-form' has been variably used in the SLA literature. It helps to distinguish three related but different senses of the term, depending on whether the perspective is a pedagogic one, a discoursal one or a psycholinguistic one. In the context of language pedagogy, focus-on-form refers to attempts to intervene in the process of acquisition by inducing learners to pay attention to linguistic form while they are primarily concerned with decoding or encoding message content. These attempts can be planned (i.e. a specific form is selected for attention) or incidental (i.e. specific forms are attended to as the need arises). In discoursal terms, focus-on-form refers to the pre-emptive and reactive devices that interlocutors use to draw attention to form while learners are engaged in performing some task that gives priority to message conveyance. These devices can consist of 'queries' (i.e. questions about linguistic form) or various types of implicit and explicit corrective feedback (e.g. reformulations of learners' incorrect utterances, known as 'recasts', and metalinguistic explanation). In psycholinguistic terms, 'focus-on-form' refers to the mental processes involved in selective attention to linguistic form while attempting to communicate. 'Noticing', discussed above, serves as a cover term for these processes.

SLA researchers argue that L2 acquisition, especially in the case of adult learners, requires a focus-on-form. There are two rationales for this claim. The first relates back to the idea that learners have a limited working memory capacity and therefore experience difficulty in attending to meaning and form at the same time (see, for example, VanPatten 1990). Because it is 'natural' for learners to give priority to meaning, they may overlook certain linguistic features, especially those that are non-salient, redundant or do not contribute to meaning. As a result they need to be induced to attend to the formal aspects of the language. The second, more controversial claim is that interlanguage development can only take place if learners attend to form *while* they are engaged with meaning. As Doughty and Williams (1998) put it 'the fundamental assumption of FonF instruction is that meaning and use must already be evident to the learner at the time that attention is drawn to the linguistic apparatus needed to get the meaning across' (p. 4). They propose that there is a 'cognitive window for the provision of focus on form' of up to 40 seconds; that is, learners are able to hold material in working memory for this length of time during which they have the opportunity to attend to the form of what they have tem-

porally stored. Doughty (2001) suggests that 'roving attention' enables learners to pay attention to form without interruption of their original speech plan. The theoretical and empirical bases for these proposals are reviewed in detail in Doughty (2001).

Providing learners with the opportunity to plan a task performance constitutes a means of achieving a focus-on-form pedagogically. It mitigates the limitations of their working memory by allowing learners the 'cognitive window' needed to attend to form while they are primarily concerned with message conveyance. In other words, it creates a context in which learners have the opportunity to map form onto meaning by accessing linguistic knowledge that is not yet automatized.

Theoretical bases for task planning

The three constructs discussed above all figure to a greater or lesser extent in the theories of language use/acquisition that I will now consider. The three theories to be considered are presented chronologically, reflecting their origins in the history of task-based research. In each case I will outline the theory and then consider how it has been applied to task planning.

1. *Tarone's theory of stylistic variation*

Tarone's theory draws heavily on Labov's account of stylistic variation in native speakers. Labov (1970) argued that 'there are no single style speakers'; that is, individual speakers manifest variation in their use of language because they are able to draw on a variety of 'styles'. Further, he argued that 'these styles can be ranged along a single dimension according to the amount of attention that speakers pay to their speech' (i.e. focus on form). Depending on the situation, speakers vary in the extent to which they monitor their speech. Attention through monitoring is greatest in speech that reflects a careful style and least in the vernacular style found in everyday speech. Labov was able to show that what he called 'style shifting' was probabilistic but also systematic and therefore predictable. That is, speakers tended to use one variant in one style and another variant in another style to a greater or lesser extent depending on whether the social context encouraged them to pay attention to what they said.

Drawing on this theory of intra-speaker variability, Tarone (1983) proposed what she called the Capability Continuum for L2 learners. This consists of a continuum of styles, ranging from the 'careful' to the 'vernacular', which Tarone saw as comprising the learner's L2 knowledge. To explain how L2 development takes place, Tarone proposed two ways in which new forms can enter

interlanguage. In one way, forms originate in the learner's vernacular style and then spread to the more careful styles over time. In the other way, forms appear initially in the learners' most careful style, manifest only when the learner is paying close attention to speech production, and then spread to the less formal styles where they replace earlier, more primitive forms. Subsequent empirical work (e.g. Tarone 1985; Tarone & Parrish 1988) was directed at showing how the choice of forms was strongly influenced by the nature of the task learners were asked to perform. However, contrary to expectations, these studies did not always show that the more target-like forms occurred with greater frequency in tasks designed to elicit a careful style.

Viewing learners' L2 knowledge as a 'capability continuum', then, can explain how planning assists L2 production and acquisition. In the case of unpressured online planning, as in conditions 3 and 4 in Figure 2, learners will be able to attend to their speech and thus access their careful style. This will be reflected in greater accuracy (i.e. a more target-like performance). However, the provision of opportunity for careful on-line planning may not in itself promote acquisition. In this respect, pre-task planning followed by the pressured performance of a task (i.e condition 2 in Figure 2) may be more effective. Pre-task planning allows learners to access their careful style but then requires them subsequently to perform the features they have accessed in real time where close attention to speech is not possible, thus encouraging the spread of these features from the careful to the vernacular style.

Nevertheless, the theory lacks explanatory power. First, it does not account for why some forms are more target-like in the learner's vernacular style. Second, the role of attention is not clearly specified. Third, the key notion of 'spread' is underdeveloped. The theory originated in a social account of language variation but planning is essentially a psycholinguistic construct. Current research on the role of planning has turned to theories that offer a fuller psycholinguistic account of L2 production.

2. *Models of speech production and writing*
By far the most influential theory where studies of task planning are concerned is Levelt's (1989) model of speech production. Many of the later chapters (i.e. Chapters 2, 4, 6, 7 and 9) draw on this model. The model was developed to account for the speech production of native speakers but de Bot (1992) has adapted it for bilingual speech production.

Levelt's (1989) model identifies three autonomous processing stages: (1) conceptualizing the message, (2) formulating the language representation, and (3) articulating the message.

The **conceptualizing** stage involves three sub-stages. First, the speaker decides upon the communicative goal. In the second substage (macro-planning) the speaker develops the communicative goal into a series of sub-goals and then identifies a speech act for each sub-goal that will achieve the intended effect. In the third sub-stage (micro-planning), the speaker retrieves the information needed to realize each of the subgoals and organizes it by determining 'the information perspective of [an] utterance, its topic, its focus, and the way in which it would attract the addressee's attention" (Levelt 1989: 5). The product of the micro planning is a preverbal message that is not linguistic in nature but contains, nonetheless, all information needed to convert the preverbal message into language. This preverbal message is then forwarded to the formulator.

Formulation involves establishing language representations of the preverbal messages by retrieving lexical items from the speaker's mental lexicon. Each lexical item is comprised of two kinds of information: 'lemma' and 'lexeme'. The lemma contains information about the meaning and syntax of each lexical item, while the lexeme contains information about its morphological and phonological properties. Thus, retrieving a lexical item serves to prompt the syntactic building procedure required for grammatical encoding. This results in a 'surface structure' (i.e., 'an ordered string of lemmas grouped in phrases and subphrases of various kinds' (Levelt 1989: 11)), which is then processed by the phonological encoder, resulting in a phonetic or articulatory plan (i.e., "an internal representation of how the planned utterance should be articulated" (Levelt 1989: 12)). Levelt (1989) calls this 'internal speech'.

Finally, this internal speech is transferred to the **articulator**. The articulator retrieves chunks of internal speech that are temporarily stored in an articulatory buffer and then "unfolds and executes [them] as a series of neuromuscular instructions" (p. 27). This leads, ultimately, to the production of overt speech.

These three stages are regulated by a self-monitoring process consisting of three subsystems. The first subsystem inspects whether the preverbal message matches the speaker's original intention. It does this before the message is sent on to the formulator to be converted into internal speech. The second subsystem inspects the internal speech before it is articulated as overt speech. Finally, the third subsystem inspects the overt speech that has been generated.

Levelt (1989) also identified two characteristics of speech production which are relevant to task planning; (1) controlled and automatic processing and (2) incremental production. According to Levelt, some of the components of the speech production process (specifically, the conceptualizer and the monitor) operate under controlled processing, while other components (specifically, the formulator and the articulator) operate automatically in the

main. In addition, he proposed that speech production processes can take place in parallel.

De Bot (1992) considers the adaptations to Levelt's model needed to account for speaking in an L2. He suggests that in the case of the conceptualizer, macro-planning is not language specific but micro-planning is (i.e. the pre-verbal message specifies which language (or languages) are to be used to encode the message). De Bot argues that there are separate systems for the L1 and L2 as far as the processing components of the formulator are concerned, although the two systems are likely to be connected in at least some areas. In contrast, given the cross-linguistic influences evident in L2 pronunciation, he considers the existence of two separate systems for articulation 'very improbable' (p. 17). We might also note that whereas L1 speakers are able to carry out the processes involved in formulation and articulation (but not conceptualisation) without attention, L2 learners (especially those with limited L2 proficiency) are more likely to need to activate and execute their linguistic knowledge through controlled processing. Thus, they are likely to experience problems during the formulation and articulation stages, as these processes are demanding on working memory.

Levelt's model is explicitly designed to account for <u>speech</u> production. However, available theories of <u>writing</u> (e.g. Bereiter & Scardamalia 1987; Hayes & Flower 1980; Grabe 2001; Grabe & Kaplan 1996; Kellog 1996; Zimmerman 2000) posit a very similar set of processes to those proposed by Levelt. There is also general acceptance that these processes will be broadly similar in both L1 and L2 writing. Kellog's (1996) model, for example, distinguishes three basic systems involved in written text production. Each system has two principal components or processes. **Formulation** entails (1) 'planning', where the writer establishes goals for the writing, thinks up ideas related to these goals and organizes these to facilitate action, and (2) 'translating', where the writer selects the lexical units and syntactic frames needed to encode the ideas generated through planning and represents these linguistic units phonologically and graphologically in readiness for execution. **Execution** requires (3) 'programming', where the output from translation is converted into production schema for the appropriate motor system involved (e.g. handwriting or typing) and (4) 'executing', the actual production of sentences. **Monitoring** consists of (5) 'reading', where the writer reads his or her own text ('a necessary but not sufficient condition for writing well', p. 61) and (6) 'editing', which can occur both before and after execution of a sentence and can involve attending to micro aspects of the text such as linguistic errors and/or macro aspects such as paragraph and text organization. The extent to which a writer is able to engage in monitoring will depend

in part on whether the writer has the time to adopt a 'polished draft strategy' or is engaged in pressured text production. Kellog, like Levelt for speaking, emphasises that writers simultaneously activate formulation, execution, and monitoring processes, although the extent to which this is achievable depends on working memory.

Kellogg also suggests how the different components of the model relate to working memory. He argues that the central executive, a multi-purpose system responsible for problem-solving (see above), mental calculation and reasoning, is involved in all the sub-processes with the exception of executing, which, he argues, is usually accomplished without the need for controlled processing. It should be noted, however, that this assumes an adult, native-like automaticity in handwriting or typing, which may be lacking in some L2 learners, especially those whose first language (L1) employs a different script. It is possible, there-fore, that the central executive may be called upon by some L2 writers during execution. Kellog suggests that the visuo-spatial sketchpad, which stores and processes visual and spatial information in working memory, is only involved in planning. Finally, he proposes that the phonological loop, which stores and processes auditory and verbal information, is required for both translating and reading. The key feature of Kellog's model is that the central executive has limited capacity, with the result that a writer may have to make decisions about which writing process to prioritise when under pressure to produce text rapidly. This is reflected in a trade off of attention directed at the different pro-cesses. Formulation demands are seen as critical, taking priority over execution and monitoring.

These models provide a basis for considering what components of lan-guage production (spoken or written) learners focus on while planning and also for examining what effects planning strategies have on actual production. Rehearsal, for example, may provide an opportunity for learners to attend to all three components in Levelt's model – conceptualisation, formulation and articulation – so it would seem reasonable to assume that this type of pre-task planning will lead to all-round improvements when the task is repeated, as found by Bygate (1996). Strategic planning can be considered likely to assist conceptualisation in particular and thus contribute to greater message com-plexity and also to enhanced fluency, as found by Wendel (1997). Unpressured within-task planning may prove beneficial to formulation and also afford time for the controlled processing required for monitoring. As a result, accuracy might increase. In other words different types of planning can be predicted to ease the pressure on the learner's limited working memory in different ways,

variably affecting the competition and trade-offs evident in different aspects of language production, as claimed by Skehan and Foster (1997).

The main advantage of these models of language production, then, is that they offer a detailed description of what is involved in speaking and writing and thereby afford relatively precise hypotheses about the effects that planning will have on task performance. In one respect, however, they are more limited than Tarone's theory of stylistic variation; they do not account for how linguistic change takes place, for, as Levelt (1989) has pointed out, they constitute steady-state models. Thus, while the models can explain the relationship between planning and language use they do not address how language use contributes to language acquisition.

3. *Cognitive models of task-based performance and learning*

Skehan's (1998b) 'cognitive approach' is based on a distinction between an exemplar-based system and a rule-based system. The former is lexical in nature and includes both discrete lexical items and ready-made formulaic chunks of language. The linguistic knowledge contained in this system can be easily and quickly accessed and thus is ideally suited for occasions calling for fluent language performance. The rule-based system consists of abstract representations of the underlying patterns of the language. These require more processing and thus are best suited for more controlled, less fluent language performance. They are needed when learners have to creatively construct utterances to express meaning precisely or in sociolinguistically appropriate ways.

Skehan also distinguishes three aspects of production; (1) fluency (i.e. the capacity of the learner to mobilize his/her system to communicate in real time, (2) accuracy (i.e. the ability of the learner to perform in accordance with target language norms) and (3) complexity (i.e. the utilization of interlanguage structures that are 'cutting edge', elaborate and structured). He suggests that language users vary in the extent to which they emphasize fluency, accuracy or complexity, with some tasks predisposing them to focus on fluency, others on accuracy and yet others on complexity. These different aspects of production draw on different systems of language. Fluency requires learners to draw on their memory-based system, accessing and deploying ready-made chunks of language, and, when problems arise, using communication strategies to get by. In contrast, accuracy and, in particular, complexity are achieved by learners drawing on their rule-based system and thus require syntactic processing. Complexity is distinguished from accuracy in that it is related to the 'restructuring' that arises as a result of the need to take risks whereas accuracy reflects the learner's attempt to control existing resources and to avoid errors.

Whereas Skehan's research assumes that learners possess a limited processing capacity such that trade-offs between fluency, accuracy and complexity (especially these last two) are likely to occur, Robinson's (2001c) research is premised on a multiple-resources view of processing – that is, that learners, like native speakers, have the capacity to attend to more than one aspect of language at the same time. According to this view, structural complexity and functional complexity are not in competition, as Skehan claims, but are closely connected such that increasing the cognitive complexity of a task is hypothesized to lead to greater linguistic complexity <u>and</u> accuracy as a result of increased output modification and input incorporation.

In Robinson's theory, task complexity is determined by two sets of features, 'resource directing' (e.g. whether or not the task requires reasoning) and 'resource depleting' (e.g. whether or not there is opportunity for strategic planning). These two factors 'interact and affect task production in measurable ways' (p. 31). Optimal attention to form arises when the task is resource directing and not resource depleting, as would be in the case when a task requires reasoning and there is no opportunity for strategic planning. Robinson argues that such a task is likely to enhance complexity and accuracy at the expense of fluency. In contrast a simple task that has no reasoning demands and allows opportunity for strategic planning is likely to promote fluency but not accuracy or complexity.

It is clear, then, that Skehan's and Robinson's models afford contradictory predictions as to the effects of planning on language performance. According to Skehan's model, strategic planning provides an opportunity for learners to access their rule-based system and thus makes them less reliant on their exemplar-based system. It may also assist them in taking the risks needed to access 'cutting edge' interlanguage features rather than relying, conservatively, on more fully acquired features. Thus, it is predicted to enhance linguistic complexity to the detriment of accuracy. In contrast, in Robinson's model, strategic planning is seen as a resource-depleting factor that works hand in hand with negative resource-directing factors to determine the overall complexity of the task and the extent to which learners attend to form when they perform the task, resulting potentially in increased fluency but decreased accuracy and complexity. However, as Robinson (2001b) admits the majority of studies of strategic planning have not supported his claim as they indicate a positive effect on complexity and, sometimes, on accuracy (see the section following). Neither Skehan nor Robinson consider the effects of unpressured on-line planning but presumably this can be hypothesized to work in similar ways to strategic planning (but see Skehan and Foster's chapter in this book).

Type of planning	Message content (Conceptualisation)	Formulation	Monitoring
1. Pre-task planning	Yes	Yes	No
2. Unpressured on-line planning	No	Yes	Yes

Figure 3. Planning and task performance

What insights do these various theories provide about how planning (1) affects task performance (spoken or written production) and (2) L2 acquisition? As shown in Figure 3, planning can impact on both the content learners communicate when performing a task and on their choice of language. In the case of the latter, planning is seen as important because of the role it can play in helping learners to access their L2 knowledge through controlled processing and, according to Skehan's theory, in promoting selective attention to form and monitoring. However, in accordance with the above discussion, the two principal types of planning – pre-task planning and unpressured on-line planning can be seen as impacting somewhat differently on these dimensions of performance. Thus, whereas pre-task planning contributes to the conceptualization of message content while also assisting controlled processing and selective attention to form, unpressured on-line planning has little impact on message content but facilitates language choice in formulation by allowing for controlled processing and selective attention to form and also monitoring.

While the theories are informative about how planning influences the performance of tasks, they are less convincing about how it contributes to acquisition. Extrapolating from performance to acquisition requires acceptance of a number of underlying assumptions:

1. Interlanguage development occurs while learners are primarily focused on message conveyance (i.e. performing tasks).
2. Interlanguage development is facilitated by selective attention to form.
3. Because learners have a limited working memory capacity, attention to form requires opportunity for controlled processing.
4. As a result of the opportunity for the selective attention made possible by controlled processing, learners are able to access more 'advanced' linguistic forms during the formulation stage of production and to achieve greater accuracy through monitoring than is possible in automatic processing.
5. One aspect of language use that fosters acquisition is the production of language that is complex and accurate (cf. Swain's Output Hypothesis).

These assumptions appear inherently reasonable, but, as we will see when I review the extant research on task planning, there is as yet very little empirical evidence in support of them. In particular, there is a notable lack of support for assumption 5, which is fundamental to the claim that planned language use assists acquisition.

Previous research on task planning

In line with the preceding typology of planning types, I will review the previous research on task planning by considering studies that have investigated pre-task planning and unpressured on-line planning. Studies of task-planning in a testing context will be considered separately.

Pre-task planning

1. *Rehearsal*

The research on rehearsal suggests that it has a beneficial effect on learners' sub-sequent performance of the same task but that there is no transference of the rehearsal effect to a different task, even when this is the same type as the original task. Bygate (1996) compared one learner's retelling of a Tom and Jerry cartoon on two separate occasions, three days apart. He found that rehearsal enhanced complexity, with the learner using more lexical verbs (as opposed to copula), more regular past tense forms (as opposed to irregular), a wider range of vocabulary and cohesive devices (e.g. words like 'then', 'so' and 'be-cause'), and fewer inappropriate lexical collocations on the second occasion. There were also more self-correcting repetitions on the second telling of the story. Bygate (2001) reports a larger study that sought to investigate the ef-fects of practicing specific types of task (involving narrative and interview) on both a second performance of the same task and on performance of a new task of the same type. The study showed that the second performance manifested greater fluency and complexity and also that the opportunity to practice that particular type of task helped. However, the practice did not appear to assist performance of a new task of the same type. In other words, disappointingly, there was no transfer of practice effect. Gass et al. (1999) report very similar findings in a study that compared learners' use of L2 Spanish in tasks with the same and different contents. In this study an effect for task repetition on rat-ings of overall proficiency, accuracy in the use of 'estar' (to a lesser extent) and

lexical complexity (type-token ratio) was found. However, again there was no transfer of these effects to a new task.

Lynch and McLean (2000; 2001) made use of a unique task that involved rehearsal. In the context of an English for specific purposes course designed to prepare members of the medical profession to give presentations in English, they designed a 'poster carousel' task. This required students to read an academic article and prepare a poster presentation based on it. Students then stood by their posters while other members of the group visited and asked questions. Altogether, each 'host' had six 'visitors'. Given that visitors tended to ask the same questions, there was substantial opportunity for retrial. Lynch and Mclean document how recycling output resulted in both greater accuracy and fluency. However, they noted that different learners appeared to benefit in different ways with level of proficiency the key factor. Thus, whereas a learner with low proficiency appeared to benefit most in terms of accuracy and pronunciation, a learner with higher proficiency used the opportunity for retrial to improve the clarity and economy of her explanations of a complex idea. Lynch and McLean also report considerable variation in the learners' awareness of the changes they were making in their production.

Task rehearsal, then, seems to have beneficial effects on learner performance. As Bygate (1999) suggests, learners are likely to initially focus on message content and subsequently, once message content and the basic language needed to encode it has been established, to switch their attention to the selection and monitoring of appropriate language. Bygate suggests that rehearsal may afford learners the extra processing space they need 'to integrate the competing demands of fluency, accuracy and complexity'. Bygate and Samuda, in Chapter 2, provide further evidence of this. However, it may not be inevitable that learners switch attention from content to form on the second performance. Nemeth and Kormos (2001) found that repeating an argumentative task influenced the number of supports the participants provided for their claims but had no effect on the frequency with which lexical expressions of argumentation were used. Also, before any strong claims can be made for rehearsal it will be necessary to show that the gains evident from repeating a task transfer to the performance of new, similar tasks.

2. *Strategic planning*
The role of strategic planning has attracted considerable attention from researchers. An effect on all three dimensions of production – fluency, accuracy and complexity – has been found. Each dimension will be considered sepa-

rately. First, though, I will consider research that has investigated what learners do when they plan strategically.

To date, only two studies have investigated what learners actually do when they are given the opportunity to plan. Wendel (1997) interviewed his learners immediately on completion of the tasks. They varied somewhat in what they reported doing during the planning time but all of them said they had focussed on sequencing the narrative events in chronological order. Only 3 reported attending to grammar but even these admitted it did not help them much when it came to telling the stories. As one learner put it: 'I feel like I'm pushing to tell you what's going on in the film. I focus on story, not grammar'. Wendel concluded that it is not useful for learners to try to plan the details of grammatical usage off-line. Ortega (1999) used retrospective interviews to investigate what learners did while they performed a narrative task. She found that they adopted an identifiable approach in their planning (e.g. they worked on the main ideas and organization first and then on the details), they attended to both content and linguistic form, and they made a conscious effort to plan at the utterance level. Ortega also reports that the learners varied considerably in the emphasis they gave to form and content, a point that she elaborates on further in Chapter 3. These two studies suggest that, when planning strategically learners tend to prioritize content. However, Ortega's study suggests that, not surprisingly, they do also attend to form.

Several studies indicate that strategic planning helps to enhance **fluency**. Studies by Foster (1996), Foster and Skehan (1996), Skehan and Foster (1997), Wendel (1997), Mehnert (1998), Ortega (1999) and Yuan and Ellis (1993) all report that giving learners the opportunity to plan results in greater fluency (i.e. a faster speaking rate and fewer dysfluencies). Foster (1996) and Foster and Skehan (1996) report that planners paused less frequently and spent less time in total silence than non-planners in all three tasks they investigated. However, the effect on fluency was stronger on the more difficult narrative and decision-making tasks than on the easier personal task. Skehan and Foster (1997), using similar tasks, replicated the result for total pauses. Wendel (1997) found that the planners in his study produced more syllables per minute and showed a lower mean length of pause in two narrative tasks. Ortega (1999) found a faster speech rate in learners of L2 Spanish on a story-telling task when they had an opportunity to plan strategically. Yuan and Ellis (2003) also report a clear effect for strategic planning on fluency. Foster (2001) found that planning resulted in learners producing a greater amount of speech whereas it led to native speakers producing less. Interestingly, Foster reports that the percentage of learner talk accomplished by means of lexicalised sequences did not change from the un-

planned to planned condition (i.e. it remained steady at about 17%) whereas that of the native speakers did change (from 32% in the unplanned to 25% in the planned). Her study suggests that planning opportunities may be used differently by learners and native speakers when the former lack an extensive store of lexicalized chunks and thus are forced to rely more on rule-based procedures in both planned and unplanned talk. Planning enables learners to access their rule-based procedures more speedily but not, so it would seem, to alter the balance of their use of formulaic and rule-based resources.

A question of obvious interest is what effect the amount of time allocated for planning has on fluency. A reasonable assumption is that the length of planning time is positively correlated with the degree of fluency. Mehnert (1998) set out to investigate this, allocating different groups of learners 0 minute, 1 minute, 5 minutes and 10 minutes of planning time. In general, she found that fluency did indeed improve in relation to the length of planning time. However, the main effect for fluency was evident between the non-planners and the planners; the differences among the three planning groups were mostly non-significant. Thus, providing learners with longer planning time did not have a major effect on the fluency of their speech.

In most of these studies, learners were simply given the task materials and told to plan what they wanted to say. However, a number of studies examined the effects of different kinds of strategic planning. Foster and Skehan (1996) investigated the effects of more guided planning. They compared the effects of 'undetailed' and 'detailed' planning, where the learners were given metacognitive advice about how to attend to syntax, lexis, content, and organization. The results showed that, in line with the overall effect of planning on fluency, for the narrative task the guided planners were notably more fluent than the unguided planners, but that there was no marked difference for the personal and decision-making tasks. This study suggests that the type of planning interacts with the type of task to influence fluency. Foster and Skehan (1999), however, found that asking learners to focus on form or meaning had no differential effect on fluency. Much may depend on the precise instructions given to the learners, as Sanguran (see Chapter 4) suggests. The study she conducted did find that focussing on form, meaning or form/meaning combined had an effect on fluency. Skehan and Foster also investigated the source of planning, comparing the effects of (1) teacher-led planning, (2) individual learner planning and (3) group-based planning on task performance. Where fluency was concerned, (2) proved most effective. However, as Batstone points out in Chapter 10, the ineffectiveness of the group-based planning may reflect the way in which the groups were constituted.

In contrast to fluency, the effects of strategic planning on **accuracy** are quite mixed. A number of studies reported that strategic planning led to increased accuracy. In Ellis (1987), I found that planning that provided opportunities for both strategic and on-line planning resulted in more accurate use of the regular past tense. Mehnert (1998) reported a significant difference in the accuracy of 1-minute planners over non-planners. However, the 5-minute and 10-minute planners performed at the same overall level of accuracy as the 1-minute planners. Other studies found no effect (e.g Crookes 1989; Wendel 1997). Yuan and Ellis (2003), using a general measure of accuracy, also found that strategic planning had no effect, a result that contrasted with that which they reported for unpressured on-line planning (see below). A number of studies found that strategic planning assisted accuracy only on some structures, some tasks and in some conditions. Ortega (1999) reported mixed findings – planning led to greater accuracy in the case of Spanish noun-modifier agreement but not in the case of articles. Foster and Skehan (1996) reported that both undetailed and detailed planners produced fewer errors than the non-planners on a decision-making task, that only the undetailed planners were more accurate than the non-planners on a personal task, while no effect for planning on accuracy was evident on a narrative task. Skehan and Foster (1997) found that planning (undetailed) led to greater accuracy on the personal and narrative tasks but not on the decision-making task. Foster and Skehan's (1999) study of the effects of source of planning found that accuracy was greatest when the planning was teacher-led. However, rather surprisingly, directing learners' attention to form as opposed to content during planning had no effect on accuracy.

It would appear from these results that whether strategic planning has any effect on accuracy may vary depending on a variety of factors, including the extent to which particular learners are oriented towards accuracy, the learners' level of proficiency, the type of task, the length of planning time available, and the particular grammatical feature. Also, with the exception of Yuan and Ellis (2003), these studies made no attempt to control for on-line planning. Thus, it is possible that the different results reflect whether learners were able to or chose to engage in monitoring while they performed the task. In terms of the Levelt model, strategic planning can be expected to aid conceptualisation but the impact of this may depend on the readiness of learners to shift attention to formulation when performing the task. If they do this, then strategic planning may lead to greater accuracy but if they do not do this no effect will be evident. Thus, the effect of strategic planning on accuracy may be linked to the kind of on-line planning that occurs subsequently during task performance. Clearly, though, more research is needed to identify how planning interacts with task

design variables, implementational procedures and individual learner factors. The variable impact of pre-task planning (rehearsal) on accuracy as a result of the learner's orientation during performance is explored by Bygate and Samuda in Chapter 2.

The results are clearer for **complexity**. As for fluency, strategic planning has a definite, positive effect; planners produce more complex language than non-planners. Crookes (1989) reports that 10 minutes of planning time led to learners producing more complex sentences and a wider range of lexis. Foster and Skehan (1996) found that detailed planners used significantly more subordination than undetailed planners who, in turn, produced significantly more subordination than the non-planners. This was broadly true for all three tasks. Skehan and Foster (1997), however, found that the planners' production was more complex on only two of the tasks. On the narrative task, where planning led to greater accuracy, no effect for complexity was evident, suggesting a trade-off between these two aspects of production. Wendel (1997) found that his planners used more complex grammatical structures but not more lexically rich language. Mehnert (1998) also found a positive effect but only for the 10-minute planners - the 1-minute and 5-minute planners performed at the same level as the non-planners. Ortega (1999) reports that mean number of words per utterance (a complexity measure) was significantly higher in the planning condition. Yuan and Ellis (2003) also found that strategic planning had a positive effect on complexity. With regard to the source of planning, Foster and Skehan (1999) found that individual learner planning worked best for complexity as it did for fluency. Again, in this study, whether the learners focused their planning on form or content had no differential effect on complexity.

These studies indicate that giving learners the opportunity to plan can increase the complexity of their production. They also suggest that this effect can be enhanced if (1) learners have a reasonable length of time to plan, say 10 minutes, (2) they are given guidance in how and what to plan and (3) they plan individually rather than in groups. It should be noted, however, that the measures of complexity used in these studies did not distinguish between propositional complexity (i.e. the content of the learners' messages) and formal complexity (i.e. the actual language used). Here too further research is needed.

What general conclusions are possible from these studies? The first is that strategic planning has a stronger effect on fluency and complexity than accuracy. This suggests that when learners plan strategically they give more attention to drawing up a conceptual plan of what they want to say rather than to formulating detailed linguistic plans. Even when asked to engage in form-focussed planning they may not do so, preferring to use the time given them

to sequence ideas and to work out the semantic linkages among propositions. Alternatively, it is possible that even when learners do attend to form when planning, they find it difficult to carry over the forms they have planned into the performance of the task, as suggested by Bygate and Samuda in Chapter 2. The second conclusion is that trade-off effects are evident. When learners plan they have to choose what aspect of production to focus on; focussing on fluency and complexity is at the expense of accuracy and vice-versa. Finally, there is some evidence to suggest that strategic planning has a greater effect on production in general when the task is cognitively demanding. If a task is easy learners are able to perform it fluently using accurate and complex language without the need for planning.

Unpressured on-line planning

Giving learners time to plan on-line and to monitor their output appears to have a clear impact on accuracy. Hulstijn and Hulstijn (1984) asked learners of L2 Dutch to perform short oral narratives under four conditions involving combinations of two variables; time (i.e. the learners were told to speak as quickly as they could or to take as much time as they wanted) and focal attention (i.e. learners were instructed to focus on form or on meaning). They found that time pressure by itself did not affect the accuracy of word order but that in combination with a focus on form it had a profound effect. This study, then, suggests that when learners use the time at their disposal to attend to formulation and to monitor the use of their grammatical resources their production becomes more accurate. However, if they use the time to plan content no effect on accuracy is observed.

    In Ellis (1987), I compared learners' performance on written and oral narrative tasks based on pictures. In the case of the written task (task 1) the learners were given as much time as they wanted to write the narrative. In the first oral task (task 2) they were asked to retell the same narrative but without recourse to their written versions. In the second oral task (task 3) they were given a different set of pictures and instructed to tell the story with minimal opportunity for prior-planning. Figure 4 summarizes the kinds of planning opportunities afforded by these three tasks. I found that the learners' use of the regular past tense forms (but not the irregular past tense or copula past tense forms) was most accurate in task 1 and least accurate in task 3, with task 2 intermediate. The difference between task 1 and 2 can be explained in terms of on-line planning; accuracy was greater when there was no time pressure. However, as

Task	On-line planning/monitoring	Strategic planning
1	Yes	Yes (Probably)
2	No	Yes
3	No	No

Figure 4. Types of planning opportunities in Ellis (1987)

Crookes (1989) and others have pointed out, tasks 1 and 2 also differed with regard to medium.

Building on Ellis' study, Yuan and Ellis (2003) set out to compare the effects of pre-task and on-line planning on learner performance of a narrative task in a more systematic way. In the pre-task planning condition learners were given 10 minutes to prepare the task and then performed it under time pressure. In the on-line planning condition, the learners were given no chance to prepare but were allowed to perform the task in their own time. There was also a control group that had no preparation time and was required to perform the task under time pressure. The results indicated that opportunities for unpressured on-line planning assisted both accuracy and complexity but, as might be expected, inhibited fluency.

These three studies suggest that the time learners are given for on-line planning improves the accuracy of their production. However, the effects may only be evident when learners are drawing on their rule-based system. In both Hulstijn and Hulstijn (1984) and Ellis (1987) the effects of time pressure were only evident on grammatical structures that are clearly rule-based (i.e. Dutch word order rules and English regular past tense); they were not evident in structures that are more lexical in nature (i.e. irregular and copula past tense forms).

Planning in a language testing context

The study of the effects of planning on the performance of tasks in a testing situation is of considerable importance given that testers in general are concerned to elicit the 'best performance' from a testee (see McNamara 1996). If planning time can affect aspects of a test-taker's performance then arguably it ought to be considered when designing the test.

Three research studies have investigated the effects of pre-task planning in a testing situation. Wigglesworth (1997) examined the performances of 107 adult ESL learners performing five tasks that were part of the Australian Assessment of Communicative Skills (Access) test. The candidates performed the tasks in a planned and unplanned condition. The performances were rated

by two trained raters using an analytic rating scale to measure fluency, grammar (or in one task vocabulary) and intelligibility. The performances of 28 candidates, who were divided into high and low proficiency groups, were transcribed and analyzed using measures of complexity, fluency and accuracy. Wigglesworth reported no significant differences in the rating scores for the planned and unplanned conditions but significant differences in the analytic discourse measures for complexity, fluency and accuracy, especially in the high proficiency candidates and especially in tasks with a high cognitive load. She concludes that at least for some learners and in some tasks planning time can help to improve the performance of test-takers but that this effect is not evident in external ratings.

In a second study, Wigglesworth (2001) sought to further investigate one of the findings of the previous study, namely that the effects of planning time were not evident in the scores obtained from raters. The study examined the effect of a number of test task variables, one of which was planning, on adult ESL learners' performance on five tasks that were routinely used to evaluate achievement in the Australian Adult Migrant Education Program. In this study an effect for planning was found on the test-takers' ratings but the effect was not as great as might have been expected. Planning proved to have a detrimental effect on tasks that were familiar to the candidates and on both structured and unstructured tasks. Wigglesworth notes that these results are inconsistent with the findings of task planning research in non-testing situations and suggests that this may reflect the fact her study used external ratings rather than discourse analytic measures. However, Iwashita, Elder and Mcnamara (2001) used both analytic discourse measures and ratings to examine the effects of three minutes of planning time on the task performance of 201 ESL students and failed to find evidence of any effects on either the discourse measures or the rating scores. Elder and Iwashita reproduce this finding in Chapter 8 and examine a number of possible explanations.

It is possible, then, that the testing context constrains the beneficial effects of planning. This suggests, more generally, that the 'psychological context' of a task constitutes an important dimension that needs to be taken into account in planning studies (see Batstone's discussion of this possibility in Chapter 10). The main conclusion to be drawn from these studies, however, is that there is a need for further research into the effects of planning in a test situation. It seems clear, however, that whatever effect planning time has on task performance it may not be reliably measured by an external rating. This is problematic where assessment is concerned, as it is not practical to calculate discourse analytic measures in testing situations.

Final comments

This review of the research suggests that the effects of planning in a testing context may be somewhat different from those reported for laboratory or classroom contexts. One reason may be that learners feel pressured in a testing context with the result that their on-line planning is hurried. To date no studies have examined whether there are any differences in on-line planning in testing and non-testing contexts. This is a fairly obvious next step.

The results of the research certainly suggest that pre-task and unpressured on-line planning may be somewhat different. Whereas opportunities for on-line planning result in more accurate and complex language use, probably because learners have the chance to monitor linguistic form, opportunities for pre-task planning generally favour fluency and complexity, possibly because it leads to an emphasis on conceptualizing what has to be communicated rather than how to say it.

As I noted in the concluding comments to the previous section, researchers have focussed their attention on investigating how different types of planning (in combination with different types of tasks) impact on learner production. They have not attempted to show how or even whether the planning of tasks assists language acquisition. Thus any claims regarding planning and acquisition can only be theoretically based. Clearly, the absence of empirical support for the key assumptions listed at the end of the previous section constitutes a major lacuna in the research to date.

Methodological issues

The task planning research to date raises a number of methodological issues. Perhaps the key one concerns how acquisition as opposed to language production can be investigated. Other issues are how to ensure that learners carry out the type of planning specified in the research design and how to measure learners' actual production when they perform the task. These issues will be considered below.

Investigating the effects of planning on acquisition

The term 'acquisition' assumes that there is some change in the learner's L2 knowledge representation. Evidence for change can be found in (1) the learner's use of some previously unused linguistic forms, (2) an increase in

the accuracy of some linguistic forms that the learner can already use, (3) the use of some previously used linguistic forms to perform some new linguistic functions or in new linguistic contexts and (4) an increase in fluency (i.e. in the speed with which linguistic forms are used in communication).

The usual method for obtaining these kinds of evidence of change is the standard experimental design involving an experimental group that completes a pre-test, a treatment and post-tests (immediate and delayed) and a control group which receives the tests without the treatment. In the case of task planning research, the treatment consists of the opportunity to plan and perform a task. Such a design, as we have already seen is rarely employed. To the best of my knowledge, the only studies that have made use of such a design are Bygate's (2001) and Gass et al.'s (1999) studies of task rehearsal. Bygate's study asked learners in the experimental groups to perform two tasks prior to the treatment (which in turn consisted of three opportunities to repeat tasks similar to one of the pre-treatment tasks) and the same two tasks following the treatment together with two new tasks of the same type. In this way, Bygate was able to assess to what extent the treatment resulted in changes in the way the learners (1) performed the same task they had completed before the treatment and (2) a similar task to the pre-treatment task. Such a design is promising as it does allow the researcher to pinpoint changes that occur as a result of the treatment. It contrasts with the standard design used in task planning research (see, for example, Foster & Skehan 1996; Yuan & Ellis 2003), which typically involves an experimental and control group performing the same task under different planning conditions (e.g. strategic planning as opposed to no planning). Such a design cannot address acquisition.

There is, however, a major limitation to the kind of design that Bygate employed. It does not provide data that can easily speak to the effects of task planning on the acquisition of <u>specific</u> linguistic features. That is, it can only provide evidence of general linguistic change, as in types (2) and (4), but not of specific linguistic changes, as in types (1) and (3). To obtain evidence of the effects of task planning on specific linguistic features it is necessary to target specific features for study. This cannot be readily achieved by means of the kinds of unfocused tasks that have figured in task planning research to date. However, it may be achievable through the use of focused tasks. Whereas unfocused tasks allow learners to choose from a range of forms focused tasks aim to induce learners to use specific forms. In Skehan's (1998b) terms they are 'structure trapping' in that they make the employment of the specific forms, natural, useful or, ideally, essential (Loschky & Bley Vroman 1993). The advantage of such tasks is that they allow researchers to construct pre- and post-tests to mea-

sure whether learners knew the targeted forms prior to performing the task and what the effect of planning tasks is on learning. The only planning studies that have investigated specific linguistic forms to date are Ellis (1987), which targeted past tense forms, and Hulstijn and Hulstijn (1984), which targeted word order rules in Dutch. Somewhat disappointingly, more recent studies have been based on unfocused tasks.

Investigating learners' planning strategies

In a typical task planning study, learners are asked to carry out planning in accordance with instructions. Below, as an example, is the description of the 'guided planning – content focus' condition in Foster and Skehan (1999):

> The students were introduced to the idea of a balloon debate. The teacher then presented ideas that each character might use to defend his or her right to stay in the balloon and students were encouraged to add ideas of their own.

Here is a description of the unpressured on-line planning condition in Yuan and Ellis (2003):

> The on-line planners were required to tell the story by producing at least four sentences for each of the six pictures after seeing the pictures for only 0.5 seconds. They were given unlimited time to enable them to formulate and monitor their speech plans as they performed the task.

Such instructions raise a number of methodological issues. The most obvious one, given the evidence that pre-task and on-line planning have been hypothesized to have somewhat different effects on learners' performance of a task, is the need to ensure that learners receive instructions relating to <u>both</u> kinds of planning. In the case of studies investigating pre-task planning this has not usually occurred. That is, the learners are given instructions relating to how to conduct strategic planning/rehearsal but are left to perform the actual task in any way they choose. Thus, it is possible that the learners interpret the task performance conditions very differently, with some engaging in unpressured and others in real-time on-line planning. This may be one explanation why studies of pre-task planning have produced such mixed results for accuracy (see previous section).

There is also an obvious methodological need to establish whether learners actually carry out the planning instructions they were given. That is, do they conform to the prescribed planning conditions? Again, few studies have attempted to establish this. However, more recently, a number of researchers have attempted to describe the different strategies learners actually use during

the pre-task planning phase of a study. All three studies in the section dealing with pre-task planning in this book (Section 3) do this. The data used for such an investigation includes the actual notes that learners make while planning strategically (see Ellis & Yuan 2004) and post-task interviews with individual learners (Ortega 1999: Chapter 3 in this book). Such research is important not just to ensure that learners plan as intended but also because it can serve as a basis for drawing up guidelines for the design of effective planning instructions. Sanguran, in Chapter 4, makes a useful advance in this direction by formulating an explicit set of assumptions that guided her in the preparation of the planning instructions she used in her own study.

Somewhat different kinds of evidence are needed to demonstrate what kind of planning – pressured or unpressured – learners engage in on-line. While it may be possible to establish this through post-task interviews (although learners may have difficulty remembering their on-line decisions even if stimulated recall techniques are used), clearer evidence may be forthcoming by inspecting the fluency properties of the texts learners produce as a result of performing the task. Yuan and Ellis (2003) considered two such properties – the number of syllables produced per minute and the number of pruned syllables per minute (i.e. after dysfluencies had been discounted). They were able to show that learners in the unpressured on-line planning condition spoke significantly more slowly than learners in the pressured on-line planning condition. In this way, they were able to demonstrate that the unpressured on-line planners had performed as required.

Measuring learner production

Learner production can be measured either by means of external ratings or by means of discourse analytic measures. In general, language testers have preferred the former and SLA researchers the latter.

External ratings are based on scales that specify (1) the specific competency being measured and (2) levels of performance for each competency (often referred to as 'bands'). In the case of ratings of task-based performance, the target competency can be specified either in behavioural terms that reflect the degree to which the learners have successfully completed the task (see, for example, Norris, Brown & Hudson 2000) or in linguistic terms. In the case of the latter, learners' linguistic competency can be described either holistically (e.g. for the highest 'band' the descriptor might be 'speaking proficiency equivalent to that of an educated native speaker') or an analytic measure, where different dimensions of performance (for example, fluency, complexity and accuracy) are rated

separately. In Ellis (2003: 298–302) I summarise the various options relating to external ratings.

In the case of discourse analytic measures, two types of measures are possible – measures of specific linguistic features and measures of general dimensions of oral and written discourse. There are a variety of well-established methods for deriving measures of specific linguistic features (e.g. error analysis, obligatory occasion analysis, frequency analysis and form-function analysis – see Ellis and Barkhuizen (2004) for a detailed account of these methods as they have been used in SLA). In the main, however, researchers have not used these, preferring instead general measures of learner production.

These general measures have been based on Skehan's model of L2 proficiency, which distinguishes two basic dimensions – meaning (fluency) and form with the latter further sub-divided into complexity and accuracy. Skehan (see Skehan & Foster 1997; Tavokoli and Skehan's study in Chapter 9 in this book) has been at pains to establish the independence of these dimensions by factor analysing scores obtained from a battery of measures. While the analyses do not always produce entirely similar results (e.g. in Skehan & Foster 1997 the analysis resulted in three distinct factors easily identifiable as fluency, complexity and accuracy while in Tavokoli and Skehan the analysis produced a somewhat different set of factors – temporal aspects of fluency, repair fluency and complexity/accuracy combined) they do broadly confirm Skehan's model. Thus, the general measures employed by Skehan and his co-researchers, have an established theoretical base.

There are nevertheless a wide range of measures of fluency, complexity and accuracy to choose from (see Figure 5 for a summary of the various measures employed in the studies reported in the subsequent chapters in this book). In one respect this is useful as, arguably, multiple measures of each dimension may yield a more valid assessment than single measures. The downside is that when researchers differ in their choice of measures it becomes difficult to compare results across studies. Ideally, work is needed to establish measures that provide the most valid assessment of each dimension (using, for example, a factor analytic approach such as that employed by Skehan), which can then be employed across studies. It is also worth noting that it may prove necessary to develop separate measures for spoken and written production, most obviously for fluency. Most of the measures used to date have been developed for oral production, as this has been the focus of the bulk of the planning studies. However, Ellis and Yuan (2004) developed measures of written production and Wolfe-Quintero et al. (1998) offer a comprehensive list of measures of all three dimensions for writing.

Type of measure	Specific measure	Description	Study
1. Fluency	Production rate	The number of syllables produced per minute of speech/ writing	Ellis and Yuan; Kawauchi; Elder and Iwashita; Sanguran
	Breakdown fluency	The ratio between number of words reformulated and total words produced	Ellis and Yuan
		Number of repetitions	Kawauchi; Elder and Iwashita
		Total silence Number of pauses greater than 1 second Number of filled pauses Length of run	Skehan and Foster; Tavakoli and Skehan
2. Complexity	Syntactic complexity	Ratio of clauses to some general unit (e.g. T-units, c-units or AS-units)	Ellis and Yuan; Kawauchi; Elder and Iwashita; Sanguran; Skehan and Foster; Tavakoli and Skehan
		Length of unit (e.g. T-unit)	Kawauchi
		Number of subordinate clauses	Kawauchi
	Complex grammatical structures	Use of comparatives and conditionals	Sanguran
	Syntactic variety	Total number of different grammatical verb forms used in the task	Ellis and Yuan
	Lexical variety	Mean segmental type/token ratio	Ellis and Yuan
3. Accuracy	Overall grammatical accuracy	Error-free clauses	Ellis and Yuan; Elder and Iwashita; Skehan and Foster; Tavakoli and Skehan
		Error-free clauses of different lengths	Skehan and Foster
		Number of errors per 100 words	Sanguran
	System-based grammatical accuracy	Correct verb forms	Ellis and Yuan
		Past-tense markers	Kawauchi

Figure 5. Discourse analytic measures used in the studies reported in this book

A final question concerns the length of the learner texts to which the measures are applied. In many cases, researchers do not use the full texts produced by learners but instead elect to use only part of the texts, typically the first five or ten minutes. The problem here, as Skehan and Foster's chapter in this book indicates, is that planning may have a markedly different effect on the first few minutes of production in comparison with later. Learners may have difficulty sustaining careful formulation and monitoring over a lengthier period of time. Skehan and Foster's study raises the awkward possibility that the findings of the research to date, which have typically been based on relatively short learner productions may not be generalizable to extended discourse.

Conclusion

Task planning has proven a rich vein for empirical study, as attested by the large number of studies that have investigated this implementational variable (larger than have investigated any other task variable) and by the current collection of studies. Why has task planning proven such a fruitful arena for SLA research? Is it just another fad in SLA, like the error evaluation studies in the 70s and 80s, that will soon lose its attraction? I think not. First, the study of task planning, as I have tried to show in this chapter has a strong theoretical basis drawing on a set of constructs (controlled processing, limited capacity memory, focus-on-form) and a number of well-established theories of L2 use and acquisition. Research, such as that reported in the subsequent chapters of this book, can both draw on this theory and help to test it. In a sense, then, the study of task planning lies at the very centre of current research in SLA. Second, the research is of obvious pedagogical relevance. Planning, whether of the pre-task or within-task kind, is a variable that teachers can easily manipulate in their day-to-day teaching. While teachers should not look to research as the only determinant of lesson design they can certainly benefit from the insights and 'provisional specifications' (Stenhouse 1975) that the task planning research offers them. Thus, for both theoretical and practical reasons I expect task planning to continue to attract attention in the years ahead.

This book constitutes an advance on the research to date. It addresses a variety of issues, some previously examined, others new:

– the role of task rehearsal in helping learners to elaborate content and to integrate the different strands of their L2 proficiency;
– the actual strategies learners employ during pre-task planning;

- the way in which learners orientate to the opportunity to engage in strategic planning;
- the extent to which learners' attention to form and meaning can be manipulated through pre-task planning;
- the effect of different types of planning (pre-task vs. on-line; detailed vs. undetailed);
- the interaction between strategic planning (a task implementation variable) and task design features (such as the introduction of a surprise element into a task);
- the effects of learners' L2 proficiency on their ability to make use of the opportunity for pre-task planning;
- the relative effects of unpressured on-line planning on oral and written production in an L2;
- the extent to which learners are able to sustain the effects of planning on performance over an extended period of time;
- the effect of context (e.g. a language test) on task performance subsequent to planning;

The range and variety of these issues testify to the richness of task planning as an area of SLA enquiry.

Section II

Task rehearsal

The chapter in this section examines the effects on task-performance of having learners repeat a task – of what was called 'rehearsal' in Chapter 1. Bygate and Samuda's paper is important both methodologically, theoretically and pedagogically.

As noted in Chapter 1, the bulk of the research that has investigated the effects of planning on task performance has examined learner productions in terms of fluency, accuracy and complexity. There is an obvious need to extend analysis to the macro properties of learner discourse. Bygate and Samuda show that one way of doing this is by examining what they call 'framing'. This is a cover term for a heterogeneous collection of linguistic resources used by speakers to convey 'perspective' (e.g. the speaker's attitude to what is being communicated) and to 'preview' (e.g. by providing an advance organizer of what is to come). In effect, framing fleshes out the bare factual bones of a discourse. The analysis of learner narratives they present in terms of framing demonstrates that this constitutes a significant addition to the tools in current use. Bygate and Samuda's analysis also points to the value of combining group-based statistical analysis with a qualitative, case study approach.

Their chapter is important theoretically because it provides a thoughtful account of how different kinds of planning (strategic planning, on-line planning and rehearsal) contribute to task performance. Bygate and Samuda argue that rehearsal offers the learners certain processing opportunities not available in the other types of planning, in particular the ability to integrate their linguistic resources. Repeating a task enables learners to reorganise and consolidate information into a richer, discursally more sophisticated performance.

Finally, Bygate and Samuda suggest that rehearsal is a useful pedagogic procedure not just because of the opportunities it affords learners to develop their L2 discourse skills but also because rehearsal arises in naturally occurring communicative activities (i.e. it has situational authenticity). The challenge facing teachers is to introduce task repetition in ways that students will find motivating.