Chapter Two

Task Complexity and Linguistic Performance: Models, Features, and Assessment Measures

2.0 Introduction

We have seen in Chapter One, that to learn a second language, students are expected to perform on tasks which are carefully designed in a task-based syllabus. The performance on these tasks requires target language use and therefore generates possibilities for interlanguage development and L2 learning. In the event of using tasks to push interlanguage boundaries, they need to be graded and sequenced.

This raises the question, how should tasks be graded and sequenced? The primary importance of this thesis is to highlight aspects like task characteristics that influence the cognitive processes underlying L2 task performance and how manipulation of tasks may give the most beneficial output with respect to L2 development. With respect to these questions, two models of task sequencing have largely been proposed which will be discussed in Section 2.1: Skehan and Foster’s Limited Attentional Capacity Model (Skehan 1998, 2001, 2003; Skehan & Foster 1999, 2001) and Robinson’s Cognition Hypothesis (2001a, 2001b, 2005, 2007). Section 2.2 presents Robinson’s Cognition Hypothesis in detail. The focus in this chapter will be on the predictions of the two models with respect to effects of cognitive task complexity and effects of interaction on L2 learners’ oral and written task performance. In order to discuss the various studies
conducted, the researcher starts with a description of measures that have typically been used to assess learner performance (§2.3). Various studies and their results in task complexity dimension have also been discussed in Section 2.4. In-depth discussion on the present study and the research questions are given in Section 2.5.

2.1 Models of task complexity

The cognitive strand of task-based research explained in Chapter One (Section 1.1.4) aims at understanding the cognitive and attentional processes during L2 task performance. Chapter One investigates how manipulations of different task characteristics may affect attentional allocation: this is important because only input that is noticed and attended to, can be used for furthering L2 development (Schmidt 1990). Research has delved on how the cognitive load of task (cognitive task complexity) determines the quantum of attentional resources drawn by task characteristics; and how this attention allocation affects task-based performance (Robinson 1995b, 2001b, 2005, Skehan 1996, Skehan & Foster 2001).

Within the cognitive strand, there exist contrasting hypotheses concerning factors of cognitive task complexity (and their predicted affects on task performance). Robinson’s (2005) theory, known as the Cognition Hypothesis, is based on the idea of multiple attentional resources. As the research presented in this thesis takes Robinson’s claims as the object of investigation, therefore, a detailed description of the Cognition Hypothesis is presented (§2.2). The following section 2.1.1 will discuss an alternative cognitive
account on task-based L2 performance that advocates the idea of limited attentional capacity (Skehan 1996, Skehan & Foster 2001).

2.1.1 Limited Attentional Capacity Model

The basic assumption of the Limited Attentional Capacity Model, as the name suggests, is that attentional capacity is limited and that there is some maximum limit in the amount of information one can keep active or pay attention to. This view of Limited Capacity of attentional resources is founded on theories of working memory (Carter 1998; Baddeley 2000; Gathercole & Baddeley 1993).

To recall from Chapter One, the processes involved in a task are input selection, goal oriented information processing, and response actions (Baddeley 2003) and the attentional resources have to be shared by all the processes the task calls for. If the task demands exceed more than the total capacity of attentional resources, the allocation of attentional resources will be in competition, and some aspects of the task will attract more attention and others less attention. The decision of which task gets more attention and which task gets less attention, depends on the nature of the task and the output requirements.

The central claim of Skehan and Foster’s Limited Attentional Capacity Model (1996) is that during task-based L2 performance, different areas of performance—linguistic complexity, accuracy, and fluency—compete with one another for the attentional resources that are available (Skehan & Foster 2001:205). As a result, only those aspects
that receive adequate attention will reach optimal performance and those aspects that receive no or limited attention will fail or become erroneous (such that overall performance will decline). The different dimensions of task performance come into competition with each other, and are likely to generate trade-off effects between the dimensions of performance.

In a complex task, learners attend to one dimension of performance i.e. either complexity, or accuracy or fluency (of language production) and that attending to one aspect may mean that other dimensions are not focused on. Attention to one aspect is therefore at the expense of the others. Skehan and Foster (1997) argue that on encountering a complex task, learner’s attentional limits are reached and this leads learners to prioritize meaning over language form. This is because it is more important for L2 learners to meet the communicative goal of the task (VanPatten 1990). The implication of this model is that in a task-based syllabus, tasks should not have high cognitive load, so that it gives rise to competition in the allocation of attentional resources, thereby, forcing learners to struggle with meaning and neglect formal aspects of the task altogether. To explain how this competition for attention occurs between complexity, accuracy and fluency, Skehan and Foster propose a step-wise generation of competition between different formal aspects: first a competition is seen between fluency and form, i.e. when required to present ideas in a consistent manner learners ignore formal aspects of complexity and accuracy. Within form-based aspects, learners tend to pay more attention to complexity than to form. This trade-off between complexity and accuracy is the most prominent capacity limitation this model talks about.
What happens when accuracy and complexity are in competition? Skehan and Foster (1997) put forward two possible consequences of this competition. First, the accuracy either shows substantial increase or remains constant, while complexity considerably drops. Second, complexity increases and accuracy declines. The former happens when the learner has stable language proficiency and has adequate control over his/her interlanguage, and therefore uses a ‘safety first’ approach, and avoids using words and structures which the learners are not sure of. The performance is therefore linguistically less complex but more accurate because the learner relies on simpler and known language forms. The latter happens when the learner tries to use complex structures and varied vocabulary, and therefore makes more mistakes (thereby reducing accuracy). Learners will produce more errors (referred to as the ‘accuracy last’ approach). Skehan and Foster (2001) present a model that illustrates these trade-offs during L2-language production, which Michel (2011) adapts in a diagrammatic representation as presented in Figure 3.

**Figure 3. Trade-off effects in Skehan and Foster’s model**

![Diagram from Michel (2011:11)](taken-from-michel-2011.png)
In contrast to what we show for complex tasks, the Limited Attentional Capacity Model predicts that cognitively simpler tasks are more likely to let learners focus on both linguistically complex and accurate performance in parallel, because the demands do not exceed limitations of their attentional capacity. Cognitively complex tasks will inevitably yield an L2 performance of lower linguistic quality than cognitively simple tasks because attention is a resource of limited capacity.

2.1.2 Skehan’s model of task difficulty/complexity


In the first area, code complexity, language is concerned with the linguistic demands of the task and refers to linguistic complexity/variety, vocabulary load/variety, and redundancy/density. Tasks which require more advanced structures, or which require a greater variety of structures, or greater densities of advanced structures, such as complex tenses or subordination or embedding, are likely to be more complex. According to Skehan (1996), complex tasks should yield greater lexical variety and diversity in the output.

Cognitive complexity is concerned with the task content and its manipulation. Problem solving tasks will involve more processing demands (such as narrating an incident/story). Here, the way the input is structured is also very important. Understanding a story with a
proper beginning, middle and end requires less processing effort than a story alternating between the present and the past. Task participant’s schema or the essential knowledge required to perform a task also determines the complexity of the task. For example, a task involving narration of one’s experiences (easy) as compared to narration of an imaginary situation (difficult). The third factor that influences complexity is: learner’s familiarity with the task, aspects like whether the task has been already performed and whether learners know what is expected of them. Also important is the input or the amount of information provided for the task.

The third area, performance conditions, is concerned with a task’s communicative stress that can affect performance. These are time allotted for the task (time pressure), the modality (listening, speaking, reading, writing or a combination of four skills), the nature of participation (group, pair or individual task), the nature of assessment (graded or non-graded), the length of the task, and the possibility of control and manipulation.

It is possible that all these factors have some influence on learner’s attention during a task and is likely to influence the performance. According to Skehan (2001), the amount of attention a task demands determines the task difficulty. Difficult tasks require more attention, and in task performance, content is prioritized over form (VanPatten 1990), so difficult content will require a lot of attention and only what is left can be devoted to language form.
Skehan (1998) reasons that inadequate attention to form will lead to non-fluent or inaccurate or non-complex language during a task. According to Skehan, a task involving a lot of complexity undertaken in conditions of extreme communicative stress, may lead learners to slow down (fluency), forget accuracy, use simple language (complexity) or adopt all the three strategies (Robinson 2001:196). Skehan believes that complexity and accuracy are in competition for attention and that they are inversely proportional to each other during the performance of a complex task.

**Figure 4. Skehan’s model of Task difficulty**

1. **Code complexity**
   - Linguistic complexity and variety
   - Vocabulary load and variety

2. **Cognitive complexity**
   - Familiarity of topic
   - Familiarity of discourse genre
   - Familiarity of task
   - **cognitive processing**
     - Information organization
     - Amount of ‘computation’
     - Clarity of information given
     - Sufficiency of information given

3. **Communicative stress**
   - Time pressure
   - Scale
     - Number of parts
     - Length of text used
   - Modality
   - Stakes
   - Opportunity for control

Taken from Skehan (1998:194-195).
In contrast to the Limited Attentional Capacity Model, the *Cognition Hypothesis* by Robinson (2005) claims that L2 learners can rely on multiple attentional pools during task performance such that complex tasks may not induce trade-off effects. As the Cognition Hypothesis is the theoretical framework under investigation in the present chapter, the following section will give a detailed description of Robinson’s ideas.

### 2.2 The Cognition Hypothesis

Robinson’s *Cognition Hypothesis* (Robinson 1995a, 1995b, 2001a, 2001b, 2003a, 2003b, 2005, 2007a, 2007b; Robinson, Cadierno & Shirai 2009; Robinson & Gilabert 2007) is in sharp contrast to the Skehan and Foster’s Limited Attention Capacity Model, and therefore has different implications for task sequencing and L2 development.

Robinson with his strong interest in the effects of inherent task characteristics on learners’ language production, argued that when tasks are cognitively and functionally demanding or difficult, learners will be encouraged to produce more complex and more accurate language production. This was known as the *Cognition Hypothesis*. In the following section, we elaborate Robinson’s view on attention, to what aspect of language (complexity, accuracy or fluency) attention is allocated in a cognitively complex task, and finally we present Robinson’s (2005) taxonomy of factors that influence task performance (the so-called ‘Triadic Componential Framework’).
2.2.1 Robinson’s view on attention and language learning

According to the Cognition Hypothesis, attention is crucial for processing linguistic information. Robinson’s perspective is based on Cowan’s (1988, 1993) idea of hierarchical relationship between memory and attention and he extends Cowans’ idea by highlighting two attentional processes that are important for SLA: detection and noticing. Detection refers to an unconscious automatic recognition process in short-term memory while noticing requires focused controlled attention. When some stimuli is noticed and receives focal attention, it is processed in working memory (Schmidt 1990).

Robinson links the notion of ‘data driven’ and ‘conceptually driven’ processes of learning to the notion of attention allocation in tasks. Robinson describes bottom-up, automatic processes that are activated by the data itself as data driven. In contrast to data driven, conceptually driven processes are participant initiated, top-down, and attentionally controlled. What controls the process of attention selection is the task itself and its characteristics. Attention could be drawn to task features in a top-down manner by the task instruction, prompting learners to pay attention to different segments of the task, or by endogenous cueing guided by the task performer’s aim to solve the task in the best possible way. On the other hand, data driven processes triggered by task inherent characteristics may draw the attention on the use of, e.g., specific linguistic structures and lexical forms.

Only when information is noticed (either consciously or unconsciously) it is learnt.
Therefore, what is focused on in a task determines what can be learnt from a task. Robinson’s Cognition Hypothesis was partially based on Cromer’s (1974) Cognition Hypothesis for First Language Acquisition (FLA). Cromer (1974: 246) argued that in FLA “our cognitive abilities at different stages of development make certain meanings available for expression” and therefore, “it is cognition which determines language acquisition”. Robinson (2003b) claimed that this argument is applicable to adult language learning. Although adults have already developed complex notions of the world, they have not yet acquired enough linguistic knowledge to express them early in their L2 learning. Therefore, Robinson (1995, 2005, 2007) claims that complex notions and high functional demands will lead adult language learners to develop or stretch their interlanguage so that they can meet the increased demands of the task and express elaborated ideas.

As for accuracy development, Robinson (2005: 9) argued that cognitively difficult tasks are “likely to draw learner attention to the ways in which the L1 and the L2 may differentially grammaticize conceptual notions, and so have positive effects on L2 accuracy of production”. Difficult tasks, therefore, are hypothesized to offer better chances for accuracy development than are simpler tasks.

Robinson, therefore, viewed syntactic complexity and accuracy as arising from functional complexity in discourse and, hence, increased functional demands imposed by the task should have detectable linguistic consequences. In this model, concurrent attention to different aspects of L2 use was considered not just possible, but natural.
2.2.2 Theory of Multiple Resources of attention and interference

The Cognition Hypothesis is underpinned by the idea that individuals have multiple-resource pools of attention (Neumann 1996; Wickens 1991, 2002, 2007) and their attentional capacity is unlimited (the interference model: Navon 1989; Navon & Gopher 1979). Recall that Wickens argued that when completing different tasks, people draw their attention from different resource pools depending on processing mechanisms (i.e., encoding or responding), codes (i.e., spatial or verbal), modalities (i.e., visual or auditory), or responses (i.e., manual or vocal) that each task requires. It is claimed that a competition for attention occurs not between but within the resource pools.

The other model, which Cognition Hypothesis draws on, is called the interference model. Researchers who advocate this model, such as Gopher (1993) and Sanders (1998), argue that it is not limitation of capacity but the limited time available to complete a task or confusion that hinders processing of multiple sets of information. In other words, the central executive of working memory, which is responsible for attention allocation, easily loses its control when it has to deal with multiple stimuli within the limited amount of time. The attentional capacity is essentially unlimited and performance problems occur due to involuntary attentional shifts that result from a loss of control over attentional allocation. The classic Stroop test is a good example, where automatic processes of reading words actually interfere with the controlled processes of naming the colours in which words are written. Robinson combined these two models and argued that individuals have multiple-resource pools of attention and the amount of attention within each pool is unlimited.
The fundamental claim of Robinson (2005) is that not every increase in cognitive task complexity generates interference. Various task demands may draw on different pools of attention and therefore increasing the cognitive task complexity does not inevitably harm overall task performance. Integrating information-processing theories (Schmidt 2001), interactionist explanations of L2 task effects (Long 1996) and psychological models such as Wickens’ model of dual task performance (Wickens 1989, 1992), Robinson proposes that increases in the cognitive demands of tasks may direct the learners’ attentional resources to language form, and input may be processed more deeply and elaborately (Gilabert 2007).

Robinson predicts that if dimensions of cognitive task complexity belong to different attentional resource pools, increases in task complexity along the so-called resource-directing variables do not cause decrease in linguistic output, but may rather lead to greater structural complexity and accuracy of learner output. Increasing task complexity along the resource-dispersing dimensions, however, depletes the attention available for the task over many specific linguistic aspects of production.

2.2.3 Task features
The Cognition Hypothesis argues that some task characteristics may require a more elaborate use of linguistic structures and forms than others. The characteristics of the task may attract the L2 learner’s attention and focus it to form. Therefore, in task design it is important to assess which task characteristics draw attention to itself (either overtly or covertly).
Robinson in his Triadic Componential Framework (2003) distinguishes between task complexity, task conditions and task difficulty. According to Robinson, task complexity refers to “the intrinsic cognitive demands of the task” (2003: 55), and can be modified in the way tasks are designed. Task complexity is an inherent feature of the task and is the cause for variation ‘within’ participants. Task condition deals with the distribution of information and the nature of participation required in performing the task. Task difficulty depends on what the learner brings to the task. Task difficulty is seen as learner’s perception of the task rather than the intrinsic cognitive demands made by the task which is not learner dependent and also affects learner’s performance (and learning). Therefore, task difficulty accounts for variations in performance ‘between’ participants.

2.2.3.1 Task complexity

The triadic framework (Figure 5) distinguishes between two dimensions of task complexity: ‘resource-directing’ dimensions, and ‘resource-dispersion’ dimensions. Resource-directing dimensions are those in which the demands on language use made by increases in task complexity can be met by manipulating the manner (directing their resources) in which the information is presented. Tasks with more complex demands by means of a resource-directing factor focus the L2 learner’s attention towards the linguistic form because the cognitively complex conceptual and performative demands require complex linguistic means. Accordingly, resource-directing cognitively complex tasks may result in an L2 output that shows a higher structural and lexical complexity because of the increased cognitive demands. Since formal aspects will receive attention, consciously or unconsciously, accuracy will also be pushed up. In Figure 5, we find three
resource-directing factors: ± here-and-now, ± few elements, and ± no reasoning demands. For example, a task that takes place in the there-and-then rather than in the here-and-now will lead to more complex use of vocabulary (e.g., references to time) and structures (e.g., past tense). A resource-directing task with many elements rather than few elements is expected to ask for a more specific lexis and to induce more complex syntactic structures because all the different elements need to be named and distinguished. Similarly, complex reasoning tasks will generate more complex language because the line of argumentation may be lexically marked (e.g., by the verbs ‘claim’, ‘propose’, ‘argue’), and syntactically expressed by means of complex sentence structures of argumentation (e.g., ‘if . . . then’ clauses). Simple tasks, that do not involve reasoning, can stick to simpler structures (e.g., clauses coordinated by ‘and then . . ’).

However, in Robinson’s model, complexity and accuracy do not compete with each other in a complex task, only fluency (in an oral task) may be expected to suffer from increased cognitive task demands, i.e. complex tasks may induce slower speech with more hesitations, pauses and self-repairs, because fluency is of a more performative nature and therefore may require more controlled processing (Levél 1989; Riggenbach 2000).
On the other hand, tasks can increase in complexity along ‘resource-dispersion’ dimensions and make extra resource demands which cannot be met by using any specific linguistic system. As seen in Figure 4, a higher cognitive demand by means of the factors ± planning time, ± prior knowledge, and ± single task will affect learner speech negatively. For example, research gathered in Ellis (2005) has shown that in a task with planning time, learners can conceptualize their propositional message prior to the actual speaking act. In a no-planning-time condition, learners will need attentional capacity for the conceptualization of the message while they are formulating it (Levelt 1989).

Increasing complexity along these resource-dispersion dimensions replicates the processing conditions under which real time language is often used. Practice along these
dimensions could be argued to facilitate real-time access to an already established and developing repertoire of language. Not providing prior knowledge or giving participants dual rather than single tasks is expected to equally disperse task performer’s attention. The resulting loss of control over attentional allocation creates an inefficient and effortful processing of information. In the end, linguistic complexity, accuracy, and fluency probably all suffer from increased cognitive task complexity on a resource-dispersion factor.

2.2.3.2 Task conditions

Just as for cognitive factors of task complexity, the Cognition Hypothesis formulates predictions with respect to how different interactive factors of task condition may affect task-based L2 performance. Robinson (2005) distinguishes two types of interactive factors: (a) participation and (b) participant variables. The latter can be seen as grouping variables based on gender, familiarity, and power relations that may or may not be shared between participants.

The first category, participation variables, is specified in terms of information-flow (e.g., one–way / two–way) and task outcomes (open / closed; convergent / divergent). The distinction one–way / two–way refers to the exchange of that information in a task (Long 1990). A classical way to manipulate this factor is the use of a split information task. If one participant has to describe a picture to an interlocutor and another participant holds only the role of speaker (the information giver), the task is characterized one–way while if there is an active exchange of information, and the participant holds both the roles of
the speaker and the hearer, the task is characterized as two-way. The latter task requires more negotiation than one–way tasks (Long 1990: 41), and therefore helps in focusing the L2 learners’ attention to form.

The distinction between open/closed tasks is based on the goal of the task. Open tasks have little or no fixed outcome and rather have a range of possible solutions, e.g. giving one’s opinion, while closed tasks have a predetermined end (or a restricted number of correct answers) that the task participants work toward. The latter task requires more negotiation (Loschky 1989; Long 1990) and therefore facilitates comprehension and promotes focus on form of input or output (Loschky & Bley-Vroman 1993).

Convergent/ divergent tasks (Duff 1986) differ in terms of the goal orientation learners are expected to have. Convergent tasks require learners to work together or come to an agreement to reach a common solution, e.g. deciding on the best candidate from a list of candidates for a job whereas divergent tasks require them to defend opposing views, e.g., a debate on what is more important, job satisfaction or money, where learners were assigned different viewpoints on the issue and they had to defend their position and refute their partner’s stance. Convergent tasks require more negotiation but divergent tasks require more explanation and justification, which in turn will lead to more complex language (Duff 1986).

Interactive factors determine the amount and nature of interaction that a certain task condition will promote. For example, closed and divergent task outcomes are related to more interaction than tasks with an open or convergent outcome (Ellis 2000).
2.2.3.3 Task difficulty

According to Robinson (2005) task difficulty is the perceived amount of cognitive effort needed to perform a task (see Figure 4). Robinson distinguishes (a) affective variables (e.g., anxiety and motivation) from (b) ability variables (e.g., working memory capacity and aptitude). In combination with the interactive factors of task condition under which a task is performed and the task inherent cognitive factors of complexity, learner factors determine how difficult a task was perceived to be by a learner. As such, the difficulty of a task of a stable cognitive complexity may vary depending on the task performer and the task conditions. Task difficulty is defined by three factors—learner, setting and task. According to Robinson, both task difficulty and task condition cannot form a basis for sequencing as they are largely learner dependent and can be determined only during rather than before the course (Robinson 2001).

Before discussing the studies in examining the relationship between task complexity and performance, it would be useful to discuss the different measures used for analysis of performance and production. Studies assessing the effect of task complexity on accuracy, fluency and linguistic complexity, and use of some variation of the measures are given in Section 2.4.
2.3 Measures of linguistic performance

Researchers such as Skehan (1998) and Robinson (2001) suggest that learners can have different goals while performing a second language task. Sometimes the learners may focus primarily on accuracy, sometimes on complexity and other times on fluency. Before discussing how each of these (accuracy, complexity and fluency) has been used in various studies, let us briefly define these three dimensions of performance.

**Accuracy** refers to “how well the target language is produced in relation to the rule system of the target language” (Skehan 1996: 23). A learner focusing on accuracy tries to achieve greater control over more stable interlanguage elements, the elements which they have already fully internalized. In this case, the emphasis is on avoiding areas where errors may occur. This leads to decreased risk taking (in language use) and increases relying on a familiar linguistic repertoire.

**Complexity** concerns the learner’s willingness to use more challenging and difficult language (Skehan 2001). For a second language learner, structures that have not become fully automated, i.e. the structures at the upper limit of their interlanguage resources, can be considered more complex than language that has been fully internalized. It also reflects hypothesis testing with recently acquired structures. Complexity may also mean that a learner is prepared to use a wider range of different structures.

**Fluency** is the production of language in real time without undue pausing or hesitation (Ellis 2005). Fluency occurs when learners prioritize meaning over form in order to
perform a task. Comprehension and communicative strategies effective in accomplishing the task will be used. No hypothesis testing or risk taking is involved while prioritizing fluency/meaning (Skehan 2001).

Several measures have been proposed in SLA research in order to assess linguistic performance. What these measures have in common is that they provide information in global, general terms, like the total number of words and the number of clauses or errors per T-unit, without further specifying the nature of the words, clauses or errors involved (Hunt 1970). However, it may well be the case that in some circumstances measures of a more specific character are to be preferred.

The accuracy, complexity and fluency of learner language of both oral and written production can be measured. Most of the specific measures of accuracy and complexity that have been developed can be applied to both skills (speaking and writing). However, fluency needs to be measured differently for written and oral language. As this research specially deals with written language, so this chapter elaborates different ways of measuring accuracy, complexity and fluency of learner production. We would not be analyzing fluency in this study because there are no measures to analyze fluency for written text. The only measure used to analyze fluency of writing, speech/ written rate can also not be used because the learners were not given any stipulated time to finish each task in the present study.
2.3.1 Accuracy

A number of different measures have been used by researchers to measure accuracy. A summary of these measures is given in Table 2.1. We discuss all these measures briefly.

(a) *Number of self-corrections*: This measure of accuracy is a count of self-corrections that learners make while speaking. The assumption is that learners’ self-corrections reveal their orientation towards accuracy. Learners who make few errors will have little need for self-correction. For this reason, this measure ideally needs to be calculated in relation to the number of errors a learner makes. Count of self-corrections with relation to errors made, is often used to assess fluency.

(b) *Percentage error-free clauses*: In this measure, the percentage of clauses which have no errors in them are calculated. The most important problem with this measure is to decide what constitutes a ‘clause’.

(c) *Errors per 100 words*: An alternative to the second measure is counting the number of errors per hundred words in the oral/written text. A common problem with both (b) and (c) is determining what constitutes an error. Also the nature of error is rarely taken into account.

(d) *Target-like verbal morphology and target-like use of plurals*: Measures that target specific features of language such as agreement/tense morphology (the number of correct finite verb phrases divided by the total number of verb phrases multiplied by
100, used by Wigglesworth 1997) and plural morphology (the number of correctly used plurals divided by the number of obligatory occasions for plurals multiplied by 100, employed by Crookes 1989) have also been used. However certain specific grammatical features are best used along with a more general measure (Ellis & Barkhuizen 2005).

(e) Target-like use of vocabulary: It can be calculated by calculating the number of lexical errors divided by the total number of words in the text (only correct words). The lower the score, the more accurate the learner’s use of lexis. This measured lexical errors. However, target-like use of vocabulary is a specific grammatical measure and is best used along with a more general measure (Ellis & Barkhuizen 2005).

<table>
<thead>
<tr>
<th>Central focus of calculation</th>
<th>Measures</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of self-corrections</td>
<td>The number of self-corrections as a percentage of the total number of errors committed</td>
<td>Wigglesworth 1997</td>
</tr>
<tr>
<td>Percentage of error-free clauses</td>
<td>The number of error-free clauses divided by the total number of independent clauses, sub-causal units and subordinate clauses multiplied by 100</td>
<td>Foster &amp; Skehan 1996; Robinson 2001; Michel, Kuiken &amp; Vedder 2007; Kuiken &amp; Vedder 2007</td>
</tr>
<tr>
<td>Errors per 100 words</td>
<td>The number of errors divided by the total number of words produced divided by 100</td>
<td>Mehnert 1998</td>
</tr>
<tr>
<td>Percentage of target-like verbal morphology</td>
<td>The number of correct finite verb phrases divided by the total number of verb phrases multiplied by 100</td>
<td>Wigglesworth 1997</td>
</tr>
<tr>
<td>Percentage of target-like use of plurals</td>
<td>The number of correctly used plurals divided by the number of obligatory occasions for plurals multiplied by 100</td>
<td>Crookes 1989</td>
</tr>
<tr>
<td>Target-like use of vocabulary</td>
<td>The number of lexical errors divided by the total number of words in the text (excluding dysfluencies)</td>
<td>Skehan &amp; Foster 1997</td>
</tr>
</tbody>
</table>
2.3.2 Syntactic Complexity

More generalized measures have been used to access complexity. These measures are discussed below.

(a) *The T-unit analysis:* Robinson et al. (1995) and Crookes (1989) used one main clause plus whatever subordinate clauses happen to be attached or embedded within it (Hunt 1965). All independent clauses in isolation (e.g., *John went to the market*) and a clause with one main and one subordinate clause (e.g., *John went to the market to get a pair of spectacles*) are treated as one T-unit. In a coordinated structure (e.g., *John went to the market and then he went to the zoo*) each clause is treated as a separate T-unit. This measure has been considered to be good for written language though not necessarily for spoken language. In spoken language, pausing/ hesitations and intonational patterns suggest that speakers may produce utterances that are longer than a single clause. Furthermore, speech is characterized by elliptical or minor utterances, which would not be considered as independent T-units.

(b) *The C-unit analysis:* Brock (1986) defined C-unit as an independent utterance (for e.g., *Go left/turn right/ Yes/ you will find the building right next to the hospital*) providing referential or pragmatic meaning, allows for ellipsis and serves as a better measure to analyze spoken language. Expressions like ‘Yes’, ‘Thank you’, ‘Now here’, etc. are all C-units.
(c) MTUL: Mean T-unit Length or MTUL is defined as the total number of T-units in a text divided by the total number of words in a text (Hunt 1965). MTUL is a measure of syntactic complexity of a text. This was a measure used to factor in the growth in functional categories in the speech of children and is widely used in child language acquisition research.

Three measures have proven to increase linearly with respect to proficiency level. These are: the number of clauses per T-unit, the number of dependent clauses per T-unit and the number of dependent clauses per total number of clauses.

Table 2.2: Frequently used measures of Syntactic complexity

<table>
<thead>
<tr>
<th>Central focus of calculation</th>
<th>Measures</th>
<th>Studies</th>
</tr>
</thead>
</table>
| Length (in words, morphemes, characters, etc.) | Mean length of utterance  
Mean length of T-unit  
Mean length of C-unit  
Mean length of clause | Brown 1973  
Hunt 1965  
Loban 1976  
Scott 1988 |
| Amount of subordination   | Mean number of clauses per T-unit  
Mean number of clauses per C-unit  
Mean number of clauses per AS-unit\(^2\)  
Mean number of dependent or subordinate clauses per total clauses | Elder & Iwashita 2005  
Skehan & Foster 2005  
Michel et al. 2007 |
| Amount of coordination   | Coordination Index                                                      | Bardovi-Harlig 1992                             |
| Total frequency of use of certain forms considered to be sophisticated | Raw tallies of certain verbal morphology (e.g. tensed forms, passive voice), classes of verbs (e.g. imperatives, auxiliaries, conditionals, modals), syntactic structures (e.g. comparatives, infinitival sentences, conjoined clauses, wh-clauses), etc. | Ellis & Yuan 2005; Robinson 2007 |
| Syntactic variety        | Total number of different grammatical verb forms used in the task       | Ellis & Yuan 2005                               |

Adapted from Norris & Ortega (2009).

\(^2\)AS-unit: Analysis of Speech unit. Single speaker’s utterance consisting of an independent clause or sub-causal unit, together with any subordinate clause associated with either (Foster et. al, 2000:365)
Syntactic variety: This is a measure that we have developed in analogy with lexical variety. The underlying assumption is that proficiency or language production can be gauged by the variety of syntactic structures that a learner controls. In this study, we measured syntactic variety using a type-token ratio. To arrive at syntactic structures, we used Quirk et al. (1995) classification of verb patterns for which the type and token was computed in learner writing. No hierarchy in learning of verb patterns was pre-supposed as only the types of verb patterns was the focus of analysis. To rule out the possibility of the script biasing the ratio, a Guiraud’s index (different types divided by twice the square root of tokens) was used to calculate Type Token Ratio.

2.3.3 Lexical complexity

(a) Type-token ratio: The traditional lexical diversity measure is the ratio of different words (types) to the total number of words (tokens), the so-called type-token ratio, or TTR (e.g., Lieven 1978; Bates, Bretherton & Snyder 1988). The basic problem with TTR is its sensitivity to text length. If a text is very long then certain words tend to be repeated, high-frequency words will be repeated more often as compared to low-frequency words, and this tendency will increase with the length of the text. In spite of this, TTR is still used for comparing text production, for instance between children's texts, or between various groups with language impairment.

(b) Guiraud index: Several measures have been proposed in order to solve the
problem with text length. One example is the index of Guiraud (Guiraud 1954), which is a type/token based measure that is supposed to be independent of text length. The index of Guiraud results from dividing the number of types by the square root of the number of tokens. For a long text, this procedure will result in a higher lexical richness than what would have been obtained with a simple TTR.

(c) **Mean Segmental Type-token ratio (MSTTR):** The mean segmental type-token ratio (MSTTR) tries to avoid the problems of the traditional type-token ratio, which has been shown to be sensitive to text length. The MSTTR divides texts into segments of forty words (section sizes are generally decided by the length of the smallest available text). The ratio of different words (types) to the total number of words (tokens) for each section is recorded. Then the mean scores for each participant’s segments are added and divided by the total number of segments in the text.

(d) **Lexical density:** Lexical density is the term most often used for describing the proportion of content words (nouns, verbs, adjectives, and often also adverbs) to the total number of words. By investigating this, we receive a notion of **information packaging:** a text with a high proportion of content words contains more information than a text with a high proportion of function words (prepositions, interjections, pronouns, conjunctions and count words). Variants of lexical density have been proposed. A popular ‘minor variant’ is to calculate the **noun density,** the number of nouns divided by the total number of tokens in
the text. Other options are for instance verb or adjective or adverb types per total lexical words.

(e) *Lexical variety/diversity/richness:* Lexical diversity is a measure of the count of different words that are used in a text.

**Table 2.3: Frequently used measures of Lexical complexity**

<table>
<thead>
<tr>
<th>Lexical variety</th>
<th>Mean segmental type-token ratio (MSTTR)</th>
<th>Type-Token ratio (TTR)</th>
<th>Gilabert 2005; Rahimpour 1997; Robinson 1995; Michel, Kuiken &amp; Vedder 2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical density</td>
<td>Proportion of new to old words in the task</td>
<td>Robinson, Ting &amp; Urwin 1995</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Percentage of open-class words as opposed to closed-class grammatical, words per utterance</td>
<td>Robinson 1995</td>
<td></td>
</tr>
</tbody>
</table>

Both (c) and (d) have the advantage of being easy to operationalize, and also practical to apply in computer analyses of large data corpora. However, both lexical diversity and lexical density have been shown to be significantly higher in writing than in speaking (Ure 1971; Halliday 1985). This could be because the two measures are interchangeable, or are a significant way to account for lexical development.

### 2.3.4 Global versus specific measures of task performance

The use of global measures of linguistic complexity, accuracy, and fluency is currently under debate (e.g., Pallotti 2009). In order to complement the results on task performance by means of complexity, accuracy and fluency, Robinson and colleagues propose to use
task specific measures (Cadierno & Robinson 2009; Robinson et al. 2009; Robinson & Gilabert 2007). The rationale of using specific measures is that “Such specific measures should be more sensitive to conception, task complexity, and its linguistic demands than general measures” (Robinson et al. 2009: 550). Especially, if tasks are designed in such a way that the demands can be met by the use of specific linguistic structures (resource-directed tasks), task specific measures in all probability would fill in the gaps left by the global complexity, accuracy and fluency (CAF) measures.

For example, the effect of an increase of cognitive task complexity by manipulation on the ‘here-and-now’ factor would be captured best by assessing specific linguistic markers of the past, e.g., past tense and temporal conjunctions. Analyzing past tense in such a task may complement the results obtained by global measures. Therefore, the use of specific measures next to global CAF may give more insights in the actual linguistic performance of an L2 learner.

Robinson, Cadierno and Shirai (2009) studied the effects of increasing complexity of task demands in conceptual domains using specific measures of the accuracy and complexity of speech. Results showed that there is an advanced use of tense–aspect morphology on complex tasks compared to simple tasks. Also more target-like-use of lexicalization patterns was seen for referring to ‘motion’ on complex tasks.

Another study by Michel (2011) examined the use of task specific measures like discourse markers and coherence. Michel (2011) in her work on task complexity assessed frequency of different kinds of conjunctions and found that neither the frequency nor the
occurrence of conjunctions was affected by cognitive task complexity. The only significant effect was found with respect to one specifically task relevant conjunction. So far, this chapter has reviewed theoretical approaches to the cognitive strand of research into task-based L2 performance. The following section will review earlier research done on task complexity dimension.

2.4 Studies on Task complexity

In this section, the various studies and their results in task complexity dimension are discussed. Section 2.4.1 gives details about the studies in the resource-directed dimension. Section 2.4.1.1 gives details of studies and their results for oral tasks and Section 2.4.1.2 gives details of studies and their results for written tasks.

2.4.1 Resource directedness

Recall that resource-directed tasks (± no reasoning demands, ± few elements) are those in which the linguistic and conceptual demands made by a complex task, can be met by certain specific features of the linguistic system. Increasing complexity along these dimensions will therefore direct learners’ attentional and memory resources to the way L2 structures occur, consequently leading to interlanguage development. Let us now look at the studies on oral tasks.
2.4.1.1 Oral performance

Much research has been done using oral tasks (Robinson 1995, 2001a, 2001b, 2005, 2007; Robinson, Ting & Urwin 1995; Rahimpour 1997; Niwa 2000; Iwashita, McNamara & Elder 2001; Deng 2005; Gilabert 2005, 2007; Ishikawa 2006). Some of the dimensions studied were: amount of cognitive load imposed, number of elements and reasoning demands.

Robinson (1995) studied the impact of manipulating Here-and-Now [-complex] on three different narratives where less fluent speech was predicted for There-and-Then [+complex] tasks but higher lexical and structural complexity as well as accuracy for There-and-Then tasks. Results showed that There-and-Then [+complex] task promoted accuracy (as measured by target-like use of articles) though not significantly. Complexity showed no difference between [+complex] and [-complex] tasks (as measured by S-nodes per T-unit) in this study. However in another study, Robinson, Ting and Urwin (1995) found that complex tasks (i.e., There-and-Then compared to Here-and-Now) generated a trade-off between complexity-accuracy and fluency: complexity and accuracy increased but at the expense of fluency. This result was confirmed in a picture description task (differing in their cognitive demands) where participants were asked to produce oral narratives by Iwashita, McNamara & Elder (2001). In another study by Robinson in 2001, lexical variety (i.e. lower type-token ratio) was also found to increase in [+complex] (direction-giving task of unfamiliar area as compared to [-complex] direction-giving task of familiar area) condition, but again at the cost of fluency. The trade-off between accuracy and fluency in oral tasks was not observed in Niwa (2000) where in complex
tasks (+reasoning demands as compared to [-complex] –reasoning demands), both complexity (greater S-nodes per T-unit) and fluency (longer mean length of phonation) increased. Interestingly in Rahimpour (1997) on crossing a complexity variable (Here-and-Now) with a condition variable (open vs. closed), results were found contradictory to previous studies. He found that closed tasks generated significantly more fluent speech, but there was only a trend for higher accuracy in closed tasks, and no differences in complexity between open and closed tasks.

In a more recent study, Robinson (2007) used interactive narrative oral task: a simple task with a small degree of reasoning demands, a neutral task with moderate reasoning demands, and a difficult task with high reasoning demands. The results indicated no statistical difference in the three different conditions in either accuracy or complexity.

Gilabert (2005) used a simple here-and-now task and a complex there-and-then task and found that the complex task elicited significantly more accurate language production (percentage of self-repairs). Fluency (percentage of self-repairs) was negatively affected by increases in cognitive complexity along both variables. However, increased task complexity did not have a significant effect on language complexity (S-nodes per T-unit). In Gilabert’s (2007), narrative task was manipulated along ±Here-and-Now, an instruction-giving task manipulated along ±elements, and the decision-making task which is manipulated along ±reasoning demands. Results show an overall effect of Task Complexity on self-repairs (measure of accuracy) behavior across task types, with different behaviors existing among the three task types. No differences were found
between the self-repair behavior between low and high proficiency groups. Complexity and fluency were not measured in this study.

Shiau and Adams (2011) replicated and extended Gilabert’s (2007) study and investigated the effects of manipulating task complexity through increasing reasoning demands of a task. However, unlike Gilabert’s (2007) study where only accuracy was studied, Shiau and Adams not only studied accuracy but also the effect of complex tasks on linguistic complexity. The results indicated that complex tasks in terms of reasoning demands had no significant effect on accuracy and syntactic complexity, but it led to significantly more lexically and syntactically varied learner production.

2.4.1.2 Written performance

A study was conducted by Ishikawa (2006) on the effects of task complexity (Here-and-Now and There-and-Then) and language proficiency on L2 written narrative production (using picture prompts). It was found that structural complexity (S-nodes per T-unit), accuracy (target length use of articles) and fluency (words/T-unit) increased when task complexity increased. Lexical complexity (square of word types to the total number of words) decreased with increase in proficiency level of learners. The low-proficiency learners received greater benefits when task complexity was modified.

A study to analyze the relation between task complexity (±elements) and writing was done by Kuiken, Mos and Vedder (2005). 62 Dutch learners of Italian were the subjects for this study. Based on cloze test scores and amount of learning experience, subjects
were divided into two proficiency groups of 31 students each. Subjects had to write a recommendation letter to a friend about the location for a holiday. Learners had to choose from 5 given holiday destinations. Certain requirements (the location, facilities provided, food, etc.) had to be fulfilled in making the choice for a holiday destination. As the complexity dimension was number of elements, in the simple task, three conditions had to be met and in the complex task, six conditions had to be met. Lexical richness (Guiraud’s index) and syntactic complexity (T units) did not show any task complexity effect. Accuracy (number of error-free T-units) increased with increase in task complexity.

Michel, Kuiken and Vedder (2007) tested the influence of complexity in monologic (a task which does not require interaction between two learners and rather involves only one learner) versus dialogic task (requires interaction or a dialogue between two learners to perform a task). Few elements were present for simple tasks as compared to complex tasks which comprised many elements. Accuracy was measured by total number of errors per AS-unit, number of lexical errors and the total number of omissions. Total number of clauses per AS-unit and sub clauses per total number of clauses, were the measures used to access structural complexity. Lexical complexity was measured by the percentage of lexical words and the number of types divided by the square root of number of tokens. Fluency was measured by speech rate A [syllables/min in unpruned speech (including reformulations, repetitions)] and speech rate B [syllables/min in pruned speech (without reformulations, repetitions)] and number of filled pauses per 100 words. It was found that accuracy, lexical complexity and fluency increased for dialogic tasks whereas for
monologues, structural complexity increased. More complex tasks resulted in more diverse speech.

Another research was done by Kuiken and Vedder (2007) on task complexity and measures of linguistic performance in L2 writing which was an extension of the study by Kuiken, Mos and Vedder (2005) study. 84 Dutch university students of Italian and of French origin were subjects for this study. This study was conducted to test the conflicting claims of Skehan’s Limited Attention Capacity Model (LACM) and Robinson’s Cognition Hypothesis (CH). Accuracy was measured by total number of errors per T-unit with respect to grammar, lexicon, orthography and appropriateness. Syntactic complexity was measured by number of clauses per T-unit and number of dependent clauses per clause. Lexical variation was measured by number of types per square root of twice of number of tokens. ANOVA was done for each task. It was found that students performed better in complex tasks than on non-complex tasks. In Italian, more frequent words were found in complex task (similar to Cognition Hypothesis). In French, more infrequent words were found in complex tasks (similar to Skehan’s LACM). Fewer lexical errors and orthographic errors (accuracy) were found in complex task.

In a similar study by Kuiken and Vedder (2008), two groups of proficiency (of Dutch university students of Italian and of French origin) performed two writing tasks with prompts of differing cognitive complexity. No interaction of task type and proficiency level were observed, though in general, complex tasks led to greater accuracy, but not
greater structural complexity and lexical variety. The results were in contrast with earlier findings where high proficiency learners showed greater increase on accuracy measures (Kuiken, Mos & Vedder 2005).

A summary of studies on written performance with respect to task complexity are presented in Table 2.4.

### Table 2.4: Summary of research on task complexity (resource-directing) and writing production

<table>
<thead>
<tr>
<th>Studies</th>
<th>Context of study</th>
<th>Task complexity</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Robinson, Ting &amp; Urwin 1995</td>
<td>Learners of Mandarin Chinese</td>
<td>Narrative± planning</td>
<td><strong>Fluency</strong>: no change</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Accuracy</strong>: no change</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Complexity</strong>: increase</td>
</tr>
<tr>
<td>Kuiken, Mos &amp; Vedder 2005</td>
<td>Dutch learners of Italian</td>
<td>Number of elements (letter writing)</td>
<td><strong>Complexity</strong>: ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Accuracy</strong>: decrease</td>
</tr>
<tr>
<td>Ishikawa 2006</td>
<td>Japanese high school learners of English</td>
<td>Here-and-Now/There-and-then(narrative)</td>
<td><strong>Accuracy</strong>: increase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Syntactic complexity</strong>: increase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Fluency</strong>: increase</td>
</tr>
<tr>
<td>Kuiken &amp; Vedder 2007</td>
<td>Dutch university students of Italian and of French origin</td>
<td>Number of elements (letter writing)</td>
<td><strong>Accuracy</strong>: increase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Complexity</strong>: increase</td>
</tr>
<tr>
<td>Kuiken &amp; Vedder 2008</td>
<td>Dutch university students of Italian and of French origin</td>
<td>Number of elements (letter writing)</td>
<td><strong>Accuracy</strong>: increase</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Structural complexity &amp; lexical variety</strong>: no change</td>
</tr>
<tr>
<td>Abdollahzadeh &amp; Kashani 2011</td>
<td>High proficiency and low proficiency Iranian EFL participants</td>
<td>Here-and-Now/There-and-then(narrative)</td>
<td><strong>Complexity &amp; accuracy</strong>: increase for High proficiency learners</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><strong>Fluency</strong>: no change</td>
</tr>
</tbody>
</table>

2.4.2 Resource-dispersion

Recall that resource-dispersion dimensions (±prior knowledge, ±planning) do not direct learners to any particular aspects of language code which can be used to meet the additional task demands. Robinson, further hypothesizes that increasing complexity along
resource-directing dimensions can facilitate accuracy and complexity though at the expense of fluency. On the contrary, increasing complexity along resource-dispersion dimensions can be expected to affect fluency, as well as accuracy and fluency (Robinson 2005: 4-7).

Most of the studies in the resource-dispersion paradigm were based on the planning time dimension- the amount of planning time given before a task. Since we are looking at pre-task (planning time) in resource-dispersion dimension, we review the studies conducted in this area.

2.4.2.1 Planning time studies

Table 2.5 gives details of all the studies done under planning dimension. Research into the effects of ± planning time on production have shown that providing learners with pre-task planning time seems to benefit their fluency and complexity, while for accuracy there is no consensus. According to studies in Skehan’s paradigm (Crookes 1989; Foster & Skehan 1996; Skehan & Foster 1997; Wigglesworth 1997; Mehnert 1998; Ortega 1999; Yuan & Ellis 2003), for cognitively simple tasks (with planning time), there is a marked increase in fluency, relatively strong effect on complexity and very little effect on accuracy.
<table>
<thead>
<tr>
<th>Study</th>
<th>Participants</th>
<th>Tasks Description</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foster 1996</td>
<td>32 intermediate ESL learners in the UK</td>
<td>3 tasks: (1) personal information; (2) narrative; and (3) decision making.</td>
<td>Fluency and complexity: increase Accuracy: increase in (1) and (3)</td>
</tr>
<tr>
<td>Foster &amp; Skehan 1996</td>
<td>As in Foster (1996)</td>
<td>As in Foster (1996)</td>
<td>Fluency: increase for tasks 1 and 3 Complexity: increase Accuracy: no effect</td>
</tr>
<tr>
<td>Robinson, Ting &amp; Urwin 1995</td>
<td>11 learners of Mandarin Chinese as a SL</td>
<td>narration/ written (narration using picture prompts)</td>
<td>Accuracy and complexity: No effect Fluency: increase only in planned speaking tasks</td>
</tr>
<tr>
<td>Wigglesworth 1997</td>
<td>107 adult ESL learners in Australia. 28 divided into high &amp; low proficiency</td>
<td>5 tasks varying in cognitive difficulty. 1 minute planning time.</td>
<td>Fluency, complexity, and accuracy: increase in the high proficiency learners. But was not evident in all measures and also not in the analytic ratings.</td>
</tr>
<tr>
<td>Skehan &amp; Foster 1997</td>
<td>40 pre intermediate ESL learners in UK</td>
<td>Similar to Foster (1996)</td>
<td>Fluency: increase in (1) and (2) only. Complexity: increase in (1) and (3) only Accuracy: increase in (3) only</td>
</tr>
<tr>
<td>Mehnert 1998</td>
<td>31 undergraduate students of German intermediate level</td>
<td>2 tasks: instruction task (structured); &amp; exposition (unstructured) task. 4 planning groups: no planning; 1 min; 5 min; and 10 min</td>
<td>Fluency: increase Complexity: no effect Accuracy: structured task (10 min planning) increase Density: structured task (10 min planning) increase</td>
</tr>
<tr>
<td>Ortega 1999</td>
<td>64 advanced level learners of Spanish as an FL</td>
<td>2 narrative tasks based on picture prompts. 10 minutes unguided planning + note- making. Retrospective semi structured interviews.</td>
<td>Fluency, complexity and accuracy: increase</td>
</tr>
<tr>
<td>Wigglesworth 2001</td>
<td>400 ESL learners at different levels of proficiency in Australia</td>
<td>5 task types. Planning - 5 min. Planning (manipulated with 2 other task characteristics: familiarity &amp; structure. 1 task performed without manipulation.</td>
<td>Fluency: increase for low proficiency learners Complexity: increase for high proficiency Accuracy: increase for high proficiency</td>
</tr>
<tr>
<td>Yuan &amp; Ellis 2003</td>
<td>42 Chinese undergraduate Students (English majors)— TOEFL 373–520</td>
<td>Oral narrative picture task. No planning &amp; pre-task planning (10 min)—written notes. Task performance pressured.</td>
<td>Fluency and complexity: increase (planning group than no-planning group). Accuracy: no effect</td>
</tr>
<tr>
<td>Kawauchi 2005</td>
<td>12 high intermediate and 11 advanced Japanese EFL learners</td>
<td>Three picture-based narrative tasks. ±planning</td>
<td>Fluency: in (-planning): low &lt; high &lt; advanced (+planning): low &lt; high = advanced. Complexity: increase only in advanced group Accuracy: increase only in low proficiency</td>
</tr>
<tr>
<td>Study</td>
<td>Participants</td>
<td>Task Description</td>
<td>Results</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------</td>
<td>------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Skehan &amp; Foster 2005</td>
<td>61 intermediate level ESL students</td>
<td>Decision-making task</td>
<td>Fluency: increase in planning groups. Complexity: (iii) increase. Accuracy: (iii) increase</td>
</tr>
<tr>
<td>Elder &amp; Iwashita 2005</td>
<td>197 adult ESL learners in Australia (TOEFL range 427–670)</td>
<td>2 narrative tasks. There were also 6 other tasks completed under different conditions. Planning = 3 min + 75 s. No planning = 75 s to read instructions. Planning was manipulated in conjunction with other task characteristics &amp; conditions.</td>
<td>Fluency: No effect Complexity and Accuracy: decrease</td>
</tr>
<tr>
<td>Tavokoli &amp; Skehan 2005</td>
<td>80 adult EFL learners in Iran elementary and intermediate level</td>
<td>Narrative tasks of differing complexity based on picture prompts. +5 minutes planning.</td>
<td>Fluency, complexity and accuracy: increase (especially in high proficiency learners)</td>
</tr>
<tr>
<td>Ojima 2006</td>
<td>Three English as a Second Language (ESL) Japanese students writing performance.</td>
<td>Effect of concept planning (as a resource-dispersion factor and as a form of pre-task planning)</td>
<td>Fluency and Complexity: increase Accuracy: no effect</td>
</tr>
<tr>
<td>Gilabert 2007</td>
<td>48 first and second-year university students with a lower intermediate proficiency level in English</td>
<td>Four conditions: (1) no- planning/Here and- Now; (2) 10 min planning/Here and- Now (3) no-planning/There and- Then (4) 10 min planning/ There-and-Then (Narrative tasks based on picture prompts)</td>
<td>Fluency and complexity: Increases in (2) and (4)) Accuracy: (1) and (2) no effect; increase in (3) and (4)</td>
</tr>
<tr>
<td>Ong &amp; Zhang 2010</td>
<td>108 Chinese EFL tertiary students enrolled in the Communication Skills Programme at a comprehensive university in Singapore</td>
<td>Argumentative writing. Four groups: extended pre-task, pre-task, free-writing, and control group.</td>
<td>Fluency: no effect Lexical complexity: increase</td>
</tr>
<tr>
<td>Nakakubo 2011</td>
<td>143 intermediate and high-intermediate university learners of Japanese</td>
<td>Narrative task under one of four planning conditions: a) no planning, b) planning before speaking performance, c) planning during performance, and d) planning both before and during performance</td>
<td>Fluency and Accuracy: no effect Lexical complexity: increase</td>
</tr>
<tr>
<td>Ghavamnia, Tavakoli &amp; Esteki 2013</td>
<td>40 intermediate EFL learners from a language center in Iran</td>
<td>Learners were randomly assigned to either the pre-planning or the online planning conditions and were required to complete a written narrative based picture prompts.</td>
<td>Fluency and complexity: increase in pre-task planning group Accuracy: increase in online planning group</td>
</tr>
</tbody>
</table>
Reviewing previous research makes it clear that there are only a few studies which examined the effects of resource-dispersion factors (e.g. planning time, number of tasks, and prior knowledge) on written language production. Since we are looking at planning dimension with writing as the modality, we present details of the writing studies where planning dimension has been studied.

Ellis and Yuan (2004) looked at the effects of planning on fluency, complexity and accuracy in second language narrative writing. In this study, planning was operationalised at three levels: No Planning (NP), Pre-Task Planning (PTP) and On-Line Planning (OLP). In the NP condition, subjects had to finish writing a story in 17 minutes and had to write at least 200 words. In the PTP condition, subjects were given 10 minutes planning time to write a story within 17 minutes. The minimum number of words remained the same (200 words). There was no fixed time limit for OLP condition. No pre-task planning time was allocated but sufficient time for on-line planning was given. Ellis and Yuan found that fluency (operationalised by syllables per minute and number of dysfluencies) increased in both cases where planning time was given. Higher syntactic complexity (ratio of clauses to T-units) was also seen in the two planning groups. Syntactic variety was measured by calculating the total number of different grammatical verb forms used in the task. Another complexity measure used was Mean Segmental Type-token ratio. In terms of accuracy (measured by error-free clauses and correct verb forms), OLP had the highest accuracy, followed by PTP group and the NP group had the lowest accuracy. Pre-task planning has a marked effect on written fluency, a relatively strong effect on linguistic complexity (more subordination and a greater variety of verb
forms in planned task than in unplanned task) and very little effect on accuracy.

In Robinson, Ting and Urwin (1995), 11 learners of Mandarin Chinese as a second language were asked to write written description of a sequence of pictures presented. For, planned (3-minute planning time) and unplanned task (no planning time), parallel but different picture strip for each condition was given. The tasks were separated by a period of few days. Accuracy was measured by target like use of three features of Chinese- use of verb morphology, use of measure words, and use of tense markers and fluency was measured in words and T-units and accuracy. It was found that there was no difference in both fluency and accuracy for both planned and unplanned task. Complexity was not measured for this task.

Ong and Zhang (2010) used resource-dispersion dimensions of task complexity to analyze the effects of task complexity on the fluency and lexical complexity of learner’s argumentative writing. They manipulated task complexity along two factors: availability of planning time and provision of ideas and macrostructure. Four groups were presented with different levels of planning time: extended pre-task, pre-task, and free-writing. The fourth was a control group. In addition, ideas and macrostructure for the task were provided along three levels: topic, ideas, and macrostructure group; topic and ideas group; and topic group. Ong and Zhang (2010) found that complex tasks (through the provision of ideas and macrostructure) led to greater lexical complexity though had no effect on fluency.
Ojima (2006) examined the effect of concept planning (as a resource-dispersion factor and as a form of pre-task planning) on three English as a Second Language (ESL) Japanese student’s writing performance. He reported that pre-task planning produced greater fluency and complexity, but did not improve grammatical accuracy.

Hence we can see that mixed results were obtained for studies on writing in planning dimension though all planned tasks led to better language. Since in this study, we will look at the interaction between resource-directing and resource-dispersion dimension, let us look at the research done in that area.

2.4.3 Proficiency

Some studies looked at the interaction between proficiency and task complexity. In Ishikawa’s (2006) study, which looked at written production for higher proficiency learners, there was approximately an increase of 1.5 words per T-unit when complexity of the tasks was increased along the ± resource-directed dimension. Lexical complexity was found to be inversely proportional to proficiency level of learners. The low proficiency learners received greater benefits when task complexity was modified.

In Kawauchi (2005), low ESL learners gained the most in accuracy whereas high ESL learners gained more in syntactic complexity and lexical variation. In the high proficiency group the complex task triggered greater lexical and syntactic complexity and accuracy, to the detriment of fluency. In the low proficiency group, on the complex task,
fluency was boosted while the other areas remained intact irrespective of cognitive task complexity (Malicka & Levkina, 2012).

In Kuiken and Vedder (2007) results reveal that it was not possible to establish any interaction effect of task complexity and proficiency level. The students of French, however, made significantly more appropriateness and other errors, but also fewer orthography errors in the complex task than in the non-complex one, whereas for students of Italian no differences were found. This different finding for Italian and French is difficult to explain. Although on the whole the students of French were more proficient than the students of Italian so that fewer Appropriateness errors could be expected, it is not clear why more Appropriateness errors were made in the complex task. On the question of the effect of complexity on lexical variation, in terms of word frequency, the students of Italian and French showed a different behaviour. The students of Italian used significantly more high frequent words in the complex task, whereas for the students of French there were more infrequent words in the complex task. Therefore the finding for French seems to be in line with Robinson’s Cognition Hypothesis, and the one for Italian with the Limited Attentional Capacity Model. Again it is difficult to explain how these different findings have to be related to the assumed higher general proficiency level of the students of French.
### Table 2.6 Proficiency studies in task complexity

<table>
<thead>
<tr>
<th>Research</th>
<th>Subjects</th>
<th>Task</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wigglesworth 1997</td>
<td>400 ESL learners at different levels of proficiency in Australia</td>
<td>Planning time (speaking)</td>
<td><strong>Complexity</strong> and <strong>Accuracy</strong>: increase in high proficiency learners. <strong>Fluency</strong>: increase in low proficiency learners.</td>
</tr>
<tr>
<td>Kuiken, Mos &amp; Vedder 2005</td>
<td>62 Dutch learners of Italian</td>
<td>Letter-writing. No. of elements</td>
<td><strong>Complexity</strong>: No effect <strong>Accuracy</strong>: increase when task complexity &amp; proficiency both high. Low proficiency group: no effect</td>
</tr>
<tr>
<td>Kawauchi 2005</td>
<td>12 high intermediate &amp; 11 advanced Japanese EFL learners</td>
<td>Three picture-based narrative tasks.</td>
<td><strong>Fluency</strong>: (-planning): was low &lt; high &lt; advanced but in (+planning): low &lt; high = advanced. <strong>Complexity</strong>: no effect <strong>Accuracy</strong>: increase only in low proficiency</td>
</tr>
<tr>
<td>Ishikawa 2006</td>
<td>52 Japanese high school learners of English</td>
<td>Narrative, Here and Now/ There and then</td>
<td><strong>Accuracy</strong>: increase. <strong>Structural complexity</strong>: increase. <strong>Lexical complexity</strong>: decreases with increase in proficiency <strong>Fluency</strong>: increase</td>
</tr>
<tr>
<td>Kuiken &amp; Vedder 2007</td>
<td>84 Dutch university students of Italian &amp; 75 students of French</td>
<td>Letter-writing. No. of elements</td>
<td><strong>Accuracy</strong> and <strong>Complexity</strong>: increase</td>
</tr>
<tr>
<td>Kuiken &amp; Vedder 2008</td>
<td>91 Dutch university students of Italian &amp; 76 students of French</td>
<td>Letter-writing. No. of elements</td>
<td><strong>Accuracy</strong>: increases. <strong>Structural complexity &amp; Lexical variety</strong>: no effect</td>
</tr>
<tr>
<td>Abdollahzadeh &amp; Kashani 2011</td>
<td>32 high proficiency and 36 low proficiency participants</td>
<td>Writing. Here-Now &amp; There-Then</td>
<td><strong>Complexity &amp; Accuracy</strong>: increase in complex task for high-proficiency learners. <strong>Fluency</strong>: No effect</td>
</tr>
<tr>
<td>Malicka &amp; Levkina 2012</td>
<td>37 (20 advanced and 17 pre-intermediate undergraduates)</td>
<td>Speaking, ±Few elements, ± spatial reasoning</td>
<td>High proficiency: <strong>Accuracy &amp; complexity</strong> increases. <strong>Fluency</strong>: decreases. Low proficiency: <strong>Fluency</strong> increases. <strong>Accuracy &amp; Complexity</strong>: no effect</td>
</tr>
</tbody>
</table>
Malicka and Levkina (2011) explored the effect of language proficiency on the language production of learners of different proficiency levels in English. 37 undergraduate students were the participants (20 advanced and 17 pre-intermediate learners of English). They were provided with two tasks manipulated along ±few elements and ±spatial reasoning dimensions. Complexity, accuracy, and fluency measures were used to analyze participants’ speech production on the two tasks. In the high proficiency group the complex task led to greater lexical and structural complexity and accuracy at the cost of fluency. In the low proficiency group, on the complex task, fluency increased while the other areas remained unaffected irrespective of cognitive task complexity. Hence, to sum up, high proficiency learners gained more by the increase in task complexity.

2.4.4 Interaction studies

As far as I know, only a few studies (Michel, Kuiken & Vedder 2007; Gilabert 2007; Levkina 2008) have combined simultaneously two cognitive variables and no written studies have specifically looked at the effects of manipulating the resource-directing (narrative vs. argumentative) and resource-dispersion dimension (±pre-task).

In Michel, Kuiken and Vedder (2007) study on the effects of task complexity (±few elements) and task condition (±monologic) on L2 production, the interaction of task complexity and task condition was only statistically significant for accuracy: in the monologic condition task complexity prompted more accurate performance.
Gilabert (2007) analyzed the effects of manipulating the cognitive complexity of tasks on L2 narrative oral production. Increasing complexity along the [±here-and-now] variable makes learners reduce their fluency of speech, but they gear their attention toward lexical complexity and accuracy, while the here-and-now dimension only slightly improves fluency. As far as [±planning time] is concerned, increasing task complexity by reducing planning time does not seem to direct learners’ attention to any grammatical features of the language. However, with pre-task planning time given to the learners they displayed improvements in lexical variety as well as in fluency.

Levkina (2008) examined the impact of manipulating task complexity simultaneously along ±planning time and ±few elements on L2 oral production. 14 upper-intermediate English level Spanish and Russian students were asked to perform four decision-making tasks manipulated along ±planning time and ±elements combinations. Data was analyzed by means of standardized measures for syntactic and lexical complexity, accuracy and fluency. The results showed no significant effects on accuracy or syntactic complexity. However lexical complexity and fluency partially corroborate earlier research. Longer pre-task planning and simpler tasks resulted in more fluent speech while shorter pre-task planning and more complex tasks increased lexical complexity. As far as the combined effects between two variables (±few elements / ±planning time) of task complexity are concerned, it is observed that participants’ oral production in this study was mostly affected by ±few elements; in particular, fluency, lexical and structural complexity dimensions; whereas accuracy did not show a significant difference. Presence of pre-tasks did not seem to show significant differences in language production. Levkina
attributed it to the reason that subjects being of high proficiency were not affected by the lack of pre-task planning time.

Simultaneous manipulation, hence led to mixed results in the three studies. As only a few studies have looked at the interaction between different dimensions, this research is therefore an attempt at analyzing a new dimension within the field of task complexity studies.

2.5 The Study

The study explores the impact of task complexity on writing. This study investigates task complexity along three dimensions:

- ±Resource directedness (± reasoning demands),
- ±Resource-dispersion (pre-tasks),
- Simultaneous manipulation along resource directedness (± reasoning demands) and resource-dispersion (pre-tasks).

The effect of these task complexity manipulations on learner writing were measured on five linguistic measures (syntactic complexity, syntactic variety, lexical density, lexical variety and accuracy) and three discourse measures (frequency of reference markers, frequency of linkers and coherence).

The research questions are:
a) Does cognitive task complexity have a beneficial effect on the accuracy and complexity (lexical and structural) on the written performance of L2 learners?

b) Do tasks involving more thinking have an effect on structural and lexical complexity and accuracy?

c) Do resource-directed tasks increase the accuracy and complexity of written production?

d) Do Pre-task help in task performance and increase structural and lexical complexity and accuracy?

e) Does a simultaneous manipulation along resource-directed dimension and resource-dispersion dimension increase the accuracy and complexity of written production?

f) Do different learners at different proficiency levels show similar language production on complex tasks?

In this study we will be working within Robinson’s Triadic Componential Framework. Most of the research conducted in this area has been done with intermediate and English as foreign Language (EFL) students. In our study we examine how users of English (ESL) at varying proficiency levels would perform the same complex task.