Effects of Task Complexity and Interaction on L2-Performance

Abstract

The Cognition Hypothesis (Robinson 2005, this volume) claims that increases in cognitive task complexity along resource-directing variables focus L2-learners' attention on language. As a result, speech is more accurate, syntactically more complex and lexically more diverse, while a cutback in fluency is expected. Interactive tasks are also supposed to heighten the joint attention to language form such that L2-learners produce more accurate speech in dialogic rather than in monologic tasks. As speakers interrupt each other, ask for clarifications, and lexically and syntactically mirror each other's speech, pair work leads to structurally less complex and lexically less diverse output while fluency increases. When manipulating both task complexity and interaction, combined effects are expected. As a complex task needs more joint clarification work, complex dialogic tasks may lead to a more accurate performance of lower structural and lexical complexity.

The present study empirically investigates the effects of increased cognitive task complexity and interaction: 64 learners of Dutch as an L2 performed a simple and a complex task either individually (+ monologic) or in pairs (– monologic). Cognitive task complexity was manipulated along the factor +/- few elements. The speech is analyzed for measures of linguistic complexity, accuracy, and fluency. The output of the L2-learners is compared to the performance of 44 native speakers of Dutch.

This study provides little support for the Cognition Hypothesis, as hardly any effects of increased cognitive task complexity become visible. Interaction, however, displays a large impact on both L2-learners' and L1-speakers' task performance. These results are discussed from a cognitive perspective on attentional allocation and L2-task performance.

1 Attention and Task Performance

Cognitive approaches to task-based research focus on how differences in the cognitive demands of a task affect the performance of second language (L2) learners on that task (Skehan 1998, Skehan and Foster 2001, Robinson 1995, 2001, 2005). Of special interest is the way changes in cognitive task complexity draw the L2-learner's attention towards language and consequently influence task performance. The present study investigates the effect of an increase in cognitive task complexity on task-based L2-performance. Furthermore, it evaluates the influence of a change in interactive setting on task performance in a second language.

The remainder of this first section gives an overview of theoretical and empirical research concerning the two factors, cognitive task complexity and interaction, within the task-based paradigm. Section 2 elaborates on the research questions, hypotheses, and method of an empirical study that investigated the effects of cognitive task complexity and interaction on the performance of L2-learners and L1-speakers of Dutch. The third section presents the results by addressing effects on linguistic complexity, accuracy, and fluency. Section 4 discusses the results of the present work from a cognitive perspective on attentional allocation. Thereby it addresses the differences between monologic and dialogic performances on simple and complex tasks by L2-learners and L1-speakers. The fifth and last section concludes with suggestions for future task-based research and practical implications for task-based language teaching.

1.1 Cognitive Task Complexity

With respect to the capacity and allocation of attention during L2-task performance there are two competing models: the Limited Attentional Capacity Model (e.g., Skehan 1998, Skehan and Foster 2001) and the Multiple Attentional Resources Model (Robinson 1995, 2007, this volume). The most important prediction of Skehan's model is, that '[...] attentional limitations for the L2-learner and -user are such that different areas of performance compete with one another for the resources that are available' (Skehan and Foster 2001:205). Based on VanPatten (a.o. 1990) Skehan states that an increase in cognitive task demands puts pressure on the attentional system. Accordingly, L2-learners have to prioritize either linguistic complexity, accuracy, or fluency. Performing on a demanding task results first in trade-off effects between form on the one hand and fluency and meaning on the other hand. Limited attention for form in turn generates a trade-off between linguistic complexity and accuracy. Consequently, a task leading to linguistically more complex L2-performance

decreases the accuracy of that performance, and vice versa. Only if more resources are available, e.g., with more planning time, is a parallel increase on all constructs possible.

Robinson (1995, 2001, 2005, this volume) has suggested a contrasting account, the Multiple Attentional Resources Model, which has become better known as the Cognition Hypothesis. Applying a cognitive psychological model by Wickens (1998, 2007), the Cognition Hypothesis proposes the availability of multiple attentional resource pools during L2-task performance. Accordingly, an L2-learner can draw on more than one attentional pool for different aspects of performance. As a result a complex task may lead to an increase in linguistic complexity without showing trade-off effects on accuracy.

Within cognitive factors of task complexity the Cognition Hypothesis distinguishes (in the so-called Triadic Componential Framework, cf. Robinson, this volume) between resource-directing and resource-dispersing variables. Increasing the cognitive demands along resource-directing factors draws the learner's attention towards the language code. When L2-learners try to meet the cognitive demands of a complex task the cognitive complexity makes them use a more elaborate and varied lexis and produce more complex linguistic structures. For example, a task with many elements, rather than a few elements, is expected to ask for a more specific lexis and more complex syntactic structures because all the different elements need to be distinguished and compared. Thus, the increase in cognitive demands guides the attention of the learner towards task-relevant (linguistic) aspects. Complex tasks may push L2-learners to shift to a syntactic mode of production (Givon 1995) such that form and meaning are enhanced in parallel. Since the cognitive processing can rely on multiple attentional pools accuracy and linguistic complexity do not compete with each other for attentional resources. Both accuracy and complexity are raised as long as the complexity is increased on a resource-directing cognitive factor. Only the more procedural measure of fluency may decrease due to a higher processing load in complex tasks.

Increasing the demands on resource-dispersing variables has a contrary effect on learner speech. A higher cognitive task complexity on this kind of factor may lead to trade-off effects between linguistic complexity and accuracy. Not providing planning time or lack of prior knowledge disperses the task performer's attention away from the linguistic code. L2-performance suffers.

Like Skehan's model, the Cognition Hypothesis thus acknowledges performance problems on these kinds of cognitively demanding tasks. However, according to Robinson (2003) these problems are not a result of resource limitations (as proposed by Skehan). Rather they stem from interference problems due to loss of control over attention (Navon 1989). This interference causes involuntary attentional shifts which draw away the second

language learner's attention from task-relevant linguistic aspects. Inefficient, unfocused processing of relevant and irrelevant information causes the linguistic complexity, accuracy, and fluency to suffer.

TABLE 1 ABOUT HERE

Several empirical investigations have focused on the effect of increased cognitive task complexity on L2-learners' task performance within Robinson's Cognition Hypothesis. As these are of particular interest to the study at hand, Table 1 summarizes work that studied effects of the resource-directing factor +/- few elements (e.g., Robinson 2001, Gilabert 2007, Kuiken and Vedder 2007). These studies indeed support the claims of the Cognition Hypothesis because accuracy and complexity were raised by an increase in cognitive task complexity. Even so, as a whole they give an inconclusive picture as e.g., a parallel increase of linguistic complexity and accuracy manifests itself as a trend-effect only (Robinson 2001). In particular, Gilabert 2007 opens new questions: why are the effects of cognitive task complexity more consistent in instruction-giving tasks than in other task types?

Nuevo 2006, Michel et al. 2007, Revesz 2008, Gilabert et al. 2009, and Kim 2009 (see Table 1) evaluated effects of cognitive task complexity in relation to interaction. Before discussing these studies section 1.2 elaborates on the theoretical basis of interactive approaches to task-based L2-research. Interaction is, after cognitive task complexity, the primary focus of attention of the present work.

1.2 Interaction

The Triadic Componential Framework (Robinson, this volume) lists different interactive factors of task condition. So-called participation variables (e.g., one-way/two-way flow of information, open/closed or convergent/divergent task outcome) are of particular interest to the present study as they can be manipulated in order to create tasks that ask for more or less interaction. Long's (1990) Interaction Hypothesis (along with others, e.g., Pica 1994, Mackey 1999, Gass et al. 1998, Gass and Mackey 2006) highlights the importance of interaction for interlanguage development. During interaction learners have to make meaningful use of their L2-knowledge because they need to understand and be understood. In their Output Hypothesis Swain and Lapkin (1995) state that while comprehension can be successful in a pragmatic or semantic mode, producing output pushes L2-learners towards a

syntactic mode of processing. In dialogues L2-learners are encouraged to try out new forms and learn through hypothesis testing. When they fail at being comprehensible they receive negative feedback from the speaking partner. Clarification requests and negotiations of form and meaning are the result. During these so-called language related episodes (LRE) interlocutors discuss the lexical items and morphosyntactic structures they use. This points their attention towards differences between the interlanguage and target L2-form. Noticing (the gap between) input and output enhances uptake and intake of new information (Schmidt 1990, Pica 1994). Furthermore, it may lead to modified output upon negative feedback. Without losing the primary focus on meaning, interaction thus draws the L2-learner's attention to the linguistic code. In the long run, all these aspects of interaction are considered to contribute to second language learning.

In the light of Robinson's (this volume) Triadic Componential Framework interaction may be studied in terms of information-flow: how task-relevant information is distributed over participants, and to what extent successful task performance involves the exchange of that information (Long 1990). In its most radical form the factor one-way/two-way flow of information turns into the factor +/- monologic. In a monologue the one-way flow of information is the default while in a dialogue a two-way flow of information is expected.

Monologues generate fewer opportunities for language learning because noticing, uptake, and intake of new information is less likely to occur than in dialogues. Hence, due to the lack of feedback monologues induce fewer instances of attention to form and meaning. In addition, generating modified output in a monologue is a result of self-monitoring only – an effortful process that needs time and attentional capacity, especially in the L2 (Kormos 2000).

Bringing attention into play, interaction can also be studied from a cognitive point of view. Here the availability of attentional resources during monologic and dialogic task performance are of particular interest. In a monologue speech production by definition relies on the knowledge and cognitive resources of a single speaker. This speaker has to plan and conceptualize the continuous message during the actual speaking process. In a dialogue, one can conceptualize one's own speech output during the interlocutor's turn. The absence of a speaking partner in a monologic task condition thus generates a need for more online-planning. As a monologue thereby may increase the cognitive load of an L2-task, Tavakoli and Foster suggest that 'a monologic task [...] makes greater demands on attentional resources than an interactive task' (2008:461).

Interactive task performance may be further eased by so-called alignment, i.e., the tendency of speakers to copy each other's language at all linguistic levels. Alignment 'greatly

simplifies production and comprehension in [native speakers'] dialogue' (Pickering and Garrod 2004:169). As their interlanguage system is neither complete nor fully automatized, L2-learners may not be able to profit from alignment to the same extent as native speakers. Even so, a dialogic task condition may free up attentional resources for non-native speakers too (Costa et al. 2008). This cognitive perspective on interaction thus assumes that in a dialogue there is more planning time available and more attentional resources may be freed by alignment processes. Consequently, creating a message in interaction is less effortful than formulating an utterance from scratch in a monologue.[1]

Earlier work on effects of interaction manipulated on the factor +/- monologic is rather limited within task-based research. Only Michel et al. (2007) systematically implemented this variable when investigating L2-task performance. The data of that study displayed large effects of interaction on measures of linguistic complexity, accuracy, and fluency. The interactive setting made participants more accurate and fluent, but decreased the syntactic complexity of L2-speech. Lexical complexity was unaffected by interaction. From a meta-analysis of their earlier work Skehan and Foster (2007) conclude that dialogues in contrast to monologues increase the accuracy of the participants at the cost of fluency. In addition dialogues show a slightly higher complexity. However, this meta-analysis compares monologic and dialogic task performances of L2-learners on a range of different tasks in different conditions, such that it is difficult to draw generalizable conclusions.

To summarize this review on interaction, dialogic in contrast to monologic tasks do allow negotiation of meaning, typically generate LREs, and may be cognitively less effortful. In dialogues L2-speakers therefore not only have more attention available, but the interactive setting also directs their attention towards the language code. Consequently, L2-learners are expected to generate more accurate speech in dialogues than in monologues. Data of Skehan and Foster (2007) and Michel et al. (2007) give support to this assumption. These empirical and theoretical accounts suggest furthermore that in dialogues fluency benefits from the lower procedural pressure. As interlocutors tend to fill each other's thinking pauses in order to keep the flow of speaking and turn taking stable (Fiksdal 2000), fluency may be further increased. With respect to linguistic complexity, the story is not as straightforward. On the one hand, the heightened attention may promote the use of more complex structures and a more elaborate lexis. On the other hand, interlocutors interrupt each other and ask for clarification. Often, dialogues consist of short utterances containing single clauses because interaction prevents complex subordinate or coordinate syntactic structures from being

[1] Note that from a pragmatic point of view dialogues may be more complex than monologues.

produced (Robinson 2001). As the work by Michel et al. (2007) shows that these effects of interaction are quite large, dialogues are predicted to show a lower linguistic complexity than monologues. Hence, the turn taking behavior weakens the effects of heightened attention in interaction. Moreover, alignment may also result in lower lexical and syntactic complexity because speakers tend to copy each other.

From the review in sections 1.1 and 1.2 it becomes clear that more research into effects of increased cognitive task complexity by means of the factor +/- few elements and the interactive factor +/- monologic respectively is needed. Both factors are interesting on their own, however when combining the two a third exciting question arises: what happens when L2-learners act on cognitively complex interactive tasks? Section 1.3 briefly discusses this point.

1.3 Cognitive Task Complexity and Interaction

The Cognition Hypothesis states that cognitive task complexity should have similar effects on L2-learners' task performance in monologues and dialogues: accuracy and linguistic complexity are promoted at the cost of fluency. However, Robinson (2005) suggests that increased cognitive task complexity possibly generates more interaction because more clarification and negotiation work is needed in a complex task. As cognitive task complexity and interaction then both guide the attention towards language form, L2-performance on complex interactive tasks may become even more accurate than on simple interactive tasks. Concerning linguistic complexity, Robinson furthermore assumes that manipulations on task complexity factors are more evident in one-way tasks than in two-way tasks: '[…] the greater amount of interaction and turn taking facilitated under complex task conditions may mitigate against the attempt of either participant to produce extended utterances and lengthy turns' (Robinson 2005:21). Thus, the promoting effect of cognitive task complexity is moderated by the decreasing effect of interaction. Similarly, fluency is expected to be more affected by interaction than by cognitive task complexity. Accordingly, in complex interactive tasks L2-learners' fluency may be high because interaction speeds them up irrespective of the fact that increased cognitive task complexity slows them down.

To recap, complex interactive L2-performances are expected to be more accurate and more fluent but linguistically less complex than simple or complex monologues or simple dialogues, respectively.

Task-based research combining cognitive task complexity and interaction has focused mostly on the amount and type of interaction due to changes in cognitive task complexity

(see Table 1: Robinson 2001, Nuevo 2006, Revesz 2008, Gilabert et al. 2009, Kim 2009). The findings reveal that increased cognitive task complexity mostly leads to more interaction, though not all task types bring consistent results (Nuevo 2006, Gilabert et al. 2009). Again, Michel et al. (2007) is the only study that investigates the combined effects of interaction and increased cognitive task complexity in a systematic way. Data from that study challenge the Cognition Hypothesis. While in monologic tasks increased cognitive task complexity fostered accuracy and lexical complexity, in the dialogic task condition this promoting effect disappeared. Again there is a lack of studies researching the combined effect of cognitive task complexity and interaction.

Another gap in task-based research concerns native speaker data. Up to now surprisingly few studies have included L1-speakers (with the exception of Dörnyei and Kormos 1998, Foster 2001 and Foster and Tavakoli 2009). Nonetheless, only if we know how task design influences native speakers' linguistic complexity, accuracy, and fluency, can we fully understand the measures we use for the evaluation of non-native performances.

Speech production in L1 needs less attention and is less effortful than speaking in L2. In terms of Levelt (1989) only the conceptualization and monitoring ask for conscious attention whereas the formulator and articulator process information automatically (de Bot 1992, Poulisse and Bongaerts 1994). Changes in cognitive task complexity that influence attentional allocation therefore are not expected to substantially affect native oral task performance.

Interaction, however, presumably influences native speakers similarly to L2- learners. Also L1-speakers do take turns and alignment presumably decreases the complexity of their syntax and lexis (Pickering and Garrod 2004). Similarly to L2-learners, the fluency measures based on conceptualization and monitoring (i.e., pausing and repair) are expected to increase in native speaker dialogues. L1-speakers' accuracy most likely is at ceiling in any task condition because it is based on native competences and skills.

To summarize, Table 2 gives a schematic overview of the predicted effects of increased cognitive task complexity and changes in interactional setting on L2- and L1-task performance, as discussed in sections 1.1, 1.2, and 1.3. It shows the predictions for both factors on their own as well as in combination. To paraphrase, cognitively complex tasks generate more accurate and linguistically more complex speech at the cost of fluency. Interactive tasks increase accuracy and fluency but decrease linguistic complexity. In combination cognitively complex interactive tasks promote accuracy and fluency but decrease linguistic complexity of task-based L2-performance.

TABLE 2 ABOUT HERE

2 The Present Study

The goal of the present study is to provide empirical data that may fill some of the gaps addressed in section 1. The main aim is to investigate the claims of the Cognition Hypothesis of Robinson (1995, 2001, 2005, this volume) with respect to the effects of increased cognitive task complexity. Furthermore, it aims to test the effects of interaction on L2-task performance. Taking the two together, it explores whether there are any combined effects of increased cognitive task complexity and interaction on task-based L2-speech. In addition, it evaluates the L2-learners' performance against a baseline of L1-speakers. The following research questions and hypotheses guide the empirical investigation:

1 *Research Question:* What is the effect of increased cognitive task complexity on L2 oral task performance?
Hypothesis: Increased cognitive task complexity results in higher accuracy, higher linguistic complexity but lower fluency of L2 oral task performance.

2 *Research Question:* What is the effect of interaction on L2 oral task performance?
Hypothesis: Interaction raises the accuracy and fluency of L2 oral task performance while decreasing linguistic complexity.

3 *Research Question:* Are there any combined effects of increased cognitive task complexity and interaction on L2 oral task performance?
Hypothesis: Cognitively complex interactive tasks promote the accuracy of L2 oral task performance. Increased cognitive task complexity enhances interaction and further decreases the linguistic complexity while increasing fluency.

4 *Research Question:* Are there any effects of increased cognitive task complexity and interaction, on their own or in combination, on L1 oral task performance?
Hypothesis: Increased cognitive task complexity does not affect oral task performance of L1-speakers. Accordingly, no combined effect of cognitive task complexity and interaction is expected. In contrast, interaction shows similar effects on L1-speakers as on L2-learners.

2.1 Experimental Design

The hypotheses were tested using a 2x2-design with cognitive task complexity (simple versus complex) as a within-participant factor and interaction (monologue versus dialogue) as a between-participants factor. The oral performance of L2-learners was investigated in simple monologues (condition 1), complex monologues (condition 2), simple dialogues (condition 3), and complex dialogues (condition 4). Baseline data come from L1-speakers performing under the same conditions. All participants performed on a simple and a complex task. Half of them acted on their own in a monologic setting (conditions 1 and 2), the other half in pairs in the dialogic setting (conditions 3 and 4). Following Robinson's (2005) framework the factor cognitive task complexity was implemented by means of the resource-directing factor +/- few elements. Table 3 gives a schematic overview of the experimental design with its four different conditions.

TABLE 3 ABOUT HERE

2.1.1 Participants

64 learners of Dutch as a second language participated in the present study. All L2-learners were of Turkish (N = 30) and Moroccan (N = 34) origin. 44 native speakers of Dutch were included as a control group. All 108 participants were attending or had finished higher levels of education.

In order to assess the general level of Dutch proficiency participants performed on a written test: in eight short texts there were in total 100 gaps where test-takers had to choose among three possible words per gap.[2] Language proficiency scores on this test revealed that the native speakers, as expected, performed at ceiling. Most learners were at an intermediate level of L2-proficiency.[3] Table 4 summarizes the background information for all participants.

[2] I thank the Language Centre of the University of Groningen which uses this task as a placement test for their language courses.

[3] The scores on the language proficiency task of one learner and one native speaker were adjusted to their group's mean +/- 2 Standard Deviations as they were outliers. This correcting procedure was adopted rather than eliminating the participants from the dataset, since a closer look at their experimental task performance revealed that they were within their group average on the performance measures.

TABLE 4 ABOUT HERE

2.1.2 The Simple and Complex Tasks

Two sets of argumentative tasks were designed on different topics (dating and study). For both topics, a simple (+ few elements) and complex (– few elements) version of the task were created. In the simple condition, participants received a full color sheet with four photographs of young people; in the complex condition there were six pictures. Next to each photograph a list of six characteristics of the depicted person was given.

The instructions for the dating task explained that the people depicted (half of them male, half of them female) were contestants on a dating show. Based on their characteristics (age, smoking, reading, sports behavior, favorite hobby and music), participants had to decide which two out of the four or six contestants would make the best male-female couple.

The study task showed four or six females, differing in age, nationality, subject of education, reading behavior, duration of attending a Dutch language course, and whether they planned to take a final exam in Dutch or not. Participants learned in the instructions that these women (half of them studying in Belgium, half of them in the Netherlands) were applying for a language student exchange program. As only two of them would be admitted, participants had to choose which Belgium-Dutch pair of students would make the best studying couple.

For both topics there were four possible combinations in the simple condition and nine possible couples in the complex condition. The Appendix gives an example of the tasks and instructions.

2.1.3 Procedure

Participants were tested individually (+ monologic) or in pairs (– monologic). Interactants in the dialogic condition shared their linguistic background and were of the same gender (e.g., both Turkish women). Task complexity and task topic were counterbalanced over participants. All started with the first experimental task (simple or complex, dating or study), performed afterwards on the paper-and-pencil language proficiency test, and closed with the second experimental task (complex or simple, study or dating).

For the oral task, participants were instructed that they should call a friend. On the phone they should reason about their decision on the best dating or study couple. In the monologic

setting they learned that their friend would not be at home so they were asked to leave a message on the answering machine. In the dialogic condition participants were asked to discuss with each other on the phone about their choice. Performers received 2 minutes of planning time and were instructed to speak for about 3 minutes in the monologic and 6 minutes in the dialogic condition. The researcher encouraged participants to be explicit and elaborate when explaining their decision. For example, they were asked to include all possible combinations in their reasoning process as they should not only give reasons why they favored a pair, but should also explain why others would not make a good couple. When they stopped reasoning within the allotted speaking time, the researcher repeated this encouragement and successfully made participants talk at least 3 and 6 minutes, in the monologic and dialogic conditions respectively.

Immediately after the experimental tasks participants rated their performance on a 5 point Likert scale on affective variables: perceived task difficulty, confidence, stress, frustration, and motivation (Gilabert 2007).

Between the two experimental speaking tasks, participants worked on the paper and pencil Dutch language proficiency task. L2-learners received 30 minutes time, native speakers only 15 minutes. In order to keep time on task equal for all participants the L1-speakers also completed a written Dutch dummy task which took them another 15 minutes.

Except for the manipulation on +/- few elements and +/- monologic all tasks were controlled for the other cognitive and interactive factors named in the Triadic Componential Framework (Robinson, this volume).

2.1.4 Transcription, Measures, and Statistical Analyses

The 108 participants produced in total 216 (simple and complex) speech samples that were recorded on a laptop with the free software Scanner Recorder 1.9 (Jacobs 2006).

Transcription

Using the CLAN program (MacWhinney 2000) recordings were transcribed and coded for measures of linguistic complexity, accuracy, and fluency. First, speech samples were divided into Analysis of Speech (AS)-units (Foster et al. 2000). This basic syntactic unit takes into account that oral task performance consists of many elliptical utterances. Afterwards, the data were coded for ten different measures by four students of linguistics and the researcher. After several joint training sessions every coder received a subset of the speech samples. Roughly 10% of the data was double coded by a student and the researcher. When they did not agree, the correct coding was discussed and adjustments were made. In

a final round, the researcher checked all transcripts. This procedure resulted in an interrater-reliability between students and the researcher of 97%.

Measures

Three measures of linguistic complexity were chosen (two structural and one lexical). The number of words per clause serves as a measure of clause complexity (Ferrari 2009) while the Subordination Index (i.e., the number of subordinate clauses per AS-unit) gives insight into the ability of participants to use complex syntax. Guiraud's Index of lexical diversity (Guiraud 1954) is a measure of lexical complexity which adjusts the commonly used type token ratio (TTR) for differences in sample length (Vermeer 2000).

The present study distinguishes three different accuracy measures: lexical errors, morphosyntactic errors, and determiner errors, all related to the number of AS-units. Lexical errors are defined as word choice errors (e.g., wrong prepositions or forms in a language other than Dutch). Morphosyntactic errors include syntactic (e.g., word order errors or omissions of obligatory constituents) and morphological (e.g., agreement, or nominal, verbal and adjectival inflection) problems. Dutch determiners are problematic for L2-learners because Dutch marks grammatical gender on the article. As the expected high occurrence of these kinds of errors would bias any other category a separate measure was used. It includes any erroneous use of definite and indefinite articles, article omissions, and mismatches of grammatical gender on articles, demonstratives, and relatives.

Following Tavakoli and Skehan (2005) the present study distinguishes measures of speed, pausing, and repair fluency. Two speed measures are used: the unpruned Speech Rate A (in syllables per second) gives a view of how well a speaker can fill time with sound, while the pruned Speech Rate B, which is cleaned for reformulations, repairs and repetitions, is a measure of 'meaningful' speech per second. Pausing fluency is calculated by the number of filled pauses (uh and uhm) per AS-unit. The number of self-repairs of errors and non-errors per AS-unit gives an insight into the speaker's ability to monitor his or her own speech (see e.g., Kormos 2000, Gilabert 2007).

Table 5 gives an overview of all the measures used in the present study.

TABLE 5 ABOUT HERE

Statistical Analyses

CLAN-output was sent to SPSS 16.0. Three mixed GLM analyses of variance (ANOVA) were computed for every construct (linguistic complexity, accuracy, and fluency) with task complexity as the within participant factor and interaction as the between participants factor. Separate analyses for L2-speakers and L1-learners were conducted.

3 Results

The following sections give the descriptives and results of the statistical analyses for linguistic complexity, accuracy, and fluency respectively.

3.1 Effects on Linguistic Complexity

Table 6 shows the absolute means on measures of linguistic complexity for L2-learners and L1-speakers. Comparing the numbers on simple and complex tasks, no obvious performance difference is visible for either group, with the exception of Guiraud's Index: both groups display a higher lexical complexity in the complex task than in the simple task.

TABLE 6 ABOUT HERE

With respect to differences due to interaction both groups show a decrease from monologic to dialogic tasks on the two structural measures: monologues have a higher number of words per clause and more subordination per AS-unit than dialogues. Concerning lexical complexity, however, L1-speakers show a decrease in dialogic versus monologic tasks while L2-learners scored higher on Guiraud's Index in dialogues than they did in monologues.

TABLE 7 ABOUT HERE

Table 7 displays the results of the mixed GLM ANOVA that tested the statistical value of the differences. This analysis indeed reveals no main effect of cognitive task complexity on either accuracy, fluency, or structural measures of linguistic complexity. Guiraud's Index,

however, is significantly affected in both groups: L2-learners reveal a smaller effect than L1-speakers but in both populations the lexical complexity increases significantly in complex tasks (L2: partial $\eta^2 = 0.091$; L1: partial $\eta^2 = 0.138$).

Effects of interaction were significant and large for both groups on both structural measures (words per clause L2: partial $\eta^2 = 0.212$; L1: partial $\eta^2 = 0.483$; Subordination Index L2: partial $\eta^2 = 0.170$; L1: partial $\eta^2 = 0.457$). A smaller effect could be detected with respect to the lexical measure (Guiraud's Index L2: partial $\eta^2 = 0.066$; L1: partial $\eta^2 = 0.139$). None of the combined effects of task complexity by interaction reached significance.

As pointed out above, non-native and native speakers produced similar results for linguistic complexity with the exception of Guiraud's Index in interactive tasks. Therefore a second mixed GLM ANOVA that included the factor native–learner as a between-participants factor was performed. Results with respect to the combined effect of the two factors native–learner X interaction are displayed in the bottom rows of Table 7. As can be seen, all three measures show a significant differential effect for the L2-learners and L1-speakers concerning monologic versus dialogic tasks. Post-hoc pairwise comparisons revealed that the two syntactic measures were significantly influenced in the same direction in both groups. L1-speakers' decrease in syntactic measures between monologues and dialogues was larger than for L2-learners. More interestingly, lexical complexity was affected in opposite directions in the two groups: Guiraud's Index in non-native speakers was lower in monologues than in dialogues, whereas in native speakers lexical complexity was higher in the monologues than in dialogues (difference monologue–dialogue: L1 = 0.428, $p < 0.05$; L2 = -0.337, $p < 0.05$).

3.2 Effects on Accuracy

Table 8 summarizes the scores of the L2-learners and native speakers on measures of accuracy. L2-learners show high numbers of all types of errors: morphosyntax is affected the most, the lexicon moderately, and determiners the least. L1- speakers display some lexical errors but the other two types of errors seem to be of negligible size. There are no obvious differences between simple and complex tasks with respect to these numbers. Interactive tasks though decrease the number of errors in both groups. Whereas native speakers show only slight gains, the L2-learners almost halve the number of errors of any type when acting in pairs.

TABLE 8 ABOUT HERE

Table 9 displays the results of the mixed GLM ANOVA that tested the statistical reality of these differences.

TABLE 9 ABOUT HERE

As expected from the absolute scores, none of the statistical comparisons of cognitive task complexity reached significance in either group. Nor did any of the combined effects of cognitive task complexity by interaction turn out significant. In contrast, interaction revealed significant results and accordingly large effect sizes for the L2-learners on all three accuracy measures (partial $\eta^2 > 0.33$). For L1-speakers the number of lexical errors and determiner errors were moderately affected by the factor interaction (lexical: partial $\eta^2 = 0.256$; determiners: partial $\eta^2 = 0.092$). Both populations made fewer errors in dialogues than in monologues.

3.3 Effects on Fluency

TABLE 10 ABOUT HERE

Table 10 gives the descriptives for the four measures of fluency. From these absolute numbers no consistent difference between simple and complex task performances is visible in either group. The interaction numbers however suggest that L2-learners as well as L1-speakers are more fluent in dialogues than in monologues. This is reflected in both speed measures (higher numbers, i.e., faster, in dialogues than in monologues) as well as in the pausing and repair behavior (lower numbers, i.e., fewer pauses and repairs, in dialogues than in monologues).

TABLE 11 ABOUT HERE

Table 11 summarizes the statistical outcome of the mixed GLM ANOVA on the fluency data. The p-values of this analysis indicate that there was no main effect of cognitive task complexity on any measure of fluency – neither for L2-learners nor for L1-speakers. Again, interaction did influence both groups.

L2-learners were affected significantly on all measures, showing larger effect sizes on repair and pausing behavior (partial $\eta^2 > 0.25$) than on the two Speech Rates (partial $\eta^2 > 0.11$). Native speakers' fluency was significantly affected by interaction with respect to the unpruned Speech Rate A (partial $\eta^2 = 0.092$) along with a significant and large effect on pausing behavior (partial $\eta^2 = 0.482$). Both pruned Speech Rate B and repair fluency were not affected significantly, but SRB showed a trend.

The only combined effect of cognitive task complexity by interaction that turned out to be significant in the dataset of the present study appeared in the fluency measures of native speakers. Planned post-hoc comparisons evaluating this large effect on both Speech Rates showed that L1-speakers were faster in simple rather than complex dialogues while no effect of cognitive task complexity was visible in monologues (SRA: monologue simple–complex = -0.022, n.s.; dialogue simple–complex = 0.254, $p < 0.01$; SRB: monologue simple–complex = -0.027, n.s.; dialogue simple–complex = 0.238, $p < 0.01$). Non-native speakers showed a different trend on their Speech Rates: increased task complexity tended to yield faster speech in dialogues, but slower speech in monologues. This comparison however was not statistically significant.

In order to summarize the results of all measures, Table 2, which gave the predicted effects, is repeated here as Table 12 based on the results of the study. As the structural and lexical measures of linguistic complexity sometimes revealed a different pattern, they are listed separately.

TABLE 12 ABOUT HERE

---

4 Discussion

The results with respect to the Cognition Hypothesis are discussed first (section 4.1), followed by the results of interaction (section 4.2). Section 5 summarizes the findings by giving suggestions for future research and formulating the practical implications of the work at hand.

4.1 Effects of Increased Cognitive Task Complexity?

The Cognition Hypothesis, as formulated in the first and third hypotheses of the present study, predicts that an increase of cognitive task complexity along resource-directing factors results in higher accuracy and linguistic complexity, but lower fluency, of L2-task performance. In interactive tasks, increased cognitive task complexity promotes interaction such that the syntactic complexity and lexical diversity are decreased in complex interactive tasks while accuracy and fluency are enhanced.

In the present study, cognitive task complexity manipulated on the factor +/- few elements only affected one measure of L2-performance: higher scores on Guiraud's Index of lexical diversity were found in complex tasks when compared to simple ones. Native speakers, in addition, showed a combined effect of cognitive task complexity with interaction on fluency: in dialogues, but not in monologues, increased cognitive task complexity slowed L1-speakers down (as measured on the two Speech Rates). No other measures were affected by an increase in cognitive task complexity nor by the combination of cognitive task complexity and interaction.

The findings with respect to lexical complexity are in line with Robinson's predictions, but no parallel increase in accuracy was found. A central claim of the Cognition Hypothesis, however, is that accuracy and complexity are both promoted by increased cognitive task complexity. Therefore, the present study gives little support for Robinson's claims. Since there was neither an increase nor a decrease in any accuracy measure, the data do not suggest the existence of trade-off effects between accuracy and linguistic complexity. Consequently, the present study does not support the Limited Attentional Capacity model of Skehan (1998) either.

This study is not in line with earlier work that manipulated the factor +/- few elements. For example, Michel et al. (2007) and Robinson (2001) did find significant increases of accuracy and lexical complexity in complex oral L2-performance, as did Kuiken and Vedder (2007) for written L2-tasks. Likewise, Gilabert (2007) found an increase in repair behavior in complex narrative tasks. The question arises, then, why the present study found such minor results. The subsequent paragraphs discuss four possible explanations, that address (1) the difference in cognitive load between the simple and complex tasks, (2) different effects of

cognitive task complexity in different task types, (3) general CAF-measures versus task specific measures, and (4) qualitative versus quantitative changes in linguistic performance.

First, perhaps the difference between the simple and the complex tasks was not large enough. The present study did single out the resource-directing factor +/- few elements and it was manipulated based on findings from cognitive psychology: Halford et al. state that our working memory and reasoning limitations share a central capacity that 'is limited to relations between four variables' (2007:240). Accordingly, the simple task of the present study, giving four elements/combinations, should be within human capacity limits whereas the complex task, with six elements/nine combinations, should be beyond it. On these theoretical grounds, the manipulation should produce a difference in cognitive load as proposed by Robinson's Triadic Componential Framework.

The work by Gilabert and colleagues (Gilabert 2007, Gilabert et al. 2009) points towards a second possible explanation. In their studies, increased cognitive task complexity did affect task performance in an instruction-giving task, but not in a decision-making task. One may assume then that the factor +/- few elements was overruled by the factor +/- reasoning. Hence, +/- reasoning is a factor of cognitive task complexity itself (cf. the Triadic Componential Framework Robinson, this volume). For other task types increasing the cognitive task complexity on the factor +/- few elements may differentiate L2-task performance. Support for this idea can be found in studies involving narrating a picture story (Robinson 2001), giving instructions in a map task (Gilabert 2007), or evaluating inanimate items (as e.g., the mobile phone tasks in Michel et al. 2007). In a complex reasoning task, however, the manipulation by the number of elements may not substantially affect the attentional allocation. Consequently, no difference in task performance is visible.

Results of the questionnaire on affective variables of the present study corroborate this explanation: participants' perceptions of task difficulty were not significantly different between the simple and complex tasks.[4] Assuming differential effects of cognitive task complexity in different task types possibly serves as a correct explanation.

Robinson (2007) mentions a third possibility: it may be that the global measures of linguistic complexity, accuracy, and fluency used in the present study are not sensitive enough to discern the differences in performance caused by increased cognitive task complexity. Although these measures have been used reliably in task-based research (e.g., Skehan and Foster 2001, Robinson 2001, Michel et al. 2007) recent work corroborates the

[4] Mean (standard deviations) on a 5 point Likert scale (1 = difficult, 5 = easy) of L2-learners: monologues simple = 3.6 (0.6) and complex = 3.5 (0.7), dialogues simple 3.9 (0.6) and complex 3.9 (0.5) and L1-speakers: monologues simple = 3.1 (0.6) and complex = 3.0 (0.6), dialogues simple = 3.7 (0.7) and complex = 3.8 (0.5).

added value of using task specific measures. For example, Cadierno and Robinson (2009) and Revesz (2008) evaluated L2-task performance by means of reference to psychological and cognitive state terms, number of wh-clauses, or clausal conjoinings. Task complexity effects were more prominent on these task specific measures than on general measures. A more qualitative look at the data at hand therefore will focus on the use of (conditional and causal) conjunctions (Michel, in preparation). This analysis by means of a more specific measure for the argumentative reasoning tasks used in the present study accordingly will evaluate this third explanation.

All the same, the data of the present study on general CAF-measures do not support the idea that increased cognitive task complexity manipulated on the factor +/- few elements leads to a more focused attention towards language form by L2-learners. Hence, the fact that lexical complexity was enhanced in complex tasks can be accounted for by a rather simple idea: the complex tasks in this study consisted of more elements (= persons) that were each associated with more words (the six characteristics per person), see Appendix. For example, the input in the simple task showed four different numbers (= the ages of the people depicted) whereas the complex task mentioned six different numbers (= the ages of six people). This pattern of input vocabulary holds for most characteristics mentioned per person. This results in twelve more words in the input of the complex task than in the input of the simple task. Calculating Guiraud's Index with twelve more types and tokens mirrors the gain from simple to complex tasks in the study at hand quite well.[5] Native speakers show a similar gain in lexical complexity and consequently corroborate this assumption.

A fourth explanation may gain support from its contrast with earlier work. Kuiken and Vedder (2007), Gilabert (2007) and Gilabert et al. (2009) did find effects of a manipulation on the number of elements. These studies operationalized the factor +/- few elements, though in a different way than the present work. For example, Kuiken and Vedder (2007) asked participants to balance reasons about different holiday destinations. When comparing possible bed-and-breakfasts in the simple task they had to take into account three criteria (e.g., close to the city centre). The complex version offered participants the same number of bed-and-breakfasts but they had to take into account six characteristics. Similarly, the design of Gilabert (2007) and Gilabert et al. (2009) increased the number of elements by means of a higher number of 'internal' criteria. In contrast, the present study added more elements that were accompanied by the same amount of characteristics.

[5] In this study, L2-learners used in the simple task on average about 95 types and 232 tokens which equals a Guiraud of types/tokens = 6.24; hypothesizing that L2-learners use every input word about three times (36 more tokens, 12 more types) the value of Guiraud (= 6.54) comes close to the average Guiraud found for complex tasks.

A more plausible explanation for the effects of the factor +/- few elements as it was manipulated in the present study then is as follows: rather than influencing the linguistic aspects of the task performances in terms of 'different' or 'more elaborate' language use, i.e., a qualitative change, the manipulation of the single factor +/- few elements results in a quantitative change only, i.e., 'more of the same'.[6] From a pedagogic perspective it may be interesting to have L2-learners practice a wider range of vocabulary in a more complex task by means of more elements. From a research-theoretical point of view, however, that seeks to understand the interplay between task characteristics and attentional allocation, the use of more words as a result of more words in the given input is rather trivial (Pallotti 2009).

A last comment concerning Robinson's Cognition Hypothesis addresses the combined effects of cognitive task complexity and interaction. While effects of cognitive task complexity were expected to be larger in L2-monologues than in L2-dialogues, in the current study, cognitive task complexity only mattered (on the fluency measures) in L1-dialogues. Thus, as in Michel et al. (2007) the only significant combined effect found contradicts the predictions of Robinson's theory.

To conclude, the data of the present study give little support for claims of the Cognition Hypothesis with respect to the effects of cognitive task complexity on their own or in combination with interaction. Other than an increase in lexical complexity, which may be explained by the input given, the study at hand therefore suggests that increased cognitive task complexity manipulated on the single factor +/- few elements does not affect L2-learners' attentional allocation and task performance as proposed by the Cognition Hypothesis (e.g., Robinson 1995, 2001, Robinson and Gilabert 2007).

4.2 Interaction and L2-Performance

The second hypothesis of the present study predicted that interactive tasks raise the accuracy and fluency of L2-task performance while syntactic complexity and lexical diversity is decreased due to interactional turn taking and alignment processes. Results indeed show significant main effects of interaction on all measures: in dialogues L2-learners were more accurate, lexically more diverse and more fluent but structurally less complex than in monologues. L1-speakers display a similar pattern with the exception of lexical complexity: as hypothesized, Guiraud's Index was lower in native dialogues than monologues.

Accuracy measures show consistently that L2-learners made fewer errors in dialogues than in monologues. All effects were large, i.e., the number of errors was almost halved in the interactive setting. Especially with respect to lexical accuracy, dialogic performances

[6] I thank Rod Ellis and an anonymous reviewer for raising this point.

generated more accurate speech than monologues. Since even the small amount of lexical and determiner errors by native speakers was significantly lowered in dialogues, the conclusion from the present work is that interaction has a beneficial effect on the accuracy of task performance of L2-learners and L1-speakers.

Interaction also promotes the fluency of L2-learners with respect to speed and pausing as well as repair behavior. Effects are the largest for pausing and repair, where interactive tasks almost halve the scores.

This positive effect of interaction is predicted by the Interaction Hypothesis (Long 1990), the Output Hypothesis (Swain 1985, Swain and Lapkin 1995), and Schmidt's Noticing Hypothesis (Schmidt 1990). The joint focus on the language code as proposed by the Cognition Hypothesis also predicts higher accuracy measures in dialogues (Robinson 2001, 2005, 2007).

The parallel increase in fluency, however, allows another interpretation: Tavakoli and Foster (2008) assume that dialogues put less procedural pressure on L2-task performance than monologues. As speakers in dialogues can plan their performance during the interlocutor's turn, they have more resources available for the actual speech production during their own turn. In a monologue, hesitations occur because the non-automatic speaking process in L2-speakers is interfered with by active online-planning. In dialogues, a joint gain in measures of accuracy and fluency (with respect to speed, pausing and repair) is possible because the formulator and articulator can rely on more attentional resources (Levelt 1989). The general assumption that dialogues are cognitively simpler than monologues is in line with the Alignment Hypothesis (Pickering and Garrod 2004, Costa et al. 2008) It also fits into the Limited Attentional Capacity Model (Skehan 2001), where increased planning time is expected to reduce trade-off effects between fluency and form (i.e., accuracy).

The fact that for L1-speakers dialogues significantly reduced the number of pauses and increased the unpruned Speech Rate A, but showed no effect on the pruned Speech Rate B or repair behavior, further corroborates this idea. Hence, pausing behavior is most directly related to conceptualizing, a process that needs attentional resources in native speakers too. SRB is cleaned for repairs and reformulations such that it is not affected by online planning, while the unpruned SRA reflects this gain in processing resources.

In addition, the task difficulty judgments of the participants revealed a greater perceived ease in monologues than in dialogues. As a whole, these results thus support the assumption that increased planning time and alignment, as a natural byproduct of interaction, leave more attentional resources for speech production in dialogues. These

cognitive aspects of the factor +/- monologic thus may serve as an explanation for the effects of interaction on learner speech: the lower cognitive complexity induced by the factor (– monologic) in interactive tasks has beneficial effects on form and meaning in the oral task performance of L2-learners and L1-speakers.

Effects of interaction on linguistic complexity were large in both populations. L2-learners and native speakers used more complex clauses and more complex syntactic structures in monologues than in dialogues. As these decreasing effects are larger for native speakers than for L2-learners, possibly a lower syntactic complexity in interactive tasks equals native behavior. Due to turn taking, sentences become short and simple and less subordination occurs in dialogues. Accordingly, clarification work, comprehension checks, and interruptions prevent speakers (whether in their L1 or L2) from using complex syntactic structures. Also Pallotti (2009) points out that lower linguistic complexity is sometimes more nativelike than complex syntactic structures - which presumably is manifested by the data of the present study.

The fluency measures in turn support the idea that interlocutors prevent each other from building complex clauses and syntactic structures: dialogues generate fewer silences than monologues. Apparently, interactants start speaking as soon as the interlocutor hesitates. Thereby they reduce the structural complexity of the joint speaking performance because the speaking partner has no chance to perform complex syntactic operations.

In contrast, the measure of lexical complexity (Guiraud's Index) was higher in L2-dialogues than in L2-monologues. At first, the fact that L1-speakers show the reversed picture is puzzling. However, taking into account the parallel gains on (lexical) accuracy and fluency the following explanation may hold: due to freed attentional resources in dialogues L2-learners benefit from each other on all linguistic levels, i.e., accuracy, fluency, *and* lexical complexity.

Pickering and Garrod (2004) assume that in dialogues the interlocutors recycle each other's speech to a large extent, which they call alignment. One central process of alignment is routinization: interlocutors establish and agree on conversation specific phrases and words, i.e., routines, which they keep using during this conversation (Pickering and Garrod 2004). The second hypothesis of the present study therefore expected that linguistic complexity would decrease in dialogues, which indeed is manifested by the native speaker data.

For L2-learners, alignment results in a different pattern. Non-natives also copy each other's words and phrases. However, the incomplete knowledge of the target language

inhibits full application of routinization: for L2-learners, copying the words of the interlocutor leads to a more diverse vocabulary use, because together they can use more different words than either would have come up with on their own. In a monologue an L2-learner only uses their own limited lexicon. In a dialogue they profit from the input of the speaking partner by incorporating his or her lexical choices. As a result, the lexical complexity increases in joint L2-task performance.

To recap, the present study shows that interactive tasks increase the accuracy, lexical complexity, and fluency of L2-speakers while monologic tasks allow for the use of more complex syntactic structures. Different processing accounts (Levelt 1989, Pickering and Garrod 2004, Costa et al. 2008), perceived task difficulty judgments, as well as the comparison with L1-speakers' performances, all indicate that monologues are cognitively more complex than dialogues.

5 Conclusion and Directions for Future Research and Practice

To conclude, the study at hand gives only little support to the claims of the Cognition Hypothesis (Robinson, 2005, this volume): increased cognitive task complexity manipulated on the factor +/- few elements resulted in a higher lexical complexity only, while no combined effects of task complexity and interaction supporting the Cognition Hypothesis were found. The expectations about the effects of interaction on its own, however, were largely confirmed: interactive tasks did promote lexical complexity, accuracy, and fluency for L2-learners, but lowered the structural complexity.

Based on the points discussed in section 4 the following three conclusions are formulated:

Firstly, this study manipulated different factors of task design (+/- few elements and +/- monologic) on their own as well as in combination. This systematic approach to investigate cognitive and interactive variables of the Triadic Componential Framework in isolation, but also the focus on the interplay of the two task variables, allows powerful interpretations of attentional allocation during task performance. Increased cognitive task complexity on the single factor +/- few elements as implemented in this study only yielded minor differences in the lexical complexity for L2- and L1-task performance. In contrast, interaction yielded large effects on all measures of task-based performance. It seems that this cognitive factor is not a task characteristic that substantially affects the allocation of attention. Rather than a qualitative change of linguistic behavior, the factor +/- few elements presumably affects the speech of L2-learners in a quantitative way only, i.e., 'more of the same language'.

Whereas the present study investigated L2-task performance at one single moment in time, the Cognition Hypothesis aims foremost at being a guideline for task sequencing and syllabus design for interlanguage development (Robinson and Gilabert 2007). By continuing to investigate the effects of different task characteristics on their own as well as in combination, future research should include aspects of task sequencing and long-term effects over time. In addition, this would give the research a more classroom oriented perspective.

Secondly, interactive tasks push L2-learners to greater accuracy, lexical complexity, and fluency while monologues give speakers the opportunity to build complex syntactic structures. For the practice of language teaching, this study reveals that both monologues and dialogues are valuable settings for L2-production, as they both promote different aspects of oral L2-performance. Language teachers and testers however should be aware of these differences, because L2-learners are often evaluated in a monologic setting, especially in a testing environment. This apparently is likely to give an underestimated picture of the L2-learner's competence with respect to lexical complexity, accuracy, and fluency.

Thirdly, the interpretation of learner speech in light of a native speaker baseline gives valuable insights into the different processes speakers are involved in when they perform oral tasks. From a cognitive perspective on L2-task performance, the results of this comparison suggest that we need to change our theoretical framework. Based on psycholinguistic models of language processing (Levelt 1989) and the Alignment Hypothesis (Pickering and Garrod 2004, Costa et al. 2008) the present study predicts that besides being an interactive variable the factor +/- monologic has a cognitive impact on L2-task performance.

References

Cadierno, T. and Robinson, P. (2009). Language typology, task complexity and the acquisition of L2 lexicalization patterns for reference to motion. *Annual Review of Cognitive Linguistics*, 7.

Costa, A., Martin, J. P., and Sorace, A. (2008). Alignment in second language dialogue. *Language and Cognitive Processes*, 23(4):528–556.

de Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics*, 13(1):1–24.

Dörnyei, Z. and Kormos, J. (1998). Problem-solving mechanisms in L2 communication. A psycholinguistic perspective. *Studies in Second Language Learning*, 20:349–385.

Ferrari, S. (2009). *Valutare le competenze orali in italiano L2: variazione longitudinale e situazionale in apprendenti a livello avanzato*. Unpublished Ph.D. Thesis, Università degli studi di Verona.

Fiksdal, S. (2000). Fluency as a function of time and rapport. In Riggenbach, H., editor, *Perspectives on fluency*, chapter 8, pages 128–140. University of Michigan Press, Michigan.

Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In Bygate, M., Skehan, P., and Swain, M., editors, *Researching pedagogic tasks: Second Language Learning, Teaching, and Testing*, pages 75–93. Longman, Harlow.

Foster, P. and Tavakoli, P. (2009) Native Speakers and Task Performance: Comparing Effects on Complexity, Fluency, and Lexical Diversity. *Language Learning* 59 (4), pages 866 – 896.

Foster, P., Tonkyn, A., and Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3):354–375.

Gass, S. M. and Mackey, A. (2006). Input, interaction and output. *AILA Review*, 19:3–17.

Gass, S. M., Mackey, A., and Pica, T. (1998). The role of input and interaction in second language acquisition: Introduction to the special issue. *Modern Language Journal*, 82(3):299–307.

Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics*, 45(3):215–240.

Gilabert, R., Baron, J., and Llanes, M. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics*, 47(3-4):367–395.

Givon, T. (1995). *Functionalism and Grammar*. John Benjamins Publishing Company, Amsterdam/Philadelphia.

Guiraud, P. (1954). *Les caractères statistiques du vocabulaire*. Presses universitaires de France, Paris.

Halford, G. S., Cowan, N., and Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences*, 11(6):236–242.

Jacobs, D. (2006). *Scanner recorder* 1.9.

Kim, Y. (2009). The effects of task complexity on learner-learner interaction. *System*, 37(2):254–268.

Kormos, J. (2000). The timing of self-repairs in second language speech production. *Studies in Second Language Learning*, 22(2):145–167.

Kuiken, F. and Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics*, 45(3):261–284.

Levelt, W. J. (1989). *Speaking: From intention to articulation*. MIT Press.

Long, M. H. (1990). Task, group, and task-group interactions. In Sarinee, A., editor, *Language Teaching Methodology for the Nineties: Anthology Series*, volume 24.

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

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*.

Michel, M. C. (in preparation). Effects of task complexity on the use of Dutch connectives (working title).

Michel, M. C., Kuiken, F., and Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics*, 45(3):241–259.

Navon, D. (1989). The importance of being visible: On the role of attention in a mind viewed as an anarchic intelligence system. *European Journal of Cognitive Psychology*, 1:191–238.

Nuevo, A.-M. (2006). *Task complexity and interaction*. Unpublished Ph.D. Thesis, Georgetown University.

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30:590–601. Special Issue, Complexity, accuracy, fluency (CAF) in Second Language Acquisition.

Pica, T. (1994). Research on negotiation: What does it reveal about second language learning conditions, processes, and outcomes? *Language Learning*, 44:493–527.

Pickering, M. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27:169–226.

Poulisse, N. and Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15:36–57.

Revesz, A. (2008). Task complexity, focus on form-meaning connections, and individual differences: A classroom-based study. *Paper presented at the AILA-conference*, Essen, Germany.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning*, 45(1):99–145.

Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1):27–57.

Robinson, P. (2003). Attention and memory during SLA. In Doughty, C. and Long, M. H., editors, *The Handbook of Second Language Acquisition*, pages 631–678. Blackwell Publishing Ltd, Malden.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, 43:1–32.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics*, 45(3):193–213.

Robinson, P. (in preparation). Second language task complexity, taxonomic description, the Cognition Hypothesis and instructional design. In Robinson, P., editor, *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*. John Benjamins Publishing Company, Amsterdam/Philadelphia.

Robinson, P. and Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *International Review of Applied Linguistics*, 45(3):161–176.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11:17–46.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press, Oxford.

Skehan, P. (2001). Tasks and language performance assessment. In Bygate, M., Skehan, P., and Swain, M., editors, *Researching pedagogic tasks: Second language learning, teaching and testing*, chapter 8, pages 167–185. Longman, Harlow.

Skehan, P. and Foster, P. (2001). Cognition and tasks. In Robinson, P., editor, *Cognition and Second Language Learning*, pages 183–205. Cambridge University Press, Cambridge.

Skehan, P. and Foster, P. (2007). Complexity, accuracy, fluency, and lexis in task-based performance: A meta-analysis of the Ealing Research. In van Daele, S., Housen, A., Kuiken, F., Pierrard, M., and Vedder, I., editors, *Complexity, Accuracy and Fluency in second language use, learning and teaching*, pages 207–226. Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten (KVAB), Brussel.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass, S. and Madden, C., editors, *Input and second language acquisition*, pages 64–81. Newbury House, Rowley, MA.

Swain, M. and Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16:371–391.

Tavakoli, P. and Foster, P. (2008). Task design and L2 performance. *Language Learning*, 58(2):429– 473.

Tavakoli, P. and Skehan, P. (2005). Strategic planning, task structure, and performance testing. In Ellis, R., editor, *Planning and task performance in a second language*, volume 11, pages 239–273. John Benjamins Publishing Company, Amsterdam/Philadelphia.

VanPatten, B. (1990). Attention to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12:287–301.

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17(1):65–83.

Wickens, C. D. (1998). Attention and skilled performance. In Holding, D., editor, *Human Skills*, pages 71–105. John Wiley, New York.

Wickens, C. D. (2007). Attention to the second language. *International Review of Applied Linguistics*, 45(3):177–191.

Appendix

FIGURE 1 ABOUT HERE

Example instruction: complex – dating – monologue

There is a new dating show on TV. Six contestants receive a couple of minutes to get to know each other. Afterwards they have to choose someone they like. Only if a man and a woman choose each other do they win a sailing trip to Spain. The audience at home can also win a prize: if you predict the winning couple you will receive 500 euro. Together with a friend you decided to take part in the viewer competition. In two minutes from now you will call him/her to give your opinion. As your friend will not be at home, you have to leave a message on her/his voice mail.

You have just read the descriptions of the candidates. Look at them again and make a decision about which two out of them (a man and a woman) make a good couple and therefore are likely to win the show. When you call your friend you have 3 minutes time to explain in detail who you would choose. Make sure you have good reasons and are able to explain why your choice is the best. Also explain why other couples are less likely to win. Note: you and your friend will have to agree on one couple in the end. Thus, take your time and have a convincing story for your friend.