

The role of task complexity, modality and aptitude in narrative task performance

Judit Kormos and Anna Trebits

Abstract

The study reported in this paper investigated the relationship between components of aptitude and the fluency, accuracy, syntactic complexity and lexical variety of performance in two types of written and spoken narrative tasks. We also addressed the question of how narrative performance varies in tasks of different cognitive complexity in the written and spoken modes. Our findings indicate a complex interaction between aptitude components and task performance under different conditions. The components of aptitude that seemed to be most strongly related to the accuracy and complexity of production were deductive ability and grammatical sensitivity. The results also show that in writing the participants used more varied vocabulary than in speech, but their performance was similar in terms of syntactic complexity.

INTRODUCTION

Research in the field of second language acquisition (SLA) has long been concerned with the question of why students show great variation in their language learning success. Studies in this area have concluded that individual differences (IDs) are the most important predictors of achievement in a second language (L2) (Dörnyei, 2005). Therefore, it is widely acknowledged that IDs have to be taken into consideration both in theoretical accounts of SLA and in practical pedagogical decision-making. Researchers often point out the necessity of making further advances into uncovering how certain IDs affect and underlie important language learning processes (Dörnyei, 2005; Kormos & Sáfár, 2008; Robinson, 2007a). Learner variables are usually divided into cognitive, affective and personality-related IDs (Dörnyei, 2005). One of the most important cognitive variables influencing the success of L2 learning is foreign language aptitude, which is assumed to be a conglomerate of cognitive factors that might potentially affect the success of language learning (Carroll, 1981). For a long time researchers were interested in the link between foreign language aptitude and global language learning outcomes (for reviews see Ehrman & Oxford, 1995; Grigorenko, Sternberg & Ehrman, 2000). In recent conceptualizations of foreign language aptitude, however, it is argued that different cognitive abilities might be useful in different phases and processes of language learning (Skehan, 2002) and that learners with different cognitive ability profiles might benefit from different types of learning tasks and instructional conditions (Robinson, 2005a).

The novelty of our study is that we investigated how various components of language aptitude are related to performance in narrative tasks that differed in their cognitive complexity in both oral and written modes. For this purpose we administered the Hungarian version of the language aptitude test and two oral and written tasks to 44 upper-intermediate learners of English in a Hungarian secondary school. Most studies examining the subtle effects of task characteristics on L2 output conclude that in addition to general measures of L2 production, task-specific measures of production reveal more precise information about how tasks can direct learners' attention to certain

linguistic forms and how IDs may differentiate the ways in which learners can benefit from the manipulation of certain task features (e.g., Kormos & Trebits, in press; Norris & Ortega, 2009; Robinson 2007b). Therefore, in our study we used both general and specific measures of performance in analyzing students' output.

The present study examines foreign language aptitude in relation to the cognitive demands of performing oral and written narrative tasks of different cognitive complexity. Therefore, we first discuss conceptualizations of foreign language aptitude and then review research on communicative tasks and task features with an emphasis on the construct of task complexity. Next, we give a brief summary of research on modality differences in task performance. Finally, we present the results of previous studies which have investigated the influence of IDs on language produced in different tasks.

Language learning aptitude

Foreign language (L2) aptitude is considered to be one of the best predictors of language learning success (Abrahamsson & Hyltenstam, 2008; Dörnyei, 2005; Nikolov & Ottó, 2006; Sawyer & Ranta, 2001). Carroll (1981) identified four components of language aptitude: i) phonetic coding ability, that is, the “ability to identify distinct sounds, to form associations between those sounds and symbols representing them, and to retain these associations”; ii) grammatical sensitivity, that is, the ability “to recognize the grammatical functions of words (or other linguistic entities) in sentence structures”; iii) rote learning ability, defined as “the ability to learn associations between sounds and meanings rapidly and efficiently, and to retain these associations”; and iv) deductive learning ability, which is “the ability to infer or induce the rules governing a set of language materials, given sample language materials that permit such inferences” (p. 105). Instruments developed to measure language aptitude, such as the Modern Language Aptitude Test (MLAT) (Carroll & Sapon, 1959) and Pimsleur's Language Aptitude Battery (PLAB) (Pimsleur, 1966), test language learners on the

above mentioned four components (for a review see Abrahamsson & Hyltenstam (2008) and Robinson, (2007a)).

The Carrollian construct of language aptitude was developed in the period when audiolingualism was the prevailing language teaching methodology and when language competence was mainly conceived of as grammatical knowledge. New language teaching methodologies (in particular communicative and task-based approaches) aim to foster the acquisition of communicative competence, which encompasses a wider range of knowledge and abilities than linguistic competence alone. These changes in language teaching methodology and in the conceptualization of the goals in language learning have also necessitated a revision and extension of the concept of aptitude. In response to this, Robinson (2005a) proposed a dynamic aptitude construct, in which cognitive resources and abilities are combined into aptitude complexes. In Robinson's (2005b) model, primary abilities include pattern recognition, speed of processing in phonological working memory and grammatical sensitivity. These general cognitive abilities, which with the exception of phonological memory, are based on the traditional construct of aptitude, help the so-called second-order abilities. The second-order abilities are specific to language learning and include noticing the gap, memory for contingent speech, deep semantic processing, memory for contingent text, and metalinguistic rule rehearsal.

Robinson (2005b) also argued that explicit and implicit learning conditions might require different combinations of cognitive abilities. In his study, which investigated the role of aptitude in different learning conditions, he found a weak link between aptitude and outcomes in implicit learning conditions and a strong relationship between aptitude and learning under explicit conditions. De Graaff's (1997) research, however, indicated that grammatical sensitivity and the ability to infer the meanings of words from a text were positively related to attainment in an artificial grammar learning experiment under both explicit and implicit conditions. Skehan (2002) also suggested that certain components of the traditional construct of aptitude such as grammatical

sensitivity and deductive ability might assist L2 learning in naturalistic contexts, where learners have few opportunities to acquire L2 linguistic rules through explicit explanation.

From this brief review of recent research on aptitude it is apparent that several of the components of the original construct of aptitude as defined by Carroll might be relevant underlying cognitive abilities that promote language learning success even in today's foreign language classrooms. Deductive learning ability and grammatical sensitivity might help learners recognize linguistic patterns in the communicative input, whereas rote learning ability might be one of the significant predictors of the success of vocabulary acquisition. Phonetic coding ability might also play an important role in the acquisition of the phonological system of the L2 and in L2 reading, in which one of the key abilities is phonological awareness (for a recent review see Grabe, 2009). It seems to be important to examine, however, not only how aptitude contributes to the overall success of language learning, but also to analyze how students with different cognitive abilities perform in communicative tasks that they commonly encounter in the process of L2 learning.

Task complexity

Cognitive abilities do not only assist in the acquisition of L2 knowledge, but they might also influence how learners utilize their acquired knowledge in performing different types of tasks. Consequently, it is important to investigate the relationships between the combination of abilities underlying L2 aptitude and the cognitive demands of pedagogic tasks (Robinson, 2005a). Such research aims to contribute to an understanding of how best to match learners with strengths in certain cognitive abilities to particular types of learning tasks by examining the interplay between the processing demands of pedagogic tasks and the components of aptitude. As cognitive processes may be closely linked to various features of task design (e.g., the availability of planning time, Ellis, 2005; Foster & Skehan, 1996; Yuan & Ellis, 2003, or the number of elements included in the task, Kuiken & Vedder, 2007), the findings of research on the interaction of individual difference factors

and task complexity may be used to inform pedagogic decisions in materials development and syllabus design.

One of the most important issues in research on language learning tasks concerns the influence of task complexity on task performance. In Robinson's (2001b) definition, "task complexity is the result of attentional, memory, and other information processing demands imposed by the structure of the task on the language learner" (p. 29). Robinson lists a number of task characteristics such as the number of elements, availability of planning time and prior knowledge, which influence the complexity of the task. In the case of speaking tasks, however, it would also be important to relate the complexity demands of tasks to the different stages of speech production. In Levelt's (1989) model, speech production has four important components, which follow each other in this order: (1) *conceptualization*, that is, planning what one wants to say; (2) *formulation*, which includes the grammatical, lexical and phonological encoding of the message; (3) *articulation*, in other words, the production of speech sounds, and (4) *self-monitoring*, which involves checking the correctness and appropriateness of the produced output. In first language (L1) production conceptualizing the message requires attention, whereas formulation and articulation are automatic, and hence processing mechanisms can work in parallel, which makes L1 speech generally smooth and fast. In the case of non-balanced bilinguals and less proficient L2 speakers, however, formulation and articulation are often not sufficiently automatic and require conscious attention, which frequently hinders parallel processing.

Task complexity is generally considered to derive from the cognitive demands a task makes in the conceptualization stage. This view is based on the assumption that complex concepts require the use of complex syntactic structures, and therefore cognitively complex tasks are complex both in terms of conceptualization and linguistic formulation (see e.g. Robinson's (2001b, 2003, 2005b) Cognition Hypothesis). It is, however, possible that tasks make separate and independent complexity demands on the conceptualization and formulation stage. If we consider two speaking tasks such as the ones used in the current study — a cartoon description task, in which the storyline

is given, and a picture narration task, in which students have to narrate their own story — we can see that different aspects of cognitive complexity make different demands on the learners. In a cartoon description task, students do not need to conceptualize the content of the story, which eases the demands in the phase of conceptualization, that is, in selecting and ordering the relevant concepts of their message (Levelt, 1989; Skehan, 2009). Nevertheless, learners have to express the content prescribed by the task in whatever linguistic resources they have available in the L2, which increases the processing load in the linguistic encoding phase of speech production. In a story narration task, however, learners need to design their own story, but they can tailor it to match their linguistic resources, which results in increased conceptualization effort and a potentially reduced load in linguistic encoding. This example illustrates that in certain cases it might be difficult to order tasks in terms of cognitive complexity because tasks might make different and non-comparable demands on different phases of language production (see Pallotti (2009) and Skehan, (2009) for a similar line of argument). Consequently, it might increase our understanding of task complexity, if the cognitive demands of tasks were also considered separately for the conceptualization and linguistic encoding stages of speech production.

A key issue in task-based language learning is that in performing a task students need to coordinate the allocation of their attention in order to successfully meet the linguistic demands of the task. This raises questions concerning how attentional resources can be used, coordinated and directed to different aspects of language production during task completion. There are two influential models of task complexity in this field, which have motivated a great number of studies that examine the effects of manipulating the different dimensions of task complexity on L2 output and interaction, and ultimately on L2 learning: Skehan and Foster's (2001) Limited Attentional Capacity Model and Robinson's (2001b, 2003, 2005b) Cognition Hypothesis. These models make contrasting predictions as to the effect of increasing task complexity along various dimensions on L2 performance. Skehan and Foster's (2001) Limited Attentional Capacity Model views attention and memory as limited in capacity; therefore, they suggest that increasing task complexity reduces

the pool of available attention and memory resources. As a result, some aspects of performance will be attended to while others will not. Skehan and Foster also claim that cognitively more demanding tasks draw learners' attention away from linguistic forms so that enough attention can be paid to the content of the message (for a recent account of the model see Skehan, 2009).

Skehan and Foster's (2001) model, however, does not consider that attention is selective and voluntary, in other words that one can have volitional control over choosing relevant stimuli and ignoring irrelevant ones (Allport, 1987; Wickens, 2007). Robinson's (2001b, 2003, 2005b) Cognition Hypothesis differs from the Limited Attentional Capacity Model in that it assumes that attention is subject to voluntary regulation. In the Cognition Hypothesis, two sets of dimensions of cognitive task complexity are distinguished: resource-directing and resource-dispersing dimensions. The resource-directing dimensions of task performance call learners' attention to the linguistic features which are needed to meet task demands (e.g., reference to events happening here and now vs. to events taking place then and there), whereas the resource-dispersing dimensions of the task act as attentional limitations in determining what aspect of the task can be heeded (e.g., reducing the pre-task planning time). The Cognition Hypothesis states that sequencing tasks from cognitively simple to complex allows students to progress towards successfully performing real-world target tasks. Robinson proposes that increasing task complexity along resource-directing dimensions can lead to greater accuracy and grammatical complexity of L2 output because such demands can direct learners' attention to how the concepts and functions required by the task have to be grammaticized using specific linguistic forms. Increasing complexity along resource-dispersing dimensions, however, depletes learners' attention without having the beneficial effect of directing it to any specific linguistic aspect of L2 production.

The hypotheses put forward by the Limited Attentional Capacity model and the Cognition Hypothesis have received mixed support, which is probably due to the fact that in some tasks, certain characteristics make resource-dispersing attentional demands on learners, whereas others simultaneously draw their attention to certain linguistic aspects of performance. Few studies have

examined systematically the combined effects of resource-directing and resource-dispersing variables on performance (but see Gilabert, 2005; Iwashita et al., 2001; Révész, 2009) although in real-world communicative tasks these two dimensions simultaneously affect performance (for a recent discussion of this issue see Pallotti, (2009)). In our research we did not intend to test the predictions of the two conflicting models of task complexity, but we were interested in comparing two types of tasks which make different conceptualization demands on L2 speakers: a narrative task, in which the learners need to design the plot of the story, and another type of narrative task, in which they need to linguistically encode a pre-determined storyline. Our aim in using tasks with different attentional demands at different stages of speech production was to extend Robinson's Cognition Hypothesis (2001b, 2003, 2005b) with respect to complexity demands concerning conceptualization and linguistic encoding.

Written vs. spoken task performance

In addition to complexity, another key factor to consider with regard to tasks is the mode of performance. Task-based language learning research has traditionally focussed on speaking tasks, and considerably fewer studies have investigated how tasks affect second language writing processes and the product of writing. There are a number of important differences between speaking and writing (Biber, 1988; Chafe, 1982), but from the point of view of psycholinguistic processing, the most crucial ones are that writing is usually not as constrained by time and that it is a recursive process, in which writers plan, linguistically encode their plans and revise them cyclically (Grabe & Kaplan, 1996). Although the availability of time is not unlimited in writing, writers are under somewhat less pressure than speakers to divide their attention between conceptualizing their message and linguistically encoding it, which allows for extensive on-line planning, that is, planning the content of the output while giving it a linguistic form (Yuan & Ellis, 2003). In writing, the time spent on planning the message (i.e. pre-task planning) is also integrated in the writing

process because writers can devote considerable time to planning the content before starting to write. Writers also have more attentional resources available for monitoring their output while they are encoding it than speakers.

Only a few studies have addressed how the mode of performance influences the output students produce on specific language learning tasks. In a study conducted with learners of French, Granfeldt (2008) found that mode did not have an effect on syntactic complexity, but students used a higher variety of words in writing and were more accurate in speaking than in writing. In Kuiken and Vedder's (2009) research, however, Dutch learners of Italian produced syntactically more complex language in writing than in speech, but the lexical variety of their spoken and written output did not differ significantly. The contrasting patterns of the findings might be related to the research design of these studies (i.e., different types of tasks used) and to the fact that different L2s and learners with different levels of proficiency were investigated.

In our study, we intended to investigate the effect of mode on performance in two types of tasks which pose different conceptualization demands. In our analyses, we focused on differences in lexical variety and in global as well as task-specific measures of syntactic complexity and accuracy. Furthermore, we were also interested in how learners with different cognitive ability profiles perform in these two modes and how various aptitude components are related to the linguistic quality of the output in speech and writing.

Individual differences and task performance

Individual differences can exert both direct and indirect influence on task performance. Individual difference factors such as anxiety, working memory capacity and aptitude might have a direct effect on students' decisions concerning the allocation of their attentional resources and on students' ability to handle their attentional limitations. Individual differences, especially those in cognitive abilities, might also influence how successfully students acquire particular aspects of linguistic

competence. As any performance in a task draws on acquired linguistic abilities, differences in underlying competence caused by individual difference factors are also expected to manifest themselves in task performance.

Few studies have addressed the issue of how IDs affect communicative task performance in general and even fewer studies exist that investigate how the different components of language aptitude relate to L2 production on tasks of differing cognitive complexity. Robinson (2007b) analyzed how input, processing and output anxiety affected students' performance on narrative tasks that increased in cognitive complexity. His findings indicated that as the tasks became more complex, the negative correlation between output anxiety and syntactic complexity grew stronger. Robinson's research, however, did not reveal any significant link between anxiety and accuracy and fluency of task performance. Niwa (2000) also investigated the relationship between a resource-directing task aspect of task complexity (simple vs. complex reasoning demands) and three ability variables: intelligence, aptitude and working memory. She found that in the cognitively more complex task students with higher aptitude and working memory spoke less fluently. She explained her findings by arguing that students with high working memory capacity made greater efforts to meet the reasoning and linguistic demands of the more complex task, which negatively affected their fluency. Kormos and Trebits (in press) also studied the effects of working memory capacity on performance on the same narrative tasks as used in the current study. The finding that students with high working memory capacity produced long clauses, which were, however, syntactically less complex indicated that working memory plays a complex role in task performance. We tentatively argued that high working memory capacity might allow students to produce narratives with high clausal complexity, but it might not be conducive to directing learners' attention to specific dimensions of the task such as subordination.

The aims of our study

As the review of literature shows, there is a scarcity of studies that investigate the role of cognitive factors in communicative task performance, and no research has been done on the differential effect of these individual variables on the quality of both written and spoken output. Therefore, one of the questions that we addressed in our study concerned how different components of language aptitude are related to Hungarian-speaking EFL learners' performance on oral and written narrative tasks. We were also interested in how individual differences in language aptitude manifest themselves in tasks which make different conceptualization demands on the learners. For this purpose we used two narrative tasks which differed with regard to whether the students were required to devise the plot of the story. Language aptitude was assessed using a Hungarian version of the MLAT (HUNLAT: Ottó, 2002) based on the consideration that the constructs operationalized in this test battery, especially grammatical sensitivity and deductive learning ability, have been found to be relevant both under explicit and implicit language learning conditions, a mixture of which characterizes current foreign language learning contexts.

In our study we also aimed to discover differences in linguistic performance measures in the above-described task types across modes. Our goal was to gain further insights into differences in spoken and written performance of L2 learners by extending the range of variables that are traditionally included in the analysis of task performance. We based our decision of selecting task-specific variables on three sources of information. First of all, we conducted an initial analysis on the tenses of verbs used by the learners, which suggested that a high percentage of the participants used the past tense to narrate their stories and that type of subordinate clause used with the highest frequency was the relative clause.¹ Second, we administered the same tasks to 10 young native speakers of British English and analyzed their performance for these features. The analyses revealed that 80% of the native speaker participants primarily used past tense for narrating the cartoon description task in speech, and all the native speakers used past tense for describing events in the oral story narration task and in both tasks in writing. The relative clause was also the most frequent type of subordinate clause in the output of the native speakers. The selection of past tense verbs and

relative clauses was also based on theoretical considerations. Although historic present is frequently used in narratives by native speakers, studies in the field of discourse analysis suggest that its use is restricted to adding a dramatic emphasis to the story and moving the events out of the past to the moment of speaking (Schiffrin, 1981). Historic present mainly occurs in situations where speakers narrate a personal story and is almost exclusively applied in describing complicating events (Schiffrin, 1981). The task used in our study did not involve the narration of personal stories, and participants rarely added complicating events; consequently, we deemed it appropriate to select past tense as the major tense in which the story is expected to be narrated. Nevertheless, our analysis focussed on target-like-use and not on suppliance in obligatory contexts (Pica, 1984), which allowed us to consider the use of present as an additional acceptable tense.

Relative clauses also have several important discourse and linguistic functions in narratives, namely to name, situate and identify old and new referents in the story, to present main characters, to motivate, enable and continue narrative actions, to set up expectations about narrative entities and events and to sum up past or upcoming events (Dasinger & Toupin, 1994). They have been found to be frequent in narratives produced by adult and child L1 speakers (Dasinger & Toupin, 1994). Relative clauses have increased importance in learning English as an L2 because they are difficult to acquire and use accurately (Pienemann, 1998), and consequently might provide relevant information on the complexity and accuracy of L2 performance.

Based on the above-described theoretical considerations and on the results of our preliminary analyses, the study addressed the following research questions:

1. How is Hungarian L2 learners' narrative performance affected by mode and task type?
2. How do the components of language aptitude relate to general and task-specific measures of accuracy, syntactic complexity, lexical variety and fluency when performing two different types of oral narrative tasks?
3. How do the components of language aptitude relate to general and task-specific measures of accuracy, complexity and lexical variety in two different types of written narrative tasks?

METHOD

Participants

The present study was conducted in a Hungarian-English bilingual secondary school for students aged between 15 and 18 in Budapest, Hungary. The participants (N=44) were students in the second academic year of a bilingual education program which consists of a so-called zero year and four years of bilingual secondary education. During the zero-year the students took part in an intensive English language training program, which aimed to prepare them for studying several school subjects in English in the following four years of secondary school. The teaching method used was predominantly communicative. The participants in our study completed the zero-year program in June 2006 before starting their secondary studies in September 2006. At the time of the data collection for the present research, they had just begun the second academic year of their studies. The participants' age was between 16 and 17 years. 27 students were female and 17 male. The teachers of the students rated the participants' level of proficiency as slightly above intermediate corresponding to B1/B2 in the Common European Framework of Reference (Council of Europe, 2001), which was also supported by the results of the language proficiency test that they administered to the students at the end of their zero year.

Instruments

The within-participants factors of the study are task type (cartoon description vs. picture narration) and mode (oral vs. written) while the components of language aptitude (as measured by the Hungarian Language Aptitude Test (HUNLAT)) constitute between-participants factors. Participants completed four narrative tasks: two involving cartoon description, and two involving picture narration (see Appendix). First they performed a cartoon-description task and a picture narration task orally (in random order). Then, a month later, they performed parallel versions of these tasks in writing (in self-chosen order, due to the limitations of the group administration

procedures). The cartoon description task involved the description of a comic strip consisting of six pictures, which had to be included in the story. The pictures were presented in the correct order and formed a coherent story line. The input to the task was provided visually, with instructions in Hungarian. In the oral version of the task the students had to narrate the story of a ship-wrecked man, who is not recognized as needing help and who in the end remains stranded on a desert island. In the written version of the task, a car which is broken down in the middle of the desert is transformed into a carriage without a horse by a wizard instead of being repaired. The stories included the same number of actors and key narrative events and both had an element of surprise. This type of task did not require the conceptualization of the plot, and was consequently considered to place a relatively low cognitive load on the participants in terms of conceptualizing their message, but as argued above, might have made greater demands on the learners in terms of linguistic encoding.

The picture narration required students to tell a story based on six unrelated pictures, all of which had to be included in the narrative. The pictures were selected carefully to include similar elements in the two versions of the task. In both versions of the task there was an object (book vs. ring), a picture depicting an adverse weather condition (a storm with lightening vs. a storm at sea), a means of transport (boat vs. airplane), a picture showing a geographical location (mountains vs. an island), a house (in the middle of a forest vs. in a town) and a door (locked vs. open). In order to successfully complete this task, the participants not only had to rely on their language skills, but they also had to use their imagination and find a way to relate the pictures to one another and invent a story around them. As argued above, this task can be characterized as cognitively more complex in terms of conceptualization and might pose smaller processing demands in linguistic encoding than the cartoon description task.

In order to avoid task repetition effects, these parallel versions of the cartoon description and picture narration task were administered to the students orally and in writing. The written and oral tasks only differed in the picture cues; as explained above, the structure of the task was exactly the

same. In a previous research project, Albert (2007) validated the two forms of the cartoon description and picture narration task orally and found no significant differences in any of the linguistic variables in the two versions of the tasks when performed in speaking.² We assumed that due to the fact that the two tasks were found to be parallel in a similar population when performed in speech, any differences between the cartoon-description versus picture-narration tasks when administered in writing could be attributed to the effect of mode.

As already mentioned, the participants' foreign language aptitude was measured using the HUNLAT (Ottó, 2002) which consists of four subtests described below (Table 1 contains the descriptive statistics of the HUNLAT scores for the participants in the study). The HUNLAT is a test battery for the Carrollian concept of language aptitude (1981), which measures four constructs: phonological sensitivity, deductive language learning ability, grammatical sensitivity, and rote-learning ability (for information on the validity of the test see Nikolov & Ottó, 2003; Hild, 2007).

Insert Table 1 around here

1. Phonological sensitivity

This test is a modified version of the 'Phonetic script' task of the MLAT and intends to measure phonetic coding ability. Participants listen to 54 consonant-vowel-consonant sequences and follow the transcription of these sequences in the test booklet. Afterwards they hear 20 similar sound sequences, and they have to indicate on the answer sheet which one of four possible transcriptions corresponds to the given sound sequence. There is also a 'none' option. This subtest takes about 10 minutes.

2. Deductive ability

This subtest, which was adapted from Pimsleur's (1966) Language Aptitude Battery, aims to measure deductive language learning ability. Participants are given a set of words and sentences in an artificial language along with their Hungarian translations. On the basis of this information, they

have to translate 20 Hungarian sentences to the artificial language and choose the correct solution from four alternatives. Participants have 15 minutes for this subtest.

3. Grammatical sensitivity

This task is based on a similar subtest of the MLAT and is assumed to measure grammatical sensitivity. Participants have 10 minutes to solve 20 items, each consisting of two Hungarian sentences. One word is underlined in the first sentence. In the second sentence, five words are underlined, and participants have to choose the one that fulfils the same function as the underlined word in the first sentence.

4. Rote learning ability

This subtest, which was adapted from the relevant section of MLAT (Paired Associates), measures rote learning ability. Participants have 5 minutes to study a list of 24 Swahili words and their Hungarian equivalents. Then they have 10 minutes to choose the Hungarian equivalent of 20 Swahili words from five alternatives.

Procedure

The aptitude test was completed by the students at the beginning of the academic year. The oral narrative tasks were administered to the students individually in a quiet room at school in the third week of the school year. They were given two minutes to prepare for each task. The planning time was selected based on the piloting of the tasks with 5 students at a similar proficiency level. In the pilot study students were given longer planning time (5 minutes), but they usually started their stories after three minutes of planning. The order in which the students performed the oral tasks was randomized. The students' performance was recorded on a digital recorder, and the recordings were subsequently transcribed for analysis by a trained research assistant. After a one month interval, the participants completed two written narrative tasks one after the other during a regular English class.³ They had 30 minutes to do the two tasks, but most of them finished earlier. They were

instructed to write a minimum of 150 words for each task. The students could decide on the order in which they completed the tasks. Approximately 55% of the students started with the cartoon description task, and 45% of the students chose to write the story narration task first.

Analysis

The following general measures of linguistic performance were used. In order to assess lexical diversity, we applied Malvern and Richards' (1997) D-formula. The calculation of the D-value is based on a mathematical probabilistic model as operationalized in the VOCD software of the CHILDES database (<http://childes.psy.cmu.edu>), which uses random sampling of tokens in calculating the type-token ratio. Malvern and Richards (1997) argue that the D-value is a valid measure of lexical diversity because it does not depend on the length of the sample, and it uses all the words produced by the participants (for a discussion of various measures of lexical diversity see Jarvis, 2002).

Accuracy of general task performance was measured with the ratio of error-free clauses, which was calculated relative to the total number of clauses. Syntactic complexity was operationalized as the ratio of subordinate clauses, which was also expressed relative to the total number of clauses, and as the length of clauses, which was calculated as the number of words within a clause (Norris & Ortega, 2009). The measure of fluency was speech rate, which has been shown to be a reliable measure of fluency (Kormos & Dénes, 2004). Speech rate was calculated as the total number of syllables uttered by each student divided by the total amount of time spent speaking (including pause and hesitation time). These accuracy, complexity and fluency measures have been widely used in task-based research and have proven to reflect the characteristics of accuracy and grammatical complexity of students' output in a reliable manner (see Bygate, 1999; Robinson, 1995; Foster & Skehan, 1996; Skehan & Foster, 1997).

The selection of task-specific measures of performance was based on theoretical considerations and on the initial analysis of data as described above. Task-specific measures of

accuracy included the ratio of error-free relative clauses, which was calculated in proportion to the total number of relative clauses, as well as the ratio of error-free verbs and the ratio of error-free past-tense verbs. The ratio of error-free verbs was obtained by dividing the number of correctly used verb forms by the total number of verbs. The ratio of error-free past-tense verbs was expressed as the ratio of error-free past-tense verbs to the total number of past-tense verbs used. This analysis was a target-like use analysis (Pica, 1984), which focused on contexts where the structure was actually used by the learner and which involved making a decision on whether the structure was accurately used. The choice of this method of analysis was based on the fact that historic present and past might both be used in certain contexts, which does not allow for a reliable analysis of this structure using the method of obligatory occasion analysis. A task-specific measure of syntactic complexity was also chosen: the ratio of relative clauses, which was calculated by dividing the number of relative clauses by the total number of clauses. This measure has been used as an index in syntactic development in the production of narratives in L1 child language acquisition research (Dasinger & Toupin, 1994). Table 2 summarizes the performance measures used in this study. The two authors coded the accuracy measures separately and the average percentage of agreement was 97%.

For the statistical analysis SPSS 13.0 (Statistical Package for Social Sciences) was used. The statistical analyses performed were correlations and paired samples t-tests.⁴ In order to compensate for multiple testing, the level of significance for this study was set at $\alpha = 0.01$. Cohen's d value was used to measure effect sizes. D values below .5 indicate small, between .5 and .8 medium, and above .8 large effect size (Cohen, 1988).

Insert Table 2 around here

RESULTS

Table 3 shows the descriptive statistics of the means and standard deviations of the measures of lexical variety, complexity, accuracy and fluency. Here we will report only on the significant differences between measures of performance. The comparison of tasks in the two different modes of performance reveals that students used more varied vocabulary in the written mode both in the cartoon description ($t(43) = 6.96, p < .001, d = .72$) and in the picture narration task ($t(43) = 9.18, p < .001, d = .81$). The participants also produced a significantly higher proportion of error-free clauses ($t(43) = 3.27, p = .002, \text{Cohen's } d = .44$) and verbs ($t(43) = 5.08, p = .001, d = .61$) in the written cartoon description task than in the parallel oral task. In the oral mode of performance, participants used significantly more varied vocabulary in the cartoon description task than in the picture narration task ($t(43) = 3.46, p < .001, d = .44$). On the other hand, the ratio of error-free verbs was significantly lower in the oral cartoon description task than in the picture narration task ($t(43) = 3.46, p < .001, d = .44$). In the written mode, students produced significantly shorter clauses ($t(43) = 2.75, p = .001, d = .38$) and used more relative clauses ($t(43) = 2.72, p = .001, d = .38$) in the picture narration task than in the cartoon description task.

Insert Table 3 around here

Tables 4a and 4b display how aptitude scores as measured by HUNLAT are related to task performance in the oral mode. In the oral cartoon description task the ratio of error-free relative clauses correlated positively with grammatical sensitivity and the overall HUNLAT score, whereas in the picture narration task students with higher levels of grammatical sensitivity used relative clauses less correctly.⁵ In the oral picture narration task, deductive ability was negatively related to D-value. Tables 5a and 5b show the results of correlations in the written mode. As shown in Table 5a, students who scored high on the grammatical sensitivity sub-test produced longer clauses in the written cartoon description task. In the case of the written picture narration task, however, no

significant relationships emerged (see Table 5b). All the significant correlations between performance measures and aptitude components were moderately strong.

Insert Tables 4a, 4b, 5a, 5b around here

DISCUSSION

Differences between modes of performance

The results indicate that in writing students were more accurate and used more varied vocabulary. Our findings are partially similar to that of Granfeldt (2008), who also concluded that mode influenced the lexical variety of output; however, they contradict the results of his study, as well as that of Kuiken and Vedder (2009), by showing a significant increase in accuracy in writing in one of our tasks: the cartoon description task. The comparison of studies on this topic is very difficult due to the nature of different tasks used and to the differing level of participants' proficiency. Some parallels between Kuiken and Vedder's (2009) and our research might, however, be drawn if we examine the tasks used in the two studies. Kuiken and Vedder (2009) administered a task in which students had to justify their choices for a particular holiday destination. This task is similar to the story narration task used in our study, in that it is also high in conceptualization demands. Therefore, it might be possible that in tasks which require increased attention in terms of conceptualization, L2 learners do not seem to produce more accurate language in writing than in speech because their attentional resources are devoted to content planning rather than to encoding and monitoring linguistic form. The cyclical nature of writing, which theoretically would allow for a closer monitoring of accuracy than the on-line characteristics of speaking, might only increase accuracy in tasks such as our cartoon description task, which does not involve high conceptualizing demands and requires the linguistic encoding of specified content.

The findings concerning the higher variety of words in writing than in speech can be explained with reference to the availability of on-line planning time in writing, which might have allowed learners to avoid repeating the same words and helped them retrieve more varied vocabulary from their mental lexicon. An alternative explanation might be that when teaching L2 writing, teachers generally encourage higher lexical variety, and students might have been aware of this encouragement.

Our results concerning the lack of difference in syntactic complexity between writing and speech are similar to those of Granfeldt (2008), but are contradictory to those of Kuiken and Vedder (2009). The lack of effect of mode on syntactic complexity might be explained with reference to the resource-directing aspects of task complexity. As the two tasks were assumed to be parallel in writing and in speech, they might have directed learners' attention to similar syntactic features and consequently might have elicited similar levels of syntactic complexity.

Differences between task types

As regards the effect of task type in the different modes, the results indicate that the picture narration task in writing elicited syntactically more complex language, as assessed by clause length and the ratio of relative clauses, than the cartoon description task. This finding is seemingly in contrast with that of Kuiken and Vedder (2008), who found no effect of task complexity in the written mode. Kuiken and Vedder (2008), however, used only a subordination ratio as the measure of syntactic complexity. Based on our results concerning syntactic complexity in different types of tasks in the two modes of performance, we can hypothesize that in writing, where the resource-dispersing dimension of task complexity might play a different role, the picture narration task, which requires students to conceptualize their own stories, has the potential to direct students' attention to syntactic complexity. In the oral version of this task, however, the demand that students need to conceptualize and linguistically encode their narrative at the same time acts as a resource-

dispersion factor, and students do not seem to have sufficient attentional resources for producing syntactically complex language.

Task type was also found to have an effect on one specific aspect of accuracy and on lexical variety in the oral mode. On the one hand, the cartoon description task seems to have pushed students to use more varied vocabulary in order to be able to express the predetermined storyline depicted in the cartoon. At the same time students' accuracy in verb forms decreased in this task, which might seem to indicate a trade-off effect between lexical and grammatical encoding in the linguistic formulation phase of speech production. On the other hand, in the picture narration task, the students could conceptualize the story line taking their own linguistic resources into account and might have used vocabulary which was easily accessible from their mental lexicon. As a consequence, they might have had more attention available for the accurate linguistic encoding of verbs. The lack of significant differences in fluency between the two tasks also seems to lend support to this trade-off effect and might suggest that the overall processing load at the linguistic encoding stage of the tasks might have been similar. If our reasoning about conceptualization and linguistic encoding demands is on the right track, these findings indicate that it is also important to consider the cognitive aspects of task complexity in the linguistic encoding phase and not only at the conceptualization stage of speech production. Table 6 gives an overview of the psycholinguistic characteristics of the two tasks and summarizes the main findings in terms of lexical variety, accuracy and complexity.

The results might also indicate that task type effects manifest themselves differently in speech than in writing. In speech not only do learners need to divide their attention between conceptualization and linguistic encoding, but they also need to carry out linguistic encoding processes under time-pressure, which requires that they share attentional resources during lexical and syntactic encoding (see Table 6). As argued above, in the oral version of the cartoon description task students' attention might have been drawn to lexical encoding, and hence, they might have had less attention available for syntactic encoding. In writing, however, syntactic and lexical encoding

do not need to be carried out in parallel, which might result in improved accuracy in general and in the use of verb forms in particular in the written cartoon description task. This suggests that there might be an interaction between task type and mode, as tasks with different cognitive and linguistic demands seem to elicit different patterns of performance in writing than in speech. As a consequence, it might be difficult to draw general conclusions concerning the effect of mode on task performance without taking the characteristics of the task to be performed into consideration.

Insert Table 6 around here

The relationship between aptitude and oral task performance

Our findings indicate that students with a given cognitive profile do not perform consistently in narrative tasks. In the oral cartoon description task, where students were not required to conceptualize and linguistically encode their message at the same time, high grammatical sensitivity and high level of language learning aptitude might have helped students to successfully encode complex and difficult syntactic structures such as relative clauses. In the oral picture narration task, however, where the HUNLAT correlations with lexical diversity and the accuracy of relative clauses was negative, the task demands in terms of conceptualizing the content of the story might have caused difficulties to students with high grammatical sensitivity in employing their linguistic knowledge efficiently. We might argue that grammatical sensitivity is a cognitive ability that aids the conscious and explicit acquisition of complex syntactic knowledge (Robinson, 2005b) and that this ability might also help learners with high aptitude to consciously employ their explicit knowledge when they have sufficient attentional capacities that they can devote to linguistic encoding. This might be the case in the oral cartoon description task, in which students did not have to conceptualize the content of the narrative, and hence learners with high grammatical sensitivity might have been able to apply their explicit knowledge of relative clauses efficiently.

From the perspective of students who scored low on the grammatical sensitivity component of HUNLAT, the findings indicate that they might be less able to encode complex syntactic structures accurately in oral tasks in which the given content requires the use of specific syntactic structures and are thus high in linguistic encoding demands. They might, however, be advantaged in the oral picture narration task, in which they have the opportunity to avoid complex structures and in which attentional demands on conceptualization are high. A possible explanation for this might be that learners with lower levels of grammatical sensitivity might rely more on implicit knowledge and memorized chunks and thus might be more efficient in employing implicit knowledge of complex syntactic structures than students with high grammatical sensitivity.

It is interesting to note the negative relationship in the oral picture narration task between deductive ability and lexical variety as measured by the D-value. We might tentatively argue that learners with high deductive ability might have devoted their attentional resources to some other aspects of performance in this task. Niwa's research (2000) also found that in the cognitively more complex task students with high aptitude scores performed worse in fluency, which she explained with reference to the fact that they might have paid more attention to the linguistic and conceptual encoding required by the task. Although the correlation between deductive ability and ratio of error-free clauses ($r = .27$ $p = .07$) is not significant, its direction is positive. Despite being weak, this correlation might indicate that students with high level of deductive abilities prioritized accuracy over lexical variety.

Finally, we would like to point out a tendency-level correlation between the ratio of error-free past-tense verbs and rote learning ability in the oral cartoon description task ($r = .38$, $p = .02$). Rote learning ability helps in memorizing words and irregular morphological information. A great proportion of past-tense verb forms at the low-frequency level of vocabulary our participants tended to employ are irregular, and consequently rote-learning ability helps their accurate acquisition. The results indicate that students with high score in rote learning ability tended to be able to accurately

retrieve the past-tense verb forms in the oral cartoon description task, in which they had sufficient attention available for linguistic encoding.

The relationship between aptitude and written task performance

Similarly to the oral mode, grammatical sensitivity was found to be positively related to certain aspects of performance in the written cartoon description task, but in writing the negative relationship of aptitude components and accuracy and lexical diversity was not observed in the picture narration task. The finding that learners with high grammatical sensitivity produced longer clauses in the written cartoon description task might be related to the complexity demands of this writing task. This task relieved students from the cognitive load of having to conceptualize the story but made high demands on the participants in the linguistic encoding phase because they had to express a given content with their existing resources. The comparison of the mean values for clause-length in the two types of narratives also shows that students in general used shorter clauses in the picture narration task than in the story narration. Considering the linguistic encoding demands of the task, it seems that students with higher grammatical sensitivity might have devoted more attention to clausal complexity than learners with lower levels of cognitive ability who might have prioritized other aspects of linguistic processing such as the encoding of sub-ordinate clauses. This seems to be supported by the finding that grammatical sensitivity was negatively, although not significantly, correlated with the other two syntactic complexity measures including the ratio of subordinate and relative clauses. This pattern of correlations might suggest that in the cartoon description task students with high grammatical sensitivity tended to produce longer clauses at the expense of greater subordination complexity, whereas students with lower levels of grammatical sensitivity used more sub-ordinate and relative clauses, but produced shorter clauses.

We also need to account for the fact, however, that in the written picture narration task no relationship between aptitude and any of the performance measures was found. It was already

shown in the case of the oral version of this task that students with better cognitive abilities tended to perform worse in terms of the accuracy of relative clauses and D-value than participants with low aptitude scores. We explained this finding by arguing that the fact that the learners had to divide their attention between conceptualization and linguistic encoding in the oral picture narration task might have disadvantaged learners with high levels of grammatical sensitivity and deductive ability. Due to the availability of extensive on-line planning time and the cyclical nature of writing, sharing attentional resources between conceptualization and linguistic encoding in the writing phase might be less demanding than in speech, and therefore this task condition might not create a disadvantage for learners with high level of cognitive abilities.

Implications and limitations

In our research we investigated the relationship of traditional components of aptitude with the accuracy, fluency, syntactic complexity and lexical variety of performance in two types of narrative tasks, which were administered both orally and in writing to 44 upper-intermediate Hungarian learners of English. Our study also aimed to reveal how narrative performance varies in tasks of different cognitive complexity in the written and spoken modes. Our findings indicate that in writing the participants were more accurate and used more varied vocabulary than in speech, but their performance was similar in terms of syntactic complexity. The effect of task type on performance differed in the two modes. In speech students used fewer correct verb forms and more varied vocabulary in the cartoon description task than in the picture narration task. In writing, however, the picture narration task elicited syntactically more complex language than the cartoon description task. The results seem to lend support to Robinson's (2001b, 2003, 2005b) Cognition Hypothesis because in writing, where the resource-dispersing demands of tasks are reduced, the task that required complex cognitive planning at the level of conceptualization had the potential to direct learners' attention to syntactically more complex language. We also pointed out, however,

that it might be difficult to rank order pedagogic tasks in terms of cognitive complexity, and that cognitive demands of tasks should not only be considered at the conceptualization phase of speech production but also at the level of linguistic encoding.

Our results indicate that aptitude components were differently related to linguistic measures of oral performance than to those of written production, and that the strongest positive link between aptitude and linguistic measures manifested itself in the cartoon description task, where students did not have to conceptualize the storyline. The results of this study might provide support for Robinson's (2001a, 2005a) Aptitude Complexes Hypothesis as they reveal that aptitude components play various roles in influencing performance under different task conditions. From our results we can see that aptitude is not a unitary construct, and having high levels of cognitive ability does not necessarily result in enhanced quality of task performance. This suggests that varying instructional tasks in classroom settings as well as tasks used in language assessment is essential in order to give learners with different cognitive profiles equal chances to perform to the best of their potential.

Our findings also highlight the importance of examining task-specific measures of linguistic performance because most of the significant aptitude and task effects were found on variables which have not been used in task-based research before. Accuracy rates in selected syntactic structures across different types of tasks can be indicative not only of the state of language development but also of how different types of tasks succeed in eliciting accurate language use.

It has to be acknowledged that our study has a number of limitations. One short-coming of our research is that tasks were not counter-balanced in the different modalities. This might have induced an effect of the content of the task in the different modes of performance. An additional concern might be that the task-specific measures of performance targeted structures the occurrence of which was limited in number, and consequently a larger corpus of L2 data would be needed to verify our findings concerning the link between aptitude and the accurate use of relative clauses. The results of our study should also be interpreted with caution due to the characteristics of the

participants. The students who participated in our research had generally high scores in language aptitude and were enrolled in a highly intensive language training programme. These circumstances necessarily warrant replication with students who show more variation in their cognitive profiles and who have different levels of proficiency. As aptitude has been found to have differential influences across different instructional settings, further studies should also be conducted to examine the interaction of cognitive factors and language teaching methodology. Finally, we also have to note that our findings reveal that even relatively similar tasks which share the same genre might elicit remarkably different performance. Therefore, it is highly important to consider the psycholinguistic characteristics of spoken and written tasks at different stages of language production if we want to advance our knowledge of how tasks can elicit targeted aspects of language.

AUTHOR NOTE

Judit Kormos, Lancaster University, Department of Linguistics

Anna Trebits, Eötvös Loránd University, Budapest, Hungary, Department of English Applied Linguistics

Acknowledgements:

We are very grateful for the three anonymous reviewers for their thorough comments on previous versions of this manuscript. The first author was sponsored by the Bolyai scholarship of the Hungarian Academy of Sciences in writing this paper.

Contact addresses

Judit Kormos, Lancaster University, Department of Linguistics, Lancaster, LA1 4YL, United Kingdom, j.kormos@lancaster.ac.uk

Anna Trebits, Eötvös Loránd University, Budapest, Hungary, Department of English Applied Linguistics, Budapest, 1088 Rákóczi út 5. panni.trebits@yahoo.com

NOTES:

1. The proportion of past-tense verbs relative to the total number of finite verbs was 71% in the oral cartoon description task, 87% in the oral picture narration task, 95% in the written cartoon description task and 98% in the written picture narration task. 88% of students used at least one relative clause in the oral cartoon description task, 95% in the oral picture narration task, 80% in the written cartoon description task and 85% in the written picture narration task.
2. Albert (2007) administered all four tasks in speech to a group of participants, who were of comparable level of proficiency. She analyzed the parallel versions of the tasks in terms of discourse complexity (the number and type of narrative events) and linguistic variables (speech rate (as measured by syllables per second), accuracy (ratio of error-free clauses), syntactic complexity (subordination ratio) and lexical variety (D-value)). Neither the qualitative nor the quantitative analyses revealed any significant differences between the discourse structure and the linguistic features of performance in the two versions of the picture narration and cartoon description task.
3. The one month interval between the oral and written task was due to organizational reasons (we did not want to disrupt the teaching routine with a brief interval). The students received approximately 16 hours of classroom instruction between the two data collection sessions, which is not expected to result in substantial linguistic development during this period.
4. The aptitude sub-tests were not significantly correlated with each other, which allowed us to examine the effect of the components of the test on task performance independently of each other.
5. The correlation between the total aptitude score and the ratio of error-free clauses is an artefact of the high correlation of grammatical sensitivity and the ratio of error-free clauses because the grammatical sensitivity score contributes significantly to the total aptitude score ($r = .59$ $p < .0001$).

References

- Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in Second Language Acquisition*, 30, 481-509.
- Albert, Á. (2007). *Creativity and task performance*. Unpublished PhD dissertation. Budapest: Eötvös Loránd University.
- Allport, D.A. (1987). Selection for action: Some behavioral and neurophysiological considerations of attention and action. In H. Heuer & A.F. Sanders (Eds.), *Perspectives on perception and action* (pp. 395–419). Hillsdale, NJ: Lawrence Erlbaum.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Bygate, M. (1999). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research*, 3, 185-214.
- Carroll, J.B. (1981). Twenty-five years of research on foreign language aptitude. In Karl C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 119–154). Rowley, MA: Newbury House.
- Carroll, J. B., & Sapon, S. M. (1959). *The Modern Language Aptitude Test*. San Antonio, TX: Psychological Corporation.
- Chafe, W.L. (1982). Integration and involvement in speaking, writing, and oral literature. In D. Tannen, (Ed.), *Spoken and written language: Exploring orality and literacy* (pp.35-54). Norwood, NJ: Ablex.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Dasinger, L., & Toupin, C. (1994). The development of relative clause functions in narrative. In R. Berman & D. I. Slobin (Eds.), *Relating events in a narrative: a crosslinguistic developmental study* (pp. 452-514). Hillsdale, NJ: Lawrence Erlbaum.
- De Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19, 249-276.
- Dörnyei, Z. (2005). *The psychology of the language learner*. Mahwah NJ: Lawrence Erlbaum.

- Ehrman, M E., & Oxford, R. L. (1995). Cognition plus: Correlates of language learning success. *Modern Language Journal*, 79, 67–89.
- Ellis, R. (2005). *Planning and task-performance in a second language*. Amsterdam: John Benjamins.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299-323.
- Gilabert, R. (2005). *Task complexity and L2 narrative oral production*. Unpublished doctoral thesis. University of Barcelona.
- Grabe, W. (2009). *Reading in a second language. Moving from theory to practice*. Cambridge: Cambridge University Press.
- Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing*. London: Longman.
- Granfeldt, J. (2008) Speaking and writing in L2 French: Exploring effects on fluency, accuracy and complexity. In A. Housen, A., S. van Daele, and M. Pierrard (Eds.), *Proceedings of the conference on complexity, accuracy and fluency in second language use, learning and teaching*. Brussels March 29-30, 2007.
- Grigorenko, E L., & Sternberg, R. J., Ehrman, M. E. (2000). A theory based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *Modern Language Journal*, 84, 390–405
- Hild, G. (2007). A Magyar Egységes Nyelvérzékmérő Teszt (MENYÉT) vizsgálata hangos gondolkodtatásos eljárással [Examining the validity of HUNLAT through think-aloud procedures] *Iskolakultúra online*.
http://www.iskolakultura.hu/index.php?option=com_content&task=view&id=32&Itemid=1
- Iwashita, N., McNamara, T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information-processing approach to task design. *Language Learning*, 51, 401-436.
- Jarvis, S. (2002). Short texts, best-fitting curves, and new measures of lexical diversity. *Language Testing*, 19, 57-84.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 146-164.

- Kormos, J., & Sáfár, A. (2008). Phonological short term-memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, 11, 261-271.
- Kuiken, F., & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *International Review of Applied Linguistics in Language Teaching*, 45, 213-236.
- Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing*, 17, 48-60.
- Kuiken, F., & Vedder, I. (2009). The influence of task complexity on linguistic performance in L2 writing and speaking. Paper presented at the 3rd Biennial Task-based Language Learning Conference, Lancaster UK, 14th September 2009.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Malvern, D. D., & Richards, B. J. (1997). A new measure of lexical diversity. In A. Ryan and A. Wray (Eds.), *Evolving models of language*. Clevedon: Multilingual Matters.
- Nikolov, M., & Ottó, I. (2006). A nyelvi előkészítő tanfolyam: Az első tanév eredményei angol és német nyelvből [The intensive language preparatory year: Results of the first cohort of students in English and German]. *Iskolakultúra*, 5, 49-67.
- Niwa, Y. (2000). Reasoning demands of L2 tasks and L2 narrative production: Effects of individual differences in working memory, intelligence, and aptitude. Unpublished M.A. dissertation, Aoyama Gakuin University, Tokyo
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 555-578.
- Ottó, I. (2002). *Magyar Egységes Nyelvérzékmérő-teszt* [Hungarian Language Aptitude Test]. Kaposvár: Mottó-Logic Bt.
- Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30, 590-601.
- Pica, T. (1984). Methods of morpheme quantification: their effect on the interpretation of second language data. *Studies in Second Language Acquisition*, 6, 69-78.
- Pienemann, M. (1998). Developmental dynamics in L1 and L2 acquisition. Processability theory and generative entrenchment. *Bilingualism: Language and Cognition*, 1, 1-20.
- Pimsleur, P. (1966). *Pimsleur Language Aptitude Battery*. New York: Harcourt Brace Jovanovich.

- Révész, A. (2009) Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, 31, 437-470.
- Robinson, P. (1995) Task complexity and second language narrative discourse. *Language Learning*, 45, 99-140.
- Robinson, P. (2001a). Individual differences. cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research*, 17, 368–392.
- Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27-57.
- Robinson, P. (2003). The Cognition Hypothesis of adult, task-based language learning. *Second Language Studies*, 21, 45–107.
- Robinson, P. (2005a). Aptitude and second language acquisition. *Annual Review of Applied Linguistics*, 25, 45-73.
- Robinson, P. (2005b). 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. (2007a). Aptitudes, abilities, contexts and practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from cognitive psychology and applied linguistics* (pp. 256-286). Cambridge: Cambridge University Press.
- Robinson, P. (2007b). 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, 237-257.
- Sawyer, M., & Ranta, L. (2001). Aptitude and second language acquisition. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319-353). Cambridge: Cambridge University Press.
- Schiffrin, D. (1981). Tense variation in narrative. *Language*, 57, 45-62.
- Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 69-93). Cambridge: Cambridge University Press.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510-532.
- Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1, 185-211.

- Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). Cambridge: Cambridge University Press.
- Skehan, P., & Foster, P. (2008). Complexity, accuracy, fluency and lexis in task-based performance: a meta-analysis of the Ealing research. In S. Van Daele, A. Housen, F. Kuiken, M. Pierrard & I. Vedder (Eds.), *Complexity, accuracy, and fluency in second language use, learning, and teaching* (pp. 207-226). Brussels: University of Brussels Press.
- Wickens, C. D. (2007). Attention to the second language. *International Review of Applied Linguistics*, 45, 177-191.
- Yuan, F., & Ellis, Rod. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 24, 1-27.

Table 1 Descriptive statistics of the HUNLAT scores

| | | Mean | SD |
|--------|--------------------------|-------|------|
| HUNLAT | Phonological sensitivity | 12.95 | 3.40 |
| | Deductive ability | 15.54 | 2.19 |
| | Grammatical awareness | 11.00 | 3.43 |
| | Rote learning ability | 16.03 | 3.27 |
| | Total | 55.51 | 7.48 |

Table 2 Summary of task performance measures used in the study

| | FLUENCY | LEXICAL COMPLEXITY | SYNTACTIC COMPLEXITY | ACCURACY |
|-------------------|-------------|-----------------------|--|--|
| GENERAL | Speech rate | D-value | Clause length Ratio of subordinate clauses | Ratio of error-free clauses |
| TASK- SPECIFIC | ----- | ----- | Ratio of relative clauses | Ratio of error-free relative clauses Ratio of error-free verbs Ratio of error-free past-tense verbs |

Table 3 Descriptive statistics of the linguistic variables

| | Written | | Oral | |
|--------------------------------------|-------------------------------|------------------------------|---------------------------------|--------------------------------|
| | Mean | | Mean | |
| | (SD) | | (SD) | |
| | Cartoon description | Picture narration | Cartoon description | Picture narration |
| <i>Lexical variety</i> | | | | |
| D value | 62.34 ^c (17.02) | 58.5 ^d (16.56) | 43.02 ^{a c} (11.85) | 36.07 ^{a d} (9.41) |
| <i>Syntactic complexity</i> | | | | |
| Clause length | 5.76 ^b (.84) | 6.21 ^b (.91) | 6.46 (.98) | 6.61 (.83) |
| Ratio of subordinate clauses | .34 (.01) | .35 (.12) | .33 (.11) | .32 (.10) |
| Ratio of relative clauses | .09 ^b (.074) | .13 ^b (.086) | .10 (.08) | .11 (.07) |
| <i>Accuracy</i> | | | | |
| Ratio of error-free clauses | .81 ^c (.08) | .82 (.09) | .75 ^c (.11) | .76 (.12) |
| Ratio of error-free relative clauses | .79 (.27) | .67 (.32) | .67 (.33) | .69 (.40) |
| Ratio of error-free past-tense verbs | .94 (.16) | .96 (.18) | .81 (.29) | .88 (.28) |
| Ratio of error-free verbs | .84 ^c (.13) | .90 (.06) | .74 ^{a c} (.16) | .83 ^a (.14) |
| <i>Fluency</i> | | | | |
| Speech rate | ----- ----- | ----- ----- | 1.62 (.39) | 1.59 (.42) |

^a Denotes significant difference between oral cartoon description and picture narration

^b Denotes significant difference between written cartoon description and picture narration

^c Denotes significant difference between oral cartoon description and written cartoon description

^d Denotes significant difference between oral picture narration and written picture narration

Table 4a The correlation of linguistic variables with language aptitude scores in the oral cartoon description task

| | Complexity | | | Accuracy | | | Fluency | |
|--------------------------|-------------------|--------------------------|---------------------------|-----------------------------|---------------------------|--------------------------------------|-------------|-------------|
| | Lexical diversity | Ratio of sub.ord clauses | Ratio of relative clauses | Ratio of error free clauses | Ratio of error-free verbs | Ratio of error free past-tense verbs | Speech rate | Speech rate |
| Phonological sensitivity | -.10 | .38 | -.10 | -.04 | -.02 | .21 | .06 | -.01 |
| Deductive ability | -.17 | .26 | .01 | -.08 | -.07 | -.01 | .06 | -.15 |
| Grammatical sensitivity | -.32 | -.07 | -.06 | -.11 | .26 | .47** | .18 | .21 |
| Rote learning ability | .01 | .25 | .05 | .01 | -.02 | .27 | -.01 | .38 |
| HUNLAT total | -.22 | .32 | -.05 | -.08 | .07 | .40** | .12 | .26 |

** p<0.01

Table 4b The correlation of linguistic variables with language aptitude scores in the oral picture narration task

| | <i>Lexical diversity</i> | | <i>Syntactic complexity</i> | | | | <i>Accuracy</i> | | <i>Fluency</i> | |
|--------------------------|--------------------------|---------------|-----------------------------|---------------------------|-----------------------------|--------------------------------------|---------------------------|--------------------------------------|----------------|--|
| | D value | Clause length | Ratio of sub.ord clauses | Ratio of relative clauses | Ratio of error free clauses | Ratio of error free relative clauses | Ratio of error-free verbs | Ratio of error free past-tense verbs | Speech rate | |
| Phonological sensitivity | -.18 | .13 | .11 | -.05 | .18 | -.21 | .28 | .14 | .11 | |
| Deductive ability | -.40** | .01 | -.13 | -.14 | .27 | .06 | .01 | .10 | -.05 | |
| Grammatical sensitivity | .14 | -.04 | .05 | -.15 | .08 | -.40** | .10 | .01 | .27 | |
| Rote learning ability | -.06 | -.06 | .10 | .29 | .14 | .26 | .19 | .15 | .04 | |
| HUNLAT total | -.15 | .02 | .08 | -.01 | .26 | -.12 | .17 | .26 | .17 | |

** p<0.01

Table 5a. The correlation of linguistic variables with language aptitude scores in the written cartoon description task

| | Lexical diversity | <i>Syntactic complexity</i> | | | | <i>Accuracy</i> | | |
|--------------------------|-------------------|-----------------------------|--------------------------|---------------------------|-----------------------------|--------------------------------|--------------------------------------|---|
| | | Clause length | Ratio of sub.ord clauses | Ratio of relative clauses | Ratio of error free clauses | Ratio of free error-free verbs | Ratio of error-free past-tense verbs | Ratio of free error-free past-tense verbs |
| Phonological sensitivity | .08 | .15 | .04 | .13 | -.02 | .10 | -.03 | .28 |
| Deductive ability | -.13 | .15 | .26 | .17 | .21 | .17 | .29 | .04 |
| Grammatical sensitivity | .16 | .46** | -.22 | -.24 | .34 | .23 | .11 | .24 |
| Rote learning ability | -.05 | -.18 | .13 | .12 | -.17 | -.13 | -.14 | .14 |
| HUNLAT total | -.03 | .24 | .05 | .05 | .13 | .15 | .06 | .31 |

** p<0.01

Table 5b The correlation of linguistic variables with language aptitude scores in the written picture narration task

| | <i>Lexical diversity</i> | <i>Syntactic complexity</i> | | | | <i>Accuracy</i> | | | |
|--------------------------|--------------------------|-----------------------------|--------------------------|-----------------------------|--------------------------------------|---------------------------|---------------------------|--------------------------------------|--------------------------------------|
| | D value | Clause length | Ratio of sub.ord clauses | Ratio of error free clauses | Ratio of error free past-tense verbs | Ratio of relative clauses | Ratio of error-free verbs | Ratio of error free relative clauses | Ratio of error free past-tense verbs |
| Phonological sensitivity | .04 | -.30 | .05 | .03 | .21 | -.13 | -.09 | -.34 | .21 |
| Deductive ability | -.03 | .12 | .09 | -.13 | -.18 | -.06 | .19 | -.01 | -.18 |
| Grammatical sensitivity | .30 | -.10 | .15 | .07 | .20 | .05 | .10 | .13 | .20 |
| Rote learning ability | .13 | -.26 | .14 | .02 | .19 | -.01 | -.29 | -.06 | .19 |
| HUNLAT total | .20 | -.27 | .18 | -.01 | .03 | -.06 | -.18 | -.12 | .03 |

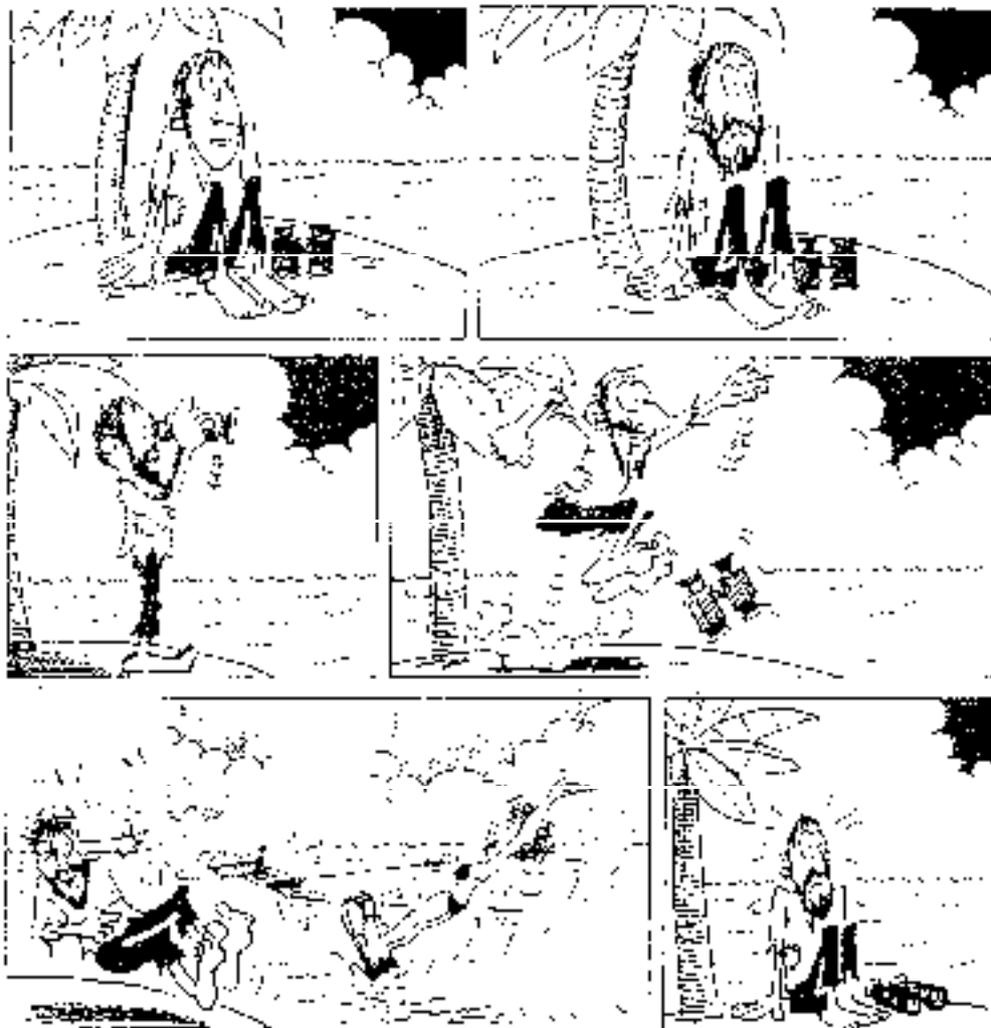
** p<0.01

Table 6. The overview of the psycholinguistic characteristics of the tasks with relation to the major findings of the study

| | Writing | | Speech | |
|--|------------------------|-----------------|----------------------|--|
| | Picture description | Story narration | Picture description | Story narration |
| Conceptualization load | Low | High | Low | High |
| Linguistic encoding demands | High | Adjustable | High | Adjustable |
| Need to share attention between lexical and syntactic encoding | Reduced | Reduced | High | High |
| Lexical variety | ++ | ++ | + | - |
| Syntactic complexity | - | + | + | + |
| Accuracy | + | + | - | + |
| Role of aptitude | Positive on complexity | Neutral | Positive on accuracy | Negative on accuracy and lexical diversity |

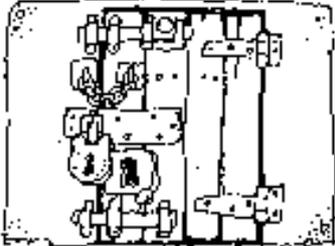
Appendix Tasks used in the study

Oral cartoon description task



Your task is to tell the story of this cartoon strip. You may also add stages not shown by the pictures. You have three minutes to think before you start.

Oral picture narration task



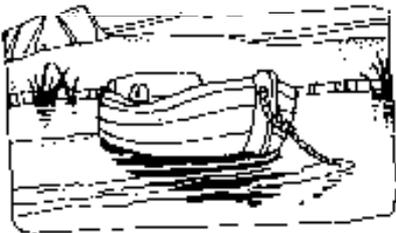
You will find six pictures in the envelope. Your task is to narrate a story which includes all the elements depicted by the pictures. You must use all the pictures from the envelope, but you may also add extra information if you wish. You have three minutes to think before you start.

Written cartoon description task



Please write a story of approximately 150 words which tells the reader what happened in this cartoon.

Written picture narration task



Please write a story of approximately 150 words, which includes all the elements depicted by the pictures below. You must use all the pictures, but you may also add extra information if you wish. Feel free to change the order of the pictures.