DIPLOMARBEIT

Titel der Diplomarbeit

„Complexity in second language classroom discourse"

verfasst von

Julia Stromer

angestrebter akademischer Grad

Magistra der Philosophie (Mag.phil.)

Wien, 2015

Acknowledgements

First, I want to thank my supervisor, Univ.-Prof. Dr. Christiane Dalton-Puffer, who took a genuine interest in this thesis and, besides her busy schedule, provided me with the best possible support and dedicated much of her precious time to guide me through this process. I am most thankful for her inspiring input, her invaluable suggestions and encouragement to change perspective, and her constructive and helpful feedback on draft versions of this thesis. I also want to thank her for bringing this topic to my attention and suggesting it for my final thesis, and additionally encouraging me to present it at a young researchers' conference about didactics of language. I feel truly grateful for her mentorship and kindness.

I also want to thank my family, my parents Monika and Josef, and my aunt Theresia, for their emotional and financial support throughout my studies and especially during the process of writing my thesis. Special thanks go to my sister and brother, Karin and Matthias, who provided me with a lot of support and who are always there for me if I need help. Stefan, my partner, who supported me enormously during this process of writing my thesis, deserves special thanks. Special thanks go also to my cousin, Marion, for proof-reading and providing feedback on parts of this thesis. I also want to thank all my dear friends who are always there for me.

I am also grateful to a number of people at the "Bundeshandelsakademie und Bundeshandelsschule Wien 11" for providing me with the possibility to conduct my research. I want to thank the headmaster of the school, Hofrat Mag. Peter Schlerith, for permitting the recording of one of the classes. Special thanks go to Mag. Daniela Weitensfelder, MA, who allowed me to observe and record her class and supported me during this project. I also want to thank the students who participated in my study for their enthusiastic participation.

Table of Contents

Abbreviations used

CAF	Complexity, Accuracy and Fluency
L1	First language
L2	Second language
WM	Working memory
PSTM	Phonological Short Term Memory
IL	Interlanguage
CIC	Classroom Interaction Competence
CEFR	Common European Frame of Reference
LTM	Long term memory
STM	Short term memory
IRF	Initiation-Response-Feedback
TBLT	Task-based language teaching
ANOVA	Analysis of Variance
sig.	Significance
S1, S2, S3	Student 1, Student 2, Student 3

Introduction

In the language classroom, the main objective is described as the acquisition of a second or foreign language. In order to learn and acquire a new language, it is important to guide students' attention towards the language and to provide enough means to learn the features of the target or second language (L2). Second language acquisition and an improvement of the interlanguage system are often assessed via specific measurements, the three performance domains complexity, accuracy and fluency. Hence, an improvement and a development of accuracy, complexity and fluency imply to some extent that the learner progresses and constantly learns new features of the target language. In order to develop the three performance domains, the student is frequently exposed to the target language, i.e. input, he/she should produce the language, i.e. output, while performing meaning-based activities which enhance output production and language acquisition, i.e. tasks and activities.

Regarding several language acquisition and learning theories, input as well as output production in the classroom plays an enormous role and, unfortunately, represents the main or only exposure to the target language for most of the students. In order to provide students with sufficient input as well as enough possibilities to produce the target language, tasks and activities where the learners are asked to perform in small groups or dyads are used and can help to enhance language development. This study is going to analyse how such tasks and activities can impact the performance domains and thus how language acquisition can be beneficially influenced in the second language classroom.

What is interesting while comparing this study to other recent research is that language production is not elicited in artificial and experimental environments, as in some previous studies (e.g. Yuan & Ellis 2003; Michel, Kuiken & Vedder 2007 among others) but during actual classroom discourse in a typical Austrian language classroom. Although some other studies (e.g. Skehan & Foster 1997) *did* analyse language performance during a classroom situation, this study focuses on an Austrian classroom in a high school context, contrary to most studies (e.g. Skehan & Foster 1997; Yuan & Ellis 2003, 2004; Norris & Ortega 2005 among others) which focused on college students in a university context.

Classroom discourse and the employment of tasks and activities in the language classroom form the basis for this thesis, showing that language performance can be influenced through various factors. Additionally, it is of utmost importance to focus on *real* classroom contexts; experimental studies might show more detailed results, however, mostly those results are distorted through the fact that the participants are eager to progress in the target language and their attention is directed towards the fact that they participate in a study regarding certain aspects of language production. During actual classroom discourse, however, such factors like motivation, willingness to improve and focused activities, may not be present to such an extent. Hence, analysing a real classroom situation might provide a more realistic picture of the language learning process in the typical language classroom.

The main objective of this thesis, thus, is to observe which tasks and activities are used in the Austrian language classroom in a high school context and to which extent such tasks and activities can influence a development of the three performance domains, complexity, accuracy and fluency. Chapter 1 presents the historical and theoretical background which paved the way for complexity, accuracy and fluency studies such as Krashen's (1984) Comprehensible Input Hypothesis, Swain's (1985) Comprehensible Output Hypothesis and several other concepts such as negotiation of meaning and noticing. I will further discuss the importance of competences in classroom discourse, especially in the Austrian classroom. In Chapter 2, I discuss several theories regarding cognition and second language acquisition and their importance for a development of the interlanguage system. After presenting a more general explanation of the analysis of language representation and the notion of explicit and implicit knowledge learning, some general positions of how explicit and implicit knowledge interact with each other are discussed. A detailed description of the role of attention and its importance for the student's language learning process and the working memory provides an insight to how attention and awareness can help the learner as well as the teacher to learn particular features of the second language and how noticing can enhance the language acquisition process. Chapter 3 will present a detailed description of the three performance domains, complexity, accuracy and fluency and their measurements. The following chapter focuses on task-based language teaching and classroom discourse. A thorough description of tasks and activities as well as the most common classification systems will be listed and analysed. The importance of task

complexity and task sequencing provides a well-founded basis for the empirical part of this study. In Chapter 5, several hypotheses about the interplay between complexity, accuracy and fluency are presented and the impact of planning time on task performance is discussed.

The empirical part of this thesis falls into four parts, (1) a description of the participants, the design and the method, (2) a presentation of the analysed tasks and activities, (3) an analysis of the data, including the dependent and independent variables as well as their measurements, and, finally, (4) a detailed discussion of the findings and results. In Chapter 8 I will discuss how the findings and results correlate with the previously formulated hypotheses and compare the results of this study to recent studies regarding the three performance domains. A particular emphasis will lie on the influence of some individual differences, the impact of planning time and the interdependent relationship of complexity, accuracy and fluency.

On the basis of the findings and results as well as my conclusions made during this research, I present my suggestions for further research and some implications for the second language classroom.

1 Conceptual background

Input is widely seen as one of the most important components of second language acquisition. Krashen (1984 [1981]) even claims that input is the only important feature of language learning, due to the assumption that the learner acquires language solely through attentive input processing. This input needs to be rich and meaningful, as well as slightly more complex and difficult than the learner's actual state in language learning, i.e. it must be linguistically challenging (cf. Swain 1985: 245; Krashen 1984 [1981]). When being exposed to this linguistically challenging input, the learner is more likely to feel the need to process the presented information and make it comprehensible. Hence, this type of input is called Comprehensible Input. However, in order to make input comprehensible, students have to understand the linguistically challenging input and process it which ideally results in a development of their interlanguage (cf. Krashen 1984 [1981]). Krashen (1985) later proposes two means of making input comprehensible, namely through the learner's use of situational context and through the teacher's simplification of provided oral input (cf. Ellis & Shintanti 2014: 177). Forms of simplified language can vary from foreigner talk to caretaker-talk to other forms of teacher's – sometimes student's – means of code modifications (cf. Gass & Mackey 2006; Ellis & Shintani 2014). However, these modifications are argued to be dynamic rather than stable, for example, language teachers tend to adapt their speech according to the learner's stage of development (Ellis & Shintani 2014: 7). In other words, at the beginning of language learning the language teacher is more likely to modify his/her language very frequently, whereas if the learner's language proficiency is high, there are fewer modifications found in the input. Regarding language material provided in the classroom, most handbooks for teachers focus on input and how to convey this particular input in the language classroom (cf. Ellis & Shintani 2014: 163). This might also be an argument for Krashen's (1984 [1981]) Comprehensible Input Hypothesis, regarding the importance of input in the second language learning, as well as the language teaching process. Another important aspect of Krashen's model concerns the importance of the affective filter during language learning and language acquisition. Factors for a *high affective filter*, i.e. less input is made comprehensible, or a *low affective filter*, i.e. more input is processed, are motivational aspects of language learning, the learner's self-esteem and the learner's anxiety (Ellis & Shintani 2014: 334; cf. Krashen 1984 [1981]). Such affective filters are not only influential in situational

contexts but affect the success of acquisition over long periods either in a positive or in a negative way (cf. Krashen 2009: 81). It seems to be clear that an important implication for the language classroom and the language teacher is the creation of a non-threatening environment. Thus, task design and sequencing play a significant role in the second language classroom in order to provide the students with a learner-friendly environment where they feel free to receive comprehensible input.

Following the Comprehensible Input Hypothesis, Swain (1985, 2005) proposes that the role of output, i.e. the student's language production in the classroom, plays as an important role. She further argues that solely representing the language does not provide sufficient possibilities of acquisition (Swain 2005). Hence, this hypothesis entails active student participation in classroom discourse. Ideally, the modification of the produced output if a communication breakdown happens, i.e. negotiation of meaning, should lead to an unravelling of misunderstandings. The notion of negotiation of meaning plays an important role in the Interaction Hypothesis which emphasises the importance of communication and interaction in the learning process (cf. Ellis & Shintani 2014: 339; Mackey 2012). Long (1996: 418) defines negotiation of meaning as the willingness to adjust the produced output due to the interlocutor's lack of comprehension, thus the learner provides means of modification until a smooth communication is assured. The modification ideally results in both student's production of comprehensible output (cf. Swain 1985) and a more complex language use which may cause a progress in the learner's interlanguage (cf. Ellis & Shitnani 2014: 9). R. Ellis (1994: 351f.) emphasises the argument that learners of a second language are actively trying out hypotheses about the target language which can either be confirmed, i.e. no corrective feedback follows the utterance, or disconfirmed, i.e. the uttered phrase is corrected or the student is asked to modify the previously produced output, thus to negotiate for meaning. Routines of negotiation of meaning typically involve clarification requests, asking for comprehension and confirmation checks (Mackey 2012: 6).

The process of noticing represents another positive side-effect of comprehensible output production. Noticing is described as the aware and conscious detection of an unknown feature, its processing by the learner and "subsequent storage in long term memory" (N. Ellis 2005: 317). Awareness and noticing can be seen slightly different, as awareness presents solely the process of recognising without further

processing. By means of the student's realisation of and reflection on their own language proficiency and use of already acquired knowledge of the L2 to analyse the represented language (cf. Swain 1985; Bialystok 1984), the learner's interlanguage is more likely to progress thanks to noticing than without it. Similarly, Schmidt (2001) argues for a beneficial effect of noticing on language learning by saying that "the more learners notice, the more they learn" (Ellis & Shintani 2014: 179). It is also pointed out that awareness is not as important as the act of noticing, but that awareness can facilitate the learning process (Robinson 2003: 637). In other words, even if the learner does not completely process a certain language feature, i.e. completely understand its functions, it is still more beneficial for the language development that the student is aware of this particularity. Hence, the teacher's guidance of the student's awareness towards certain language features might prepare the way to noticing but cannot assure that the noticing process takes place. Robinson (2003: 638) further supports this argument by stating that the process of noticing may prepare the way for language intake but a conclusive co-existence seems improbable. In other words, conscious reconstruction or analysis of represented language is not crucial for language acquisition, but noticing of an unknown structure is enough to stimulate the acquisition process. Thus, conscious learning and focus on form might provide even more support for the learning process (Ellis & Shintani 2014: 179). In recent research, it has been stated that noticing-the-gap, i.e. student's attention on the difference between own produced output and target language, represents a stimulation to progress interlanguage development (e.g. Swain 1985; Mackey 2012; Pica 2005). Negotiation of meaning can be used as a means to direct the student's attention to the incorrectly or inappropriately used linguistic features (Ellis & Shintani 2014: 11). One criticism, however, which needs to be voiced, is that one can never ascertain the beneficial influences of the noticing process, as it is impossible to verify whether or not the student is consciously or unconsciously noticing particular features of the L2. Nevertheless, means to direct the learner's attention to certain aspects of the language or the learning process are likely to raise his/her awareness of the deliberately emphasised properties of the target language. Schmidt (2001: 7) further argues that learners of a second or foreign language tend to be overwhelmed by the amount of new information. Therefore, it seems to be obvious that basic knowledge needs to be automatised first before new linguistic features can be incorporated into the learner's knowledge of the L2.

1.1 Competences in the second language classroom

Competences and their descriptors seem to have reached every aspect of the Austrian language classroom. Regarding the curriculum for English as a second/foreign language, strikingly, almost every part is in connection with language and communicative competences. Hence, it is not surprising that the use of interaction does not only include communicative competences *per se,* but also classroom interactional competences (CIC) in particular. It can be argued that such competences include the Interaction Hypothesis even more into the classroom by emphasising the importance of interactional competences. Walsh (2012) accentuates the notion that not only the target language is in the centre of attention but he further defines CIC as "teachers' and learners' ability to use interaction as a tool for mediating and assisting learning" (Walsh 2006a: 130).

Interactional competence describes a specific language skill, the use of language to negotiate meaning by the incorporation of several language features, i.e. codes, forms and functions, semantics and semiotics, and paralinguistic features (Seedhouse & Walsh 2010: 140). Walsh (2012: 2) argues that it is interactional competence that might receive the least attention in typical language classrooms but what is most needed in real life situations. Regarding research what seems to be of utmost importance of today's classroom concerns the production of accurate, fluent and appropriate language, but if the interaction goes smoothly or if codes of communication are interpreted correctly only plays a minor role (cf. Walsh 2012). It is further stated that for example, grammatical and syntactical features of the L2 might represent items for assessment or practice, but interactional competence does not (Walsh 2012). Contrasting this perspective might be that evidence for interactional competence *can* be found in the communicative language classroom, as negotiation of meaning provides opportunities to involve such strategies.

Classroom interactional competence involves all of the features found in interactional competence, including an additional aspect: the situational context of the classroom interaction (Seedhouse & Walsh 2010). Taking into consideration the competence descriptors of the Austrian English language classroom (cf. Lehrplan Englisch), one can observe how classroom interaction and communicative goals orient themselves towards competence achievement. Seedhouse and Walsh (2010: 140) further

argue that CIC might become an additional skill which will constitute another competence of the communicative language classroom. Another beneficial aspect of CIC is the awareness of both pedagogical goals and interactional strategies while using semantically and grammatically appropriate language for communicative purposes (Seedhouse & Walsh 2010: 141). If CIC is present in the language classroom and teachers' as well as students' awareness of it is raised, students are provided with more opportunities for output production, productive feedback, support and more time (ibid.). All of those aspects appear to have beneficial effects on the student's performance and language development. Some examples of CIC include techniques like scaffolding, modification of teacher initiation or paraphrasing which might replace frequently used and fast techniques like repair and recast or teacher-echo (Seedhouse & Walsh 2010: 141). Moreover, typical error correction becomes less important, whereas negotiation of meaning, output modification and student's self-correction is brought to the fore (cf. Seedhouse & Walsh 2010: 143). Referring back to Swain's (1985) Output Hypothesis and Long's (1996) understanding of negotiation of meaning, it seems likely that CIC has already entered the communicative language classroom and exerts its influence on classroom and task design.

2   Cognition and second language acquisition

The notions of complexity, accuracy and fluency, have a rather complex structure. An understanding of the relationship between cognition and second language acquisition is a helpful tool for a detailed description of the complexity of learner talk. The cognitive process of language learning has multiple factors that influence output production and its complexity. Some of those factors are the cognitive analysis of language, the influence of working memory, the relationship of cognition and task, the importance of attention and consciousness and the difference between explicit and implicit knowledge of the L2.

Schmidt (2001: 16) defines language analysis as the constructive process of internal and formal language representations. There is a transition between unanalysed and analysed language, given that unanalysed language becomes more and more analysed, therefore, providing the learner with more control of certain concepts and contributing positively to the acquisition of fluency (Schmidt 2001: 16). According to Bialystok (1984: 45) the analysis of knowledge is "the extent to which the learner is able to represent the structure of knowledge along with its content". In other words, if the learner is familiar with the coding of the presented language this means that it is analysed and he/she does not only pay attention to the meaning but also to its form. If language representation is unanalysed, the learner might only focus on its meaning, ignoring the form completely (cf. Bialstok 1984: 45). A manipulation of the previously analysed form-meaning relationships provides the learner with more opportunities for language use (Bialstok 1984: 45). Therefore, the learner gains more and more structural forms by analysing the given input of how to produce more complex output. Bialstok (1984: 46) points out that the students need to reach a certain level of proficiency, equal to mastery, which is called "cognitive control" in order to use certain structures or forms appropriately, including a balanced form-meaning relationship. However, it might be possible that students achieve a certain level of complexity even though mastery of language or of a particular rule is not yet achieved. Individual factors, like observational and imitational skills might play an equally important role for complex output production, as well as the ability of cognitive control. Another interesting point is that cognitively controlled ways of communication might hinder the production of complex language. Some examples for control-based strategies present evidence for this

assumption such as the use of context for filling a lexical gap, gestures and mimic, the use of the L1 etc. (Bialstok 1984: 46). Such strategies might assure the conveying of meaning; however, they do not push the learner to modify his/her output by producing more complex language.

Bialstok's (1984) definition of analysed and unanalysed language correlates astonishingly well with the rather popular notion of explicit and implicit knowledge. Both of those two concepts would be part of the analysed representation of language, representing different stages of the analysis. Explicit knowledge describes the beginning of the analysis, as the student learns a new form or function of a language, but cannot apply this form or function fluently. It can be seen as knowledge which the learner is aware of and which helps to monitor output production (cf. R. Ellis 1994: 355). In the following stage, the form or function is acquired and automatised (cf. N. Ellis 1994) – which means that it forms a part of the learner's implicit knowledge of the L2. Some researchers (e.g. Schmidt, N. Ellis) described implicit knowledge as "true competence in an L2" (Schmidt 2001: 16), or a key factor for fluency. Implicit knowledge can be divided into two main parts, formulaic knowledge, i.e. automatised language chunks, and rule-based knowledge, i.e. internalisation and automaticity of rules and structures (R. Ellis 1994: 355). The main difference of implicit and explicit knowledge is, therefore, the stage of the learner's automatisation of specific features and functions of the target language. Explicit knowledge functions rather as a language monitor, basically used for correction and editing, whereas implicit knowledge provides the learner with forms and utterances, so-called chunks (N. Ellis 2001), which can be employed automatically in fluent speech. A similar categorisation of knowledge is presented by the distinction of declarative and procedural knowledge, where declarative knowledge acts like explicit knowledge, i.e. the knowledge of the facts, and procedural knowledge depicts the ability of language performance in certain situations (cf. DeJong 2009: 97), which correlates with the notion of implicit knowledge.

Regarding the relationship of explicit and implicit knowledge, the literature provides a considerable variety of different points of view. Three of the most common but very oppositional positions will be explained briefly (cf. Schmidt 2001).

Krashen (1984 [1981]) supports a *non-interface position*, i.e. a clear cut distinction of acquisition and learning with no transfer from explicit to implicit knowledge. According to this hypothesis the knowledge of rules and grammar of an L2

has solely a monitoring function and can be used for editing purposes and error correction (Krashen 1984[1981]: 1f.). What is of importance for the learner's understanding of the language is specifically modified input (done by teachers, native speakers etc.) which helps the learner to acquire language solely based on meaningful or comprehensive input presentation (ibid.). An additional argument is that there is some sort of naturally stable order of acquisition where learning of explicit knowledge has no influence on language acquisition. Krashen (1984[1981]: 5) claims that conscious language learning does not result in communicative competences but only serves as a language monitor. In this particular case, time constitutes a crucial aspect, meaning time pressure might have negative effects, for being accurate the learner needs sufficient time for rule application and/or self-correction, if necessary (cf. Krashen 2009: 84). According to this hypothesis, a distinction between the acquisition and the learning process where no transfer of explicit and implicit knowledge exists seems to be inevitable. In other words, learning a second language represents the notion of explicit knowledge, whereas acquiring it solely depends on implicit knowledge.

A rather different approach is proposed by DeKeyser (1998: 61f.) who argues that the learner needs to master a rule in order to acquire it and use it fluently. Therefore, explicit knowledge is the basis for implicit knowledge as automatisation can only be achieved through practice *and* knowledge of the rule (DeKeyser 1998: 62). This point of view has been called the *strong interface position* (cf. Schmidt 2001: 12). It is further supported, as Schmidt (2001) also points out that even though acquisition seems to happen unconsciously, consciousness might be of importance in order to process implicit knowledge (cf. Ellis & Shintani 2014: 178).

A more open position, the *weak interface position*, is presented by R. Ellis (1994) and N. Ellis (2005). N. Ellis (2005: 306) suggests that most learning happens implicitly through both the production and reception of language. He further argues that language acquisition which happens mostly unconsciously, i.e. implicitly, represents the basis for fluency and is predominantly acquired through practice (N. Ellis 2005: 307). Even though evidence supports the statement that explicit and implicit knowledge are stored in different parts of the brain (Paradis 1994), it can be argued that there exist some sort of interaction or transfer between explicit and implicit knowledge (N. Ellis 2005: 307). R. Ellis (1994) supports this argument by quoting Gass's (1988) hypothesis that explicit knowledge may be a representation of the L2 which the learner is not yet

able to process properly. Hence it is stored in "the form of some kind of explicit representation of L2 items and Rules" (R. Ellis 1994: 349). In other words, fluent communication with a strong meaning focus is more likely to draw from implicit knowledge. However, if a misunderstanding occurs which forces the learner to reflect on, for example, accuracy or complexity, explicit knowledge can act like a sort of monitor and contribute positively to language acquisition, therefore, to the acquisition of implicit knowledge (cf. N. Ellis 2005: 308). This notion of interface is supported by a strong emphasis on the dynamics of the learning process (N. Ellis 2005). The dynamic of the process shows that the interlanguage system changes constantly over time and that new language items or features can have an effect on the entire representation of the knowledge of the second language (R. Ellis 1994: 350). Regarding classroom discourse and the range of activity types which are common in most of today's classrooms, it seems to be more likely that the weak interface position is more present in the minds of most language teachers of the communicative classroom than the non-interface position. In a second or foreign language classroom the learner is provided with opportunities to practise and automatise language features, implying that repetition and practice can help to acquire the target language. When learning a new item, the learner is exposed to examples of language features, either derives a rule or is presented with explicit knowledge about those features, practises the items and, ideally, stores them in the long term memory and includes the new features into the interlanguage system. Comparing this procedure to the weak interface position, it can be observed that the learner goes through several stages, ranging from learning and understanding explicit knowledge to using the language feature automatically, i.e. implicitly. Hence, it may be argued that a transfer between explicit and implicit knowledge is taking place when learning a new language.

Concluding, explicit as well as implicit knowledge reinforce a change in the interlanguage system and provide the learner with the means to produce complex language (Skehan 2009b: 93). The use of previously acquired and automatised chunks and a simultaneously available rule-based system result in a beneficial fluency in learner talk (Skehan 2009b: 94). In other words, on the one hand, the learner can employ well-known structures without taking any risks of incorrectness, on the other hand, new hypotheses can be tested and more complex language can be produced. This sort of beneficial fluency seems to depict another stage in the interlanguage

development process, as "previous restructuring becomes automatized" (Skehan 2009b: 94). In order to be capable of such a cognitive process, the student's working memory plays an important role. Task designers need to take these notions into consideration when creating activities for the language classroom. Sufficient opportunities which provide the learner with enough time to restructure and, further, automatise those language chunks should occur frequently and recurrently. Supporting the student to use a wider repertoire and providing sufficient opportunities to actively reuse previously uttered complex chunks will have beneficial effects on a balanced accuracy, fluency and complexity development (Skehan 2009b: 98). Skehan (2009b: 95) points out that communicative problems at "the right level of processing difficulty" constitute an example of such tasks. In order to prevent the learner from being either too focused on form or too focused on meaning, a wide range of different tasks assure a balanced interlanguage development (Skehan 1992). It is further argued that complex language needs a certain amount of attentional capacity in order to be integrated into automatised structures, i.e. to become part of fluency (Skehan 2009b: 98). To avoid the learner to digress from focus on form to focus on meaning, attention needs to be guided towards complex language acquisition and a conscious intake of complex structures.

2.1   The role of attention

A crucial factor in second language learning seems to be attention and conscious learning or focusing. Schmidt (2001: 11) states that almost all factors which influence language learning are united when it comes to attention and being focused on particular features of the language. Hence, the allocation of attention can have beneficial effects on both, form and meaning, as accuracy, fluency and complexity can become the focus of the learner's attention (cf. Schmidt 2001). One important aspect of attention however is that just globally paying attention does not benefit the learner, but that it has to be focused on particular features in order to raise the student's awareness and to benefit the student's interlanguage (IL) (Schmidt 2001: 30). Foster and Skehan (2009: 276) argue that it seems rather unlikely that learners are paying attention to both, meaning and form, but that attentional space is limited and students might prioritise meaning over form. Thus, it seems to be obvious to draw the learner's attention to focus on form and structure, in order to complexify and advance the IL system. Regarding teaching

practice this may result in controlled activities to direct the learner's attention to particular domains that are learned and/or need to be practised. Another way of directing the learner's attention is the use of negotiation of meaning, as the student is more likely to produce a more complex and correct structure when restructuring and modification occurs. If this is done in an accuracy-focused way, the student's attention is not only directed towards meaning modification but also towards a more accurate use of certain forms. N. Ellis (2005: 312) further argues that consciousness can be seen as a "gateway", which means that if the student is consciously learning, numerous unconscious links are connected to this new information. "Consciousness is the interface" (N. Ellis 2005: 312) and helps the learner to "[access, disseminates] and [exchange] information and [to exercise] global control" (ibid.). Students thus need to be attentive to decide on whether or not the learnt information is important (Robinson 2003: 635). In other words, without being aware of what is learnt, it seems to be unlikely that the learner acquires the new information as effectively as by consciously paying attention. Moreover, consciousness is in some way seen as the link between implicit and explicit knowledge. Assuming that there is a transfer between implicit and explicit knowledge acquisition, noticing and attention have an enormous influence on how language is acquired, both either implicitly, i.e. passively through stimuli, the so-called process of priming and tallying, or explicitly, i.e. actively through focused attention on certain L2 features, also called noticing (N. Ellis 2005: 311).

Robinson (2003: 652) draws a connection between attention and Swain's (1985) output hypothesis by stating that learners need to put an "effort" into both language production and noticing, thus being conscious and attentive to input as well as output. By further developing this hypothesis, one might argue that the learner, in order to produce more complex language, needs attentive and conscious use of the L2, hence showing willingness to avoid the usage of automatised chunks and to take risks to produce more complex language, by a possible employment of explicit knowledge as a monitor. This process in combination with the willingness to complexify or modify the output seems to be likely to promote a development in the learner's interlanguage system.

Additionally to attention, it is mentioned that planning, as a form of pre-task attention paying to a particular form, may lead to improvements in accuracy, complexity, as well as fluency (Schmidt 2001: 13f.). This perspective correlates

positively with the noticing hypothesis and might be an indication that more complex output is produced when students have enough time to prepare and edit their ideas and arguments. According to Ortega (2005: 77) there was more evidence of fluency and complexity during task performance among adult learners who were provided with planning time. It is further argued that planning time as a form of pre-task activity gives the learners the possibility to activate prior knowledge, which allows the learner to direct more attention towards the actual task (cf. Mackey 2012: 67f.; cf. Robinson 2003). Although, planning time might entail positive aspects for complex language production; it may hinder students development of language automaticity and limits preparation for real-life situations which require spontaneous and immediate reaction implying a rather limited time for planning (Mackey 2012: 68). Robinson (2001: 31) states that if more attention is directed towards language and code to fulfil a (more complex) task, the learner can acquire language features progressively which are needed for language use outside the school. Complex tasks might act as a starting point to learn and test complex language which, successively, can be automatised during practice.

R. Ellis (2009: 474) differentiates between two kinds of planning, pre-task planning and within-task planning. Pre-task planning is further divided into rehearsal, i.e. repetition of a task, and strategic planning, i.e. preparation for actual task performance (R. Ellis 2009: 474, 476). On-line planning is defined as the speaker's attention towards articulation and structure during speech production and modification (Yuan & Ellis 2003: 6). Within-task or on-line planning can be either pressured by providing only limited time for task performance, or unpressured, implying that students do not have any time limits (ibid.). Learners provided with enough time to prepare, i.e. plan for task performance showed an improvement in both fluency and complexity during actual task performance, whereas the beneficial effects on accuracy are not that clear (Ellis & Yuan 2004: 60). Yuan and Ellis (2003) further stated that not only pre-task but also on-line planning has a positive influence on language complexity.

One main characteristic of attention and its influence on language acquisition is that there seems to be an interdependent relationship between form and meaning. It does not suffice if the learner's attention is solely directed to the form of a structure without presenting a meaningful context (cf. Schmidt 2001: 30f.). For example, vocabulary can only be acquired if its form and its meaning are present and the learner can pay attention to both (ibid.). However, this might not be true for familiar and well-known forms and

functions, as Schmidt (2001: 31) points out that only "comprehending what is *new* requires conscious processing" (own emphasis added). Referring back to the noticing hypothesis, noticing would be found at the beginning stage of the acquisition of a new feature of the target language (Schmidt 2001: 31), whereas automatised, i.e. implicit use can be seen as the stage of mastery. Nevertheless, there is some criticism towards this hypothesis, as one can never be certain if and how the learner is processing features to which attention was directed and if and how the planning time was used for attentive preparation (cf. Schmidt 2001: 18). Attention, however, has not only a beneficial effect on the language learning process *per se*, but likewise represents an important part of some working memory models (Mackey 2012: 100).

2.2   Working memory and language acquisition

Working memory (WM) is commonly described as "the ability to temporarily store *and* process information in order to carry out a cognitive task" (Mackey 2012: 93). It is the system that enables the learner to perform complex tasks without constantly digressing during task performance (Baddeley 2010: 136). It is used for both processes, noticing and encoding of unknown language features as well as drawing on explicit knowledge during the modification process (N. Ellis 2005: 317). Although there are several different models of how working memory actually operates, there is some general agreement on the following points: working memory has an influence on higher-order cognition, attentional resources are distributed and monitored by an executive attention, individual differences influence working memory and, finally, it affects second language acquisition (Mackey 2012: 94). Working memory can be seen as the home of language development, application and modification and it helps to redirect the focus if attentional digression occurs (N. Ellis 2005: 337).

Baddeley (1974; 2010; 2003) divides the working memory into three parts which are influential for language acquisition: the phonological loop, the visuo-spatial sketchpad, and the central executive (cf. N. Ellis 2001: 33f.; Baddeley 2003: 190f.). The phonological loop operates by processing auditory input through a phonological analysis and storing the information into the phonological short-term storage (Baddeley

2003: 193). The stored information is further used for both rehearsal of the information or spoken output production (ibid.). The process of rehearsal encodes the new information and transfers it into the long-term memory (LTM) (Robinson 2003: 654). In other words, the rehearsal process enables the learner to become aware and notice particular features of the L2 which helps him/her to store this information into the LTM, hence, to develop his/her IL system. Additionally, the information is not only stored in the phonological short-term store, but after rehearsal and/or spoken output production, it becomes continually part of the long-term memory and the learner's representation of the L2 (cf. Baddeley 2003: 196).

The visuo-spatial sketch pad, on the other side, is referring to the mental processing of imagery and symbols of a language (cf. Baddeley 1992). It is assumed that attentional control is the responsibility of the central executive, and that these three "slave systems" (Baddeley 1992) are interdependently correlated. A fourth component which is in close relationship with the central executive but more likely to be concerned with storing information is the episodic buffer (Baddeley 2003: 203). Additionally, Baddeley (1992) points out that the working memory is not only responsible for the process of long-term memory development but seems to act as interface between attention and perception in the language acquisition process. Working memory capacities are thus an influential aspect for the noticing process in the second language classroom.

There are two main but very opposing models which attempt to explain the operational processes of working memory, the "less is more" model, supported by Newport (1990) and the "more is more" model (Ludden & Gupta 2000). The "less is more" model supports the argument that smaller working memory capacities may analyse less input in return in more detail (Mackey 2012: 94). On the contrary, Mackey (2012: 94) explains that greater WM capacity may facilitate L2 learning, supporting the "more is more" model. The concept of verbal working memory describes the assumption that higher working memory capacity equals higher scores on perceptive skills, i.e. reading and listening (Mackey 2012: 97). Similarly it is argued (Mackey 2012: 97f.) that working memory capacity and an expertise in communication strategies may be related to each other. However, one needs to take into consideration that contextual and situational influences may "[mediate] the relationship between WM capacity and L2 learning outcomes" (Mackey 2012: 98). The Phonological Short Term

Memory Model (PSTM), however, considers the phonological loop as starting point and main focus (Mackey 2012: 99). The PSTM can be seen as the ability to reproduce and/or remember language due to phonological features and is measured through testing reproduction or repetition of fictional or non-words (ibid.). Both concepts, verbal working memory and phonological short term memory, appear to have influential effects on (second) language acquisition. Concurrently, external factors, for example instruction giving, classroom atmosphere et cetera, and internal factors, such as level of proficiency, communicative pressure et cetera, can function as means of mediation during the learning process (Mackey 2012: 99).

Robinson et al. (2014) hypothesised that recast might have greater influence on the interlanguage system of learners with higher working memory capacities than those with lower WM capacities (cf. Mackey 2012: 101). It seems to be more likely that students with higher WM capacities tend to modify or edit their output more frequently. Mackey (2012: 103) explains this hypothesis by arguing that the process of modification engenders an enormous cognitive load. In order to modify the output (1) the student's attention needs to be directed towards form (away from meaning), (2) the preceding utterance needs to be compared to corrective feedback as well as to explicit and implicit knowledge stored in the long term memory, implying that the preceding utterance as well as ruled-based knowledge or ready-made chunks are remembered, (3) identification of the incorrect utterance or feature concludes in (4) modification of previously produced output (cf. Mackey 2012: 103). Assuming that the learner works through all of these steps almost simultaneously and under a great pressure of time, modification might be more successful if working memory capacities, i.e. in this case storage and processing of the knowledge of the L2 are at a higher state (ibid.). Lack of communication or misunderstandings, therefore, could be described as a failure of successful implicit knowledge application and the learner is obliged to focus on form rather than on meaning in order to modify the utterance, i.e. to draw on stored explicit knowledge (cf. N. Ellis 2005: 316). Such situations, further, function as opportunities for negotiation of meaning to occur and to provide the student with time to restructure what has been said. Concluding, the learner needs to draw on working memory and prior knowledge to solve the communication problem and it seems to be obvious that higher WM capacities can be beneficial in this case. Mackey (2012: 106) summarizes that higher working memory may facilitate the noticing process, enhance output

modification and promotes beneficial effects of task-based interaction. Even though working memory appears to have a great influence on language acquisition, it cannot be influenced through the teacher, the student and/or classroom discourse and due to limitations of this study cannot be included.

Given that working memory capacities are extremely dependent on the individual learner, it seems to be obvious that other individual factors or differences exert a huge influence on whether language acquisition and/or language production progress successfully (cf. Robinson 2003: 662; Larsen-Freeman 2009). Individual factors include introversion vs. extroversion, emotional stability, experience or openness to experience, i.e. willingness to risk taking, and, finally, intelligence as well as emotional intelligence (cf. Dewaele 2009: 626-630). According to Robinson et al. (2014: 258f.) aptitude can influence the learner's language learning ability, for example, focus on form depends on the learner's grammatical sensitivity and noticing; allocution of attention, therefore, is influenced by a learner's aptitude. Additionally, grammatical sensitivity plays an important role for explicit and implicit learning of the L2. It has also been argued that the individual differences might impact the learning of explicit knowledge more significantly than implicitly learned language (Robinson 2003: 662f.). However, Robinson (2003: 664) points out that there is yet too little evidence and research done in this field to provide a fully developed theory about the influences of individual differences on implicit learning. In addition to those cognitive individual factors, affective variables, for example motivation, anxiety or confidence (Robinson 2001: 32) can have an additional impact on task performance. Larsen-Freeman (2009; 2006) emphasises that studies on second language acquisition might give an insight into how an interlanguage system develops, however, it is important to keep in mind that language learning is strongly learner dependent. Therefore, one might argue that general statements or hypotheses are likely to be challenged by learner-dependent individual factors when researching the three domains of performance, i.e. complexity, accuracy and fluency. However, in order to provide a complete picture of influences on interlanguage development, individual factors need to be included into a thorough analysis. Due to the limited scope of this study, only some aspects of individual difference can be taken into account, such as willingness of risk taking and individual choices of speaking styles.

Interestingly, not only language aptitude seems to be important for successful language learning, but it might be hypothesised that a positive attitude towards the target language has influential effects on interlanguage development as well. In other words, in order to assure complex language production, it appears to be inevitable to create positive attitudes towards the language, including the culture that is learned. Assuming that the learner has reached a certain stage of complexity, Skehan (2009b: 93) argues that the student will not only articulate his/her ideas more precisely, but also those ideas will reach another level of complexity. In contrast, individual differences and preferences might result in situations where learners do not wish to enhance their proficiency level and show no willingness to restructure or modify produced output. If then output is produced at all, it seems to be unlikely that negotiation of meaning will take place (cf. Skehan 2009b: 93). The implementation of useless and poor instruction and/or input may result in both, unwillingness to learn the language and fear of performance, which brings language acquisition to a standstill (cf. Skehan 2009b: 93). However, this willingness to modify output and to complexify language production can have enormous beneficial effects on the development of the IL system (cf. Skehan & Foster 2001: 190). Task design, content information, as well as context and general classroom atmosphere can counter those negative attitudes and evoke the learner's interest not only for the language but also for positive interlanguage change. An interesting choice of content, a non-threatening classroom atmosphere and appropriately challenging tasks thus represent the starting point for successful language acquisition.

In connection with classroom discourse and task performance, a reduction of the cognitive load of output modification or task performance might support learners with lower WM capacities and can be accomplished through guiding the learner towards a noticing-the-gap process or through providing more possibilities to reuse pre-structured utterances. R. Ellis (2009: 494) provides some support for this argument by pointing out that fluency benefits more from planning time during simpler tasks than more complex ones. Thus, planning time can be incorporated to support students with lower WM capacities as well as to improve fluency of output production (R. Ellis 2009: 494). Planning time is assumed to free the learner's working memory capacities and can have an effect on various trade-offs between accuracy, fluency and complexity (R. Ellis 2009: 502). However, it has been argued that planning time can have different effects on CAF among more proficient learners than among beginners, implying that more

proficient learners' fluency benefits more from strategic planning rather than their accuracy and complexity (cf. R. Ellis 2009: 503; Bygate 2009: 493).

3 Accuracy, Fluency and Complexity

Widdowson (1989) initially started with the opposition of analysability and accessibility to provide a measurement for language performance which correlates to some extent with the notion of explicit and implicit knowledge and Bialstok's (1984) concept of analysing represented language (cf. Skehan & Foster 1997). Analysability, similar to explicit knowledge refers more to the form and structure of the used language, whereas accessibility refers to meaningful as well as automatised language use (cf. Peterwagner 2005: 13f.). In other words, there is a difference between focus on form, i.e. accuracy and focus on meaning connected to fluency. Regarding the literature, one can observe the development that the initial distinction of accuracy and fluency developed more and more towards a triadic distinction of the performance domains, adding the component of complexity (Housen & Kuiken 2009: 1). Those three notions, namely accuracy, fluency and complexity, are used to represent a means of measurement for both performance as well as the learner's L2 proficiency level (ibid.).

3.1 Accuracy

As mentioned above, the initial focus of research as well as language teaching lay in accuracy and fluency activities, the latter promoting spontaneous usage of the L2, whereas during accuracy activities the emphasis was on accurate and grammatically correct production of particular linguistic structures (Housen & Kuiken 2009:1). It can be argued that this distinction represents the general assumption that either meaning or form can be in the focus of attention. However, through the addition of another component, complexity, a more precise image of language acquisition and language performance can be presented. Housen and Kuiken (2009: 3) argue that accuracy as well as complexity represent the learner's linguistic knowledge of the L2, and are therefore both related to form. Accuracy, in general, is described as the correctness of language performance, i.e. the absence of errors and mistakes (cf. Housen & Kuiken 2009: 3; Foster & Skehan 2009: 279). The only aspect which might offer some room for discussion is the question of which standard or norm to use for assessing language production, i.e. Standard English which in most language classrooms is either the British or the American Standard. The foundation for accuracy is seen as explicit

knowledge or *"learned linguistic knowledge"* (Towell 2007: 260) of the second language, such as grammar rules, syntactic and pragmatic structure, et cetera. However, even though a student might speak very accurately, the language used might be relatively simple. In other words, the learner may either employ a rather conservative approach to language use (Foster & Skehan 2009: 279) or shows more willingness to take risks thus producing more complex output. For instance, simply structured and short answers are more likely to score high on accuracy measures. However, a longer and more complex output production might help the student to accomplish a development in his/her interlanguage system. Hence, accuracy can only function as a measurement for error-free language production but does not provide information about the complexity of the analysed utterances.

3.2 Fluency

The notion of fluency is more difficult to define. Fluency can be divided into different sub-categories which are reflected through their means of measurements: (1) speed fluency, measured by rate and density of delivery, (2) breakdown fluency, measured through pauses in speech and (3) repair fluency, measured through false starts and repetitions (Housen & Kuiken 2009: 4). Foster and Skehan (2009: 280) emphasise another important aspect of fluency and fluency activities, the meaning component and the correlation to real-life situations. In other words, fluency enables the learner to perform more native like language by the use of idioms or prior automatised language chunks. Schmidt (2001: 9) explains that chunking can be seen as a result of attention and working memory capacities. Briefly, a chunk is a language unit which the learner has already acquired and which is stored in the LTM to be used automatically during language performance (cf. N. Ellis 2001; Newell 1990). In other words, the language learned is mentally represented and ready for automatic use during spontaneous language production (cf. Towell 2007: 260). Given that this process enables the learner to express him/herself more fluently by the use of previously automatised chunks (cf. N. Ellis 2001: 38), activities to internalize those chunks provide the learner with an opportunity to produce more native-like and maybe even more complex language. In order to produce more complex language, however, the student needs to utter different language structures and test new hypotheses instead of constantly relying on already

automatised chunks. Chunking can only contribute to complex language production if the student is willing to use more new language structures after having automatised already learned language chunks.

Besides the above mentioned definitions, fluent speech might be described as the skill to produce speech at a normal rate (cf. Pallotti 2009: 591). Strikingly, a definition of normality or normal speech rate is not further described. However, the measurement of speed fluency is frequently used, without detailed definition if fast or slow speech might be taken as a reference point. The rate of speed can differ significantly in different contexts, for example, a casual talk between friends might include rather fast speech, whereas a more formal context requires the learner to speak more clearly and slowly. Additionally, speed of speech might be an individual characteristic, given that some speakers tend to talk rather quickly, whereas others might prefer pauses and speaking at a slower rate. Taking into consideration the notion of pauses, some students might prefer to use end-clause pauses to emphasise an argument or to provide the interlocutor with time to ask questions. An avoidance of such pauses would increase speed fluency scores, though at the expense of other rhetorical skills.

3.3 The notion of complexity

Complexity, as the term per se might imply, offers a more complex way of defining this performance domain. These multifaceted layers are mirrored not only in a more complex way of defining but also a variety of measurements that can be applied. It is used to describe both, task complexity and language complexity, i.e. L2 performance and proficiency (Housen & Kuiken 2009: 4). Taking into consideration recent studies (e.g. Skehan 2009b; Foster & Skehan 2009; Robinson 2003, 2001), an interdependent relationship between the meanings of complexity seems to be obvious, given that more complex tasks appear to foster more complex language production. Complexity, in further detail, does not only describe task complexity or the form of the produced language, but also its semantic complexity and the semantic relationships that are built through language use. Gass (1984: 11ff.) points the reader to the existence of semantic and structural complexity. The learner can thus add complexity to the produced utterance either semantically or syntactically.

Another aspect of complexity refers to the diverse features of the target language, i.e. code or linguistic complexity and the subcategory cognitive complexity which involves individual factors. The first aspect, linguistic complexity describes the complex language system used, whereas the later, cognitive complexity employs a learner-centred approach (Housen & Kuiken 2009: 3). In contrast to accuracy and fluency, cognitive complexity engenders more diverse aspects of the process of language production, such as individual learner differences, for example the willingness of risk-taking. Hence, cognitive complexity engenders learner-centred characteristics, meaning that the features of the target language may be well-known; however, the learner is yet not willing to use those features during language production. In other words, although the student might be perfectly aware of the linguistic structure, during task performance he/she could be still too anxious to employ it, thus he/she is more likely to rely on well-known and already automatised chunks instead of trying to complexify the output by the use of new and untested items. Even if this aspect of cognitive complexity refers to the individual learner, it plays an important role in complex language production. Teachers need to be prepared to support anxious learners and to provide a non-threating environment where the student feels free to test new hypothesis and to complexify his/her output.

Additionally, Skehan (2009b: 92f.) adds the aspect of restructuring, i.e. modification, to the notion of complexity, hence employing a more dynamic way of defining the performance domain of complexity (cf. Norris & Ortega 2009: 556). Modification just like using complex language is dependent on the student's willingness to negotiate and to take risks. However, as mentioned above, some students are more likely to approach language production in a more conservative way and *avoid* complex structures (Robinson 2001: 37). Hence, a lack of willingness to use "cutting edge" language (Skehan 2009b: 279), the so-called safety-first approach prevents the student to negotiate for meaning and to further complexify the output (cf. Skehan and Foster 1997). A reason for such a behaviour might be that the use of linguistically challenging, i.e. complex language can stress the learner when he/she is publicly testing L2 hypotheses about previously learned but yet not automated structures (cf. Skehan & Foster 2001: 190). Foster and Skehan (2009: 279f.) support the view that complexity reflects the student's willingness to use complex language by arguing that students who are prone to take risks are more likely to "use forms closer to the cutting edge" which

may push forward the learner's IL system (cf. Skehan & Foster 2007: 202). Hence, students should be encouraged to test language hypotheses rather than to rely on already automatised language chunks during task performance in order to progress learner language development. In contrast, Towell (2007: 260) describes the notion of complexity by stating that its basis is the learner's linguistic competence, i.e. cognitive skills to analyse input and language features. Complexity, in the sense of combining linguistic competence and willingness of risk taking, might allocate the student's attention to a more sophisticated or rich language production, thus provides the student with more ways of structuring the output (Foster & Skehan 2009: 279). Hence, a push towards the use of cutting edge language and the presentation of a variety of different language structures enables the student to feel free to produce more complex language which is assumed to prepare the way to acquisition (cf. R. Ellis 2009: 504).

Linguistic complexity, on the other side, can be seen as more stable than cognitive complexity, as it is not influenced through individual differences but refers to the properties of the language per se (Housen & Kuiken 2009: 5). In order to analyse complex language and complexified output production, it seems reasonable to define linguistic or code complexity in greater detail. Complex language generally is defined as "*size, elaborateness, richness* and *diversity* of the learner's linguistic L2 system" (Housen & Kuiken 2009: 5). In other words, linguistic complexity refers to the subcategories of lexis, structure, grammar and syntax of a language system.

Regarding lexis, pragmatic choices and lexical diversity might have an influence on syntactic and/or grammatical complexity, given that a more sophisticated choice of words may render complex phrasal structure redundant (cf. Pallotti 2009: 598f.). Syntactic complexity is often analysed by mean number of clauses per unit, i.e. subordination and coordination or mean length of clause (cf. Norris & Ortega 2009: 558). Additionally, utterance length, sophistication and diversity of structures help to define syntactic complexity (ibid.). According to Greenbaum and Quirk (1990: 283-303) a complex sentence may include the following features:

- o Subordination and coordination
  - o including finite, non-finite and verbless clauses
- o Direct or indirect speech
- o Transferred negation

Given that syntactic complexity is a multifaceted construct, it seems to be obvious that the measurements of complexity need to include this multidimensionality as well (cf. Norris & Ortega 2009: 562). Additionally, structural or syntactic complexity might be the result of functional discourse complexity, which implies that a more complex task might engender more complex language structure (cf. Robinson 2001: 35).

3.4 Measures of CAF

Regarding the definitions of accuracy, complexity and fluency, it is obvious that complexity and fluency present more than one means of measurement, whereas accuracy is solely measured by either the absence or presence of errors. It has also been argued that accuracy measures are based on errors, contrary to fluency and complexity measures which have a closer relation to development (Wolfe-Quintero et al 1998: 118). Housen and Kuiken (2009: 3) define accuracy as a form of correctness and argue that thus accuracy can be measured through deviation from the norm. Hence, a more adequate accuracy measure might be the number of error-free clauses, or error-free t-units (Wolfe-Quintero et al 1998: 62). Bygate (2009: 261) defines a t-unit as "a finite clause together with any subordinate clauses dependent on it". Given that in spoken speech students often do not finish their own sentences, i.e. t-units, errors per clause or error-free clauses appear to represent a more appropriate and comparable way of measurement.

Fluency represents several subcategories which can be measured separately; three which are rather popular measurements are the following:

1. Rate and density of delivery measuring *speed of fluency*
2. Speech pauses indicating *breakdown fluency*
3. False starts and repetition to measure *repair fluency* (cf. Housen & Kuiken 2009: 4; Skehan & Foster 2009: 281)

Regarding the number of pauses during speech, it seems to be natural, especially for native speakers, to make pauses during speech, particularly at the end of clause. On the contrary, second language learners show a tendency to pause mid-clause which would be rather odd for native speaker behaviour (Skehan & Foster 2009: 281). Hence, when

measuring fluency by the number of pauses, it seems to be advisable to focus on mid-clause pauses than on end-of-clause pauses.

In order to measure complexity, typically syntactic complexity is in the centre of attention, thus the number of subordination and coordination per clause or unit is frequently measured. The number of subordination per clause, t-unit or AS-unit is, therefore, the most common way of measuring complexity in spoken interaction (cf. Ellis & Barkhuizen 2005: 152ff.). Norris and Ortega (2009: 563) stated that coordination is rather likely to indicate complex language when analysing beginners, whereas measuring subordination represents a useful tool for measuring complex language produced by intermediate learners. Even though measuring subordination and coordination presented itself as a rather useful and successful tool, some researchers (e.g. Schleppegrell 2004; Pullum 2014; Deutscher 2009) argued that subordinate clauses in the context of measuring complexity are described too simplistically. In other words, subordination does not equal subordination and different subordinate clauses can have different influences on language complexity (cf. Schleppegrell 2004: 13f.). Geoffrey K. Pullum argued during a guest lecture (*English Grammar as a Domain of Scientific Exploration*, 26.11.2014, Universtiy of Vienna) that the traditional way of defining subordination and coordination should be seen critically. He further mentioned that noun clauses, adverb clauses and adjective clauses might be represent the wrong choice in terminology as those clauses do not at all act like a noun, an adverb or an adjective. The individual features of the languages are, according to Pullum (2014 guest lecture), spuriously and multiply categorised. He proposes another classification of subordinate clauses:

- o Declarative clauses

    e.g. Billy said that he was not home.
- o Closed interrogated content clauses

    e.g. She does not know whether she likes it or not.
- o Open interrogated content clauses

    e.g. wh-phrases or gaps
- o Relative clauses

    e.g. phrases that include that or which
- o Comparative clauses

Regarding Pullum's point of view, this study will include this argumentation into the analysis. Deutscher (2009: 199ff.) argues that nominalization as an important aspect of subordination plays as an important role for complex output production. Taking into consideration learner proficiency, it might be argued that nominalization is more likely to occur in the speech of more proficient and advanced learners, thus it will not be included in this study as the participants are at an intermediate level. Robinson (2001) also proposed to include the use of logical connectors as well as verb-argument structure into the analysis of language complexity. Regarding lexis, a frequently used measure to indicate lexical complexity is the type token ratio, i.e. word or word type per total number of words (cf. Wolfe-Quintero et al. 1998: 102). One negative aspect of this measurement is that task length can influence the results, therefore, a word limit, e.g. word type per 300 words, might be added to avoid a distortion of the results (ibid.). Taking into consideration the employment of different tenses, avoidance of base forms and inclusion of a range of syntactic forms can indicate grammatical variety and complexity (cf. Skehan & Foster 2001: 281).

Regarding the different measurements and ways of displaying accuracy, fluency and complexity, there seems to be a tendency to interpret higher scores on the performance domains to equal better speech production. However, Pallotti (2009) raised an important critique towards such interpretations by arguing that high scores do not always indicate a better or more appropriate speech production. Situational influences might require different speed of language production, i.e. fluency, as well as a highly complex speech may not be adequate in some contexts. It is further argued that instead of solely analysing the complexity of utterances, there is a need to compare the produced output to goal achievement and communicational skills (Pallotti 2009. 597). Furthermore, at a certain level learners tend to employ less syntactic complexity (measured in subordination and coordination) but more elaborate and sophisticated lexis. At this stage, sophisticated lexis might be more appropriate and indicate a higher proficiency level than complex language structure and long phrases.

Additionally, a criticism regarding the measurement of the performance domains and the question of adequacy and appropriateness was pointed out (cf. Pallotti 2009: 596ff.). Achieving communicative goals during task performance might be seen as successfully performing the task and helps the learner to develop his/her IL language. Further, Pallotti (2009: 595) invited the reader to question the idea that any form of task

complexity increases complexity levels of all linguistic features. Taking into consideration recent studies (e.g. Skehan 2009a; R. Ellis 2009a; Skehan & Foster 2009; Norris & Ortega 2009 among others) adequate or appropriate language along with complex language use was not considered and or included during the interpretation and discussion of most results (cf. Pallotti 2009). Taking into consideration that language should be used adequately, one need to keep in mind that appropriateness is strongly connected to context (cf. Widdowson 2003: 93-109). In other words, in the context of a dialogic task between students, highly complex language might not be as appropriate as rather simple and quick turn taking. Thus causing a decrease in complexity scores, however, rendering task performance more adequate.

4 The interplay of Complexity, Accuracy and Fluency

The three main goals for learners of a second language which can be positively influenced, either through priming or through task design and activities, are accuracy, fluency and complexity. In order to promote a beneficial development of complexity, accuracy and fluency, it is important to understand the interdependent relationship of cognition and task. Regarding task design and the knowledge of attention, the relationship of explicit and implicit knowledge, as well as working memory, it appears to be obvious that attentional factors need to be included into task design and performance (Skehan 2009b: 90). But not only attention and consciousness influence successful language development, the learner his/herself needs to aim for an improvement of CAF (Skehan 2009b: 93). By complexifying his/her interlanguage system, as well as becoming more accurate and fluent, the learner is not only more likely to be "accepted as a speaker of the language concerned" (Skehan 2009b: 93), it can also be argued that the employment of complex structures are seen as the manifestation of positive interlanguage development (ibid.).

The development of fluency, complexity and accuracy is strongly connected to working memory capacities. It is argued that interactions between the visual long-term memory and the visuo-spatial sketchpad act as some sort of basis for the automatisation process, i.e. chunking for fluency development (cf. N. Ellis 2001: 35). In other words, familiarity, language chunking and frequent practice to incorporate such chunks into the interlanguage system contribute to fluent speech in the L2. In contrast, it is claimed that complex and unknown situations, where students might not have ready-made schemes and chunks for problem solving and where they need to reason and articulate complex ideas simultaneously during task performance, may be more likely to engender the development of more complex language (cf. N. Ellis 2001: 37). Similarly, Foster, Tonkyn and Wigglesworth (2000: 356) argued that remembering and reciting automatised chunks is not an indication for complex language. However, students who are prepared and can rely on language chunks may show more willingness to take risks and to test new hypotheses about the target language. Regarding the interplay of the three performance domains, there are several hypotheses on how accuracy, fluency and complexity correlate.

4.1 The Trade-Off Hypothesis

It has been argued that it might be difficult for the learner to focus on all three domains of performance concurrently. Hence, a trade-off effect between accuracy, complexity and fluency appears to be a legitimate possibility during task performance (cf. R. Ellis 2005: 140). Limitations to working memory capacities might cause such trade-offs, given that content processing and linguistic processing happen concurrently (cf. R. Ellis 2005: 141). Assuming that the learner's attentional capacities are limited, it seems to be rather unlikely that the learner is able to pay equal attention to all three domains while performing a task (Skehan 2009a: 512; Skehan & Foster 2007: 202). One aspect of the trade-off hypothesis supports the view that complexity might benefit at the expense of accuracy, due to the learner's limited attentional space causing a lack to focus on both domains at the same time (cf. Skehan & Foster 2001; Skehan 2009a). Skehan and Foster (1997: 201f.) state that learners showed a clear tendency to produce less accurate language if the task as well as preparation during planning time demanded a more complex output production, whereas tasks with less cognitive challenges provided a better opportunity to focus on accuracy. Hence, the trade-off hypothesis provides a more competitive view of accuracy, complexity and fluency and claims that those three performance domains are rather unlikely to benefit simultaneously (cf. Skehan and Foster 1997: 204).

Assuming an interdependent relationship between complexity and fluency, it can be argued that the more complex the produced language is, the less fluent it becomes due to a need for pauses to think about what is said (cf. Bygate 2009: 264). Similarly, Bygate (2009: 264) claims that also during task repetition complexity increases at the expense of fluency and accuracy. Foster and Skehan (2009: 295) try to explain this phenomenon by arguing that the students might feel a need to prefer one domain, thus, other domains suffer under the allocation of attention towards the favoured domain. Housen and Kuiken (2009: 7) agree by stating that complexity, accuracy and fluency are more likely to have a competitive relationship, where fluency and accuracy similar as accuracy and complexity contend against one another.

Another influential factor of the trade-off hypothesis is that different tasks might have different influences on complexity, accuracy and fluency as well. Interactive or dialogic tasks appear to elicit more complex language among learners than narrative or

monologic tasks, which, in contrast, might enhance fluent speech (cf. Bygate 2009: 264). Given that some trade-off effects might be present during task performance, studies have shown that appropriately challenging tasks elicit both an increase in accuracy and complexity, although this does not happen as frequently as the co-occurrence of fluency and accuracy or fluency and complexity (Skehan & Foster 2007: 203). Skehan and Foster (2001: 188) support the idea of task performance and task demands high interdependency by stating that cognitively challenging tasks can "push" the learner to particular performances. In other words, complex tasks might cause complex and maybe also more accurate language, simple tasks might equal simple though fluent and automatised language production and practice (Skehan & Foster 2001: 188; Bygate 2009). During the performance of complex or demanding tasks, for example reasoning tasks, the learner has to be more attentive, implying that more memory capacities are used (Robinson et al 2014: 255). Michel, Kuiken and Vedder (2007) compared dialogic and monologic tasks along with task complexity and observed that the complex tasks prompted the students to utter more accurate language in connection with an increase in lexical complexity. Regarding fluency, their results (Michel, Kuiken & Vedder 2007: 248) correlate with Skehan's and Foster's (2001) as well as with Bygate's (2009) assumption that simpler tasks evoke more fluent language than complex tasks. Considering the difference of dialogic and monologic tasks, dialogic tasks proved to engender more fluent language and fewer occurrences of errors than monologic tasks, however, at the expense of structural complexity (Michel, Kuiken & Vedder 2007: 252ff.). Michel, Kuiken and Vedder (2007: 255) argued that fluency might be increased due to "quick turn-taking", however, stating that time pressure reduced structural complexity. With regards to negotiation of meaning, interactive tasks appear to "push" (cf. Swain 1985) the learners to modify more frequently than during monologic task performance (cf. Michel, Kuiken & Vedder 2007: 254f.). Similarly, Robinson (2007b: 209) observed that there tended to be more interaction as well as a greater amount of negotiation of meaning in complex tasks.

If the content (or the task design) is unfamiliar or too challenging, the learner needs to divide his/her attentional capacities to reasoning and language production. In other words, output production might be neglected as focus lies primarily on content comprehension (Bygate 2009: 252). Considering this argument, form *or* meaning can be the focus of attention, and even if the emphasis lies on form, it is argued that the learner

might attentively produce either accurate *or* complex language, but it seems very unlikely that they co-occur (Robinson et al 2014: 254f.; Skehan & Foster 2007: 203). Skehan and Foster (1997: 206) agree by stating that "the need […] to express the complex relationships seemed to push learners to complex language", simultaneously limiting attentional space for accurate language production (ibid.). It has also been argued that students with the ability to produce more complex language are also more likely to articulate more complex ideas (cf. Skehan 2009b: 93) which can be interpreted as an indication for an interdependent relationship of complex language and complex tasks. The results of another study (Foster & Skehan 2009: 293) correlate with this assumption, since the performance during the more complex tasks seemed to engender higher scores for complexity, however, at the expense of accuracy scores. Similarly, as already mentioned, negotiation of meaning seems to occur more frequently during complex task performance as the student's need for negotiation is higher and, consequently, more opportunities for output production and reflection on own output are provided (cf. Robinson 2003: 651; 2001).

Hence, the learner seems to be "pushed" (cf. Swain 1985: 249) to do both, produce output and notice the difference between own output and the target language, i.e. input (cf. Gass & Mackey 2006: 13; Robinson 2003: 651). Regarding fluency, however, it seems to be more likely that students who are familiar with task design and/or content speak more fluently than learners who need to focus on task encoding and content comprehension (cf. Bygate 2009). Unfamiliarity with either content or task design appears to cause speech dysfluency (Robinson 2003: 649). Complex tasks as well can have a disadvantageous effect on fluency, given that cognitively demanding tasks cause more negotiation and disruption. The above described disagreement is in some way supported by the results of Skehan's and Foster's (1997) study, which show that – even though the focus was on planning – there is mixed evidence for the correlation of task complexity and complex, fluent and accurate language production (Skehan & Foster 1997: 198). Strikingly, Foster's and Skehan's (2009) study unravels that the task with the highest cognitive load – a decision making task – seemed to cause both, a high level of accuracy as well as complexity. Although the scores for both were not at their highest, given that other task types caused either higher complexity with lower accuracy scores or *vice versa*, they were still at a beneficial level for the development of the IL system. It is further suggested that task type and planning

conditions, when combined successfully, allow the student to produce complex *and* accurate language at the same time (Foster & Skehan 2009: 295). Robinson (2001: 49) agrees by indicating that there is a tendency for greater accuracy in complex tasks.

Recently, studies have shown that even though a trade-off might be happening, simultaneous increase in accuracy and complexity appears to be legitimate (cf. Skehan & Foster 2007: 213). Complexity and accuracy might co-occur during task performance, task difficulty may not be the only influence though, given that task structure promotes accuracy, whereas complexity benefits from integration of information (Skehan & Foster 2007: 214).

4.2   The Cognition Hypothesis

In contrast to the trade-off hypothesis, Robinson (2001; 2003) proposes the Cognition Hypothesis, where attentional space is not limited and task complexity stimulates rather than hinders the learner to produce both, more complex and more accurate language. Especially lexis is assumed to benefit from more complex tasks, as a wider range of vocabulary and a more appropriate and accurate language might be used to fulfil the cognitively more demanding task (Robinson 2001: 46f.). Michel's, Kuiken's and Vedder's (2007: 248) confirm this assumption to some extent by arguing that there were higher accuracy scores along with an increase in lexical complexity during complex tasks, however, accompanied by a decrease in structural complexity. In other words, both domains of performance which concern form, i.e. accuracy and complexity, can increase concurrently when a task is cognitively more demanding and challenging. Given that the student is more consciously involved, it appears to be evident that more attention is directed towards output production and modification (Robinson et al 2014: 255; Robinson 2001). Hence, it seems to be legitimate that allocation of attention is assumed to occur more likely during complex tasks, i.e. cognitively demanding tasks (Robinson et al 2014: 248), rather than during simple and familiar tasks. The cognition hypothesis (Robinson 2001; 2003) further proposes that more complex and demanding tasks do not only increase the learner's attention but simultaneously create a more real-life like situation in the classroom and approach the needs students are likely to encounter outside of the classroom (Robinson 2003: 648; Robinson 2001). However, more complex tasks might engender more need for

negotiation, i.e. more confirmation checks and clarification requests which, due to rather short yes/no answers, may disrupt fluent speech and, consequently might have a negative influence on complexity due to shorter utterances (Robinson 2001: 36, 47). Taking into consideration that those utterances operate as clarification checks or confirmation and that such "elliptical answers" (Robinson 2001: 36) may distort the results, a qualitative approach might be more appropriate and give more insight into the complexity of the modified output. Hence, single-word answers will not be included in the analysis and when negotiation of meaning takes place, a more qualitative and detailed analysis will be applied.

In order to complexify cognitive demands of the task, several factors can be changed to increase task complexity. Taking into consideration the triadic componential framework for task classification (cf. Robinson 2007a), There-and-Then tasks are more challenging for the learner than Here-and-Now tasks (Gilabert 2007: 51; Robinson & Gilabert 2007: 164f.; Robinson 2007a: 17.). Considering the trade-off hypothesis, Iwashita et al. (2001) assumed that the less complex the task, the more fluent and accurate the outcome. However, the study showed that in the more complex There-and-Then task, students tended to speak more accurately (cf. Gilabert 2007: 52) thus supporting the cognition hypothesis. Similarly, Gilabert's (2007: 62) results confirm the cognition hypothesis to some extent, given that students seemed to focus more on form at the expense of fluency during the There-and-Then tasks. However, during the simpler Here-and-Now task fluency scores only increased slightly (Gilabert 2007: 62), which might indicate a general dysfluent speech among the participants of this study.

When comparing the trade-off hypothesis and the cognition hypothesis, it is obvious that there is some disagreement concerning the relationship of task complexity and output production. It is argued by some researchers (e.g. Skehan 2009b; Bygate 2009) that students are more likely to use complex language if the cognitive load of the task is not too demanding and more attentional space can be used for language production, i.e. focus on accuracy, complexity and fluency, rather than on problem solving, i.e. focus on content (cf. Schmidt 2001; Mackey 2012; Skehan 2009b). Evidence for this argument can be found in beneficial effects of rehearsal or repetition of tasks on accuracy and complexity (cf. Bygate 2009; R. Ellis 2009; Skehan & Foster 1997; Foster & Skehan 1996).

Attention and language learning are both dynamic concepts, where task complexity and its influence might vary between the different stages in the development of the interlanguage system. Comparing now the trade-off hypothesis and the cognition hypothesis to the *Comprehensible Input* and *Comprehensible Output Hypothesis* it seems to be obvious that a task should be neither too simple nor too complex but slightly challenging − linguistically *and* pragmatically − in order to stimulate a developing process in the learner's interlanguage system (cf. Swain 1985; Krashen 1985; Skehan & Foster 2009). Moreover, a diverse offer of activities ranging from rather simple tasks to cognitively demanding ones might prevent the students from either feeling unchallenged or overwhelmed and contributes to a varied language classroom, thus, a balanced development of all the three performance domains, accuracy, fluency and complexity.

4.3   The influence of planning

Planning time has been in the focus of several recent studies and has been argued to positively influence some performance domains. Given that planning time is also seen as a beneficial factor for the creation of long term memory, it additionally helps the student to focus his/her attention during actual task performance on certain aspects (cf. N. Ellis 2005: 318). Regarding planning time and attention allocution, it seems to be legitimate to argue that attentive or guided planning might be more beneficial than planning time which does not provide a certain focus (R. Ellis 2009: 495). Planning, however, seems to have a more beneficial effect on fluency during simple tasks than during complex task performance (cf. R. Ellis 2009). A similarly positive effect of planning on accuracy can be found in Foster's and Skehan's (2009) study, where accuracy increased at least in some tasks where planning time was provided. In contrast, Yuan and Ellis (2003:7) as well as Gilabert (2007: 62) found that pre-task planning only had insignificant impact on accuracy scores. However, online planning caused an increase in accuracy scores as well as complexity scores (Yuan & Ellis 2003: 23). It might be argued that the lack of an increase of accuracy despite planning time does not show limited attention capacities but that the planning time was not used to prepare for more accurate speech (Gilabert 2007: 62). R. Ellis (2009: 501) observed that dysfluency and repairs reduced as the task progressed, implying that the

student's need for online planning diminishes during task performance. Regarding fluency, it can be argued that tasks with planning time cause more fluent speech among students than tasks without planning time (Skehan & Foster 1997: 196). Similarly, Gilabert (2007: 49) argues that pre-task planning appears to benefit fluency and complexity, whereas accuracy is more likely to benefit from on-line planning, i.e. planning during task performance. Even though different forms of planning seem to have an effect on different dimensions of performance, task type and task familiarity as well play an important role for a good preparation for actual task performance (cf. R. Ellis 2009: 497). Additionally, planning time may provide the learner with enough time to decide on more appropriate semantic choices and can help him/her to employ a wider range of vocabulary (cf. Skehan 2009b: 515). Gilabert's (2007: 62) study confirm this assumption by showing the result that pre-task planning time seemed to cause more fluency as well as a greater variety of lexis. Taking into consideration time pressure, i.e. lack of either pre-task planning time or on-line planning time, Foster, Tonkyn and Wigglesworth (2000: 356) stated that such time pressure can have some negative effects on complexity, accuracy and fluency. Interestingly, chunking and idiomatic language appears to occur more frequently when students feel pressured, given that "word-by-word processing" seems to decrease along increasing pressure (Foster, Tonkyn & Wigglesworth 2000: 356). This phenomenon might be explained through the hypothesis that when the learner is under time pressure, the communicative goal is given prominence and the student does not have enough time to modify the output or to produce a more complex structured utterance (Skehan 2009b: 93). Hence, time pressure might add a more real life-like aspect to the task, however, when the emphasis lies on complex language production, planning time might be a useful addition to task preparation.

Complexity and fluency seem to be positively affected through rehearsal – with and without feedback during the first performance – whereas a progress in accuracy appears to depend on some sort of corrective feedback (R. Ellis 2009: 476f.). Interestingly, Bygate (2009: 253) also found evidence that task repetition benefited the learner's accuracy during second task performance, though solely in certain groups. Fluency seems to benefit most from a familiarisation process with task types, which means that practice of a certain type of task showed increased fluency, however, a similar tendency was found in accuracy and complexity during performance of familiar

task types (Bygate 2009: 267). Other studies (e.g. Foster & Skehan 2009) have also shown that a general familiarity with task design does not cause a significant increase in complex language production. Planning, especially detailed planning, however, can enable the learner to produce more complex language (Foster & Skehan 2009: 290; Skehan & Foster 1997: 198). Contradicting R. Ellis (2009) hypothesis, Bygate's (2009: 265f.) study suggests that complexity seems to benefit more significantly from task repetition, accompanied by a *loss* in fluency. It is further argued (Bygate: 2009: 266) that a trade-off effect in task repetition is more likely, given that students seem to focus more on complex than fluent or accurate language production when doing a task a second time.

Generally, it can be argued that planning time might have positive effects on fluency, however, for complexity and accuracy the results are not that clear and need further research. Given the above mentioned hypotheses about the influence of planning time, during this study, planning time was included into the analysis.

5   Task-based language teaching and Classroom discourse

5.1   Classroom discourse

The foreign or second language classroom represents a very wide range of different discourses and means of exploring language acquisition and learning. It is not only different from ordinary discourse in many ways but it entails quite complex processes regarding the acquisition process of particular features of the L2. How those features are then absorbed and acquired by the individual learners in some cases might be described as "the frightening complexity" (Seedhouse & Walsh 2010: 135) of classroom discourse. A crucial feature of classroom discourse analysis which should be kept in mind while analysing learner language is the concept that learning and interaction cannot be simply separated but need to be seen as two sides of one coin (cf. Seedhouse & Walsh 2010: 138). Additionally to this interdependence, environmental and situational context too can have an enormous influence on language performance (cf. Seedhouse & Walsh 2010: 139). An exclusion of the listed aspects from the analysis might distort the results and therefore the learner's language development or current state in the language learning process would not be presented appropriately.

Classroom discourse differs from out-of-the-classroom discourse in certain ways which are important for learner language analysis. According to Walsh (2011: 4; 2007) there are four main features of classroom communication which influence learner talk: control of interaction, speech modification, elicitation and repair. All of which are organised and mostly employed by the teacher who typically dominates classroom talk. In other words, even though language production seems to be the main goal in the communicative language classroom, learners' own output production is rather limited and generally supervised. This asymmetrical contribution of learner-teacher roles in the classroom leads to classroom-typical move structures, for example the Initiation-Response-Feedback (IRF) structure (cf. Walsh 2011: 4, 17). However, such turn taking sequences where solely one person initiates and provides feedback, whereas the interlocutor is frequently limited to one or two words per answer, is scarcely found in communication outside of the classroom. R. Ellis (1994: 575) supports the view of such an asymmetrical contribution and typical classroom discourse connected with IRF, however he further states that quality has a great influence on learner participation,

given that it seems more likely for a student to participate if a certain proficiency level is achieved (R. Ellis 1994: 592). If typical classroom interaction mainly consists of Initiation-Response-Feedback sequences, free and qualitative output production is limited to a high extent. This might hinder language acquisition, assuming that qualitative output production provides a major contribution to interlanguage development and represents one of the main principles of Swain's (2005; 1985) output hypothesis. Therefore, R. Ellis (1994: 598), as well as many others, implicitly proposes the inclusion of a wide range of different activity types, for example small-group work and pair interaction, to provide the students with varied opportunities to produce more output.

An illustrative example of teacher domination in classroom talk is the management of topic and turn taking (Walsh 2011: 4). This kind of management is very typical for classroom discourse where topics and aims of the lessons are defined by the teacher, whereas in informal talk the interlocutors might jointly decide on topic and speaking time for the individual participants. Similarly, Seedhouse and Walsh (2010: 131) point out that in a classroom, contrasting ordinary conversation, discourse is more likely to be goal or objective oriented and does not solely serve the purpose to convey meaning. Hence, classroom discourse is not only teacher dominated but also goal oriented, as aims and objectives – although defined partly by the teacher and partly by the government – are in the main focus of the language learning process as well as the individual lessons (cf. Walsh 2002). Taking into consideration that teachers as well as learners focus on established goals, it seems to be obvious that the produced language is not only monitored for correctness but also to fulfil pedagogical purposes (cf. Seedhouse & Walsh 2010: 131).

Classroom discourse thus provides learners of a second language with many opportunities to take input in, as well as to test own hypotheses about language structures. Given that output production and hypotheses testing might cause some evidence of incorrect utterances, corrective feedback or error correction represents another aspect of classroom discourse which might not be found as frequently in ordinary discourse. A popular and prominent form of corrective feedback is *recast*, i.e. mostly a teacher-provided correction of a form that has been uttered incorrectly by a student without any explicit information about the correction or the form. It is very similar to the so-called teacher echo, i.e. a teacher-repetition of a student's utterance.

However, recast involves some sort of modification due to incorrect and/or improper use. Even though it might seem like a rather economical technique, studies (e.g. Mackey and Philp 1998) have shown that recast seems to be more beneficial for advanced learners than beginners, since recast appears to cause an increase of L2 production and modification in interaction of advanced learners than interaction without any evidence of recast (cf. Mackey 2012: 26). In other words, advanced learners are more likely to recognise recasts as a form of correction, thus tend to restructure their previously uttered output more often when recasts occur. One criticism against this form of correction was uttered by Mackey (1999: 575ff.) and Gass and Mackey (2006: 9) by the argument that the learner might not recognize the recast as a form of correction but might assume that it is solely a way of rephrasing what has been said. Braidi (2002: 31) supports this view by stating that one cannot confirm whether the student recognises recasts as corrective feedback or see them as confirmation checks on what has been said. Even though it might seem that not all of those recasts are perceived as a form of corrective feedback, studies have shown that learners detect those forms, even if not immediately, and reuse them, hence integrate them into their interlanguage system (cf. Mackey 2012: 19, 31). This phenomenon is called *priming* (Mackey 2012; McDonough 2006; DeJong 2009: 97). Although there might not be an immediate effect on the interlanguage system, corrective feedback seems to enhance language development. In particular, the notion of priming might also have beneficial effects on the complexification of output production. In other words, students may internalize some complex structures when they are provided by the teacher during feedback sessions even if it does not result in immediate modification. However, there is no assurance that recast evokes a noticing process, thus contributes positively to language uptake. It might be argued that recast is more an *economic* way of pointing towards a more appropriate language use rather than a beneficial and influential type of corrective feedback (cf. Lyster & Ranta 2013: 170ff.). Recast, therefore, might be a useful tool to enforce fluency practice (cf. Walsh 2002: 10f.), but in a classroom where solely recasting is used to provide error correction, students' accuracy as well as complexity could suffer. Hence, a variation of forms of corrective feedback might be considered as a more appropriate way of approaching error-treatment in the language classroom (cf. Lyster & Ranta 2013: 180).

Considering all those classroom particularities, one of the most striking features still needs to be elaborated: the fact that the tool of learning and teaching is at the same time the object of learning – a second or foreign language. It is obvious that there is a difference in proficiency of teacher talk and learner talk which results in another asymmetry and stresses the importance to distinguish teacher talk from learner language, the so-called interlanguage. Some researchers (e.g. Gass 1984) argued that there is some sort of "universal order of acquisition independent of native language background" (Gass 1984: 3) which strongly accounts for a universal form of interlanguage. However, this study will not take this universal hypothesis into consideration. In the literature, interlanguage (IL) is defined as "some kind of 'in-between' grammar and language" (Selinker 1992: 212f.) In other words, learner language does not simply consist of correct and incorrect phrases or utterances, but it has a structure of its own (Selinker 1992: 222). It is further argued (Ellis & Shintani 2014: 5) that incorrect utterances are not happening arbitrarily but form part of the learner's "creative construction of the L2" (Ellis & Shintani 2014: 5). Another typical characteristic of the acquisition process of a second language is that the language most certainly does not become simpler over time. Individual grammatical features as well as the complete interlanguage system develop into a more and more complex language representation which is supported by previously learned and/or acquired structures (R. Ellis 1994: 350). In classroom discourse, this understanding of learner talk entails the hypothesis that simple error analysis and corrective feedback without any acknowledgement of the existence of an interlanguage might lead to learner discouragement. Errors might not simply indicate the absence of learning or the unwillingness to learn, but that the learner is in a process of acquiring this particular feature of the second or foreign language (cf. Ellis & Shintani 2014: 6). Additionally to potential discouragement, one particularity of classroom discourse is that the student's utterances and language production can be evaluated by the teacher, who in many cases is not only the assessor but also the interlocutor (cf. Seedhouse & Walsh 2010: 132). This might create anxiety among some learners and can have negative effects on language production as affective filters might be high instead of low. Assessment of L2 use, however, is a domain which is not solely dominated by the teacher. Students might use self- or other-student-initiated modification and/or correction as well to present their own ability of assessment simultaneously with their awareness of the institutional interplay of pedagogy and interaction (cf. Seedhouse & Walsh 2010: 137). In other

words, output production in the classroom consists of different layers of performance which the student needs to be aware of to allow successful interaction. However, what is important is that even though the term classroom discourse might give the impression that it is a more or less collective performance, students' utterances and produced outputs are always the result of "individual performance[s]" (Walsh 2012: 2).

Interlanguage, besides other implications of classroom discourse, entails some important consequences for the interpretation of data, given the fact that it is likely there are more categories than just correct or incorrect utterances. In the case of complexity this would mean that an utterance could still be complex, even though it might not be accurate. As interlanguage cannot be seen as a static concept but a dynamic process, means of progress making are essential for developing the learner's language proficiency. This implies that the learner of a second language may have to repeat several stages in the language learning process in order to positively influence his/her interlanguage development. In order to provide an accurate picture of the learner's performance in the classroom, one needs to keep in mind that the language learning process has not only a complex structure but is additionally very student as well as situation dependent. A very significant means to provide the student with numerous ways for language practice is the use of tasks and activities (cf. Skehan & Foster 2001: 186).

5.2   Task-based language teaching

There are numerous ways of describing or categorising tasks and activity types and it seems that several researchers recently focussed increasingly on the influence of tasks or task-based language teaching (TBLT) on language performance (e.g. Foster & Skehan 2009; Gass & Mackey 2007; Mackey 2012; R. Ellis 2009, 2003; McDonough 2006; Van den Brandon 2006). Task-based language teaching is comprised of several features: the main aspects are that the core focus lies on meaning, a need for communication with another person is present, the learner draws from his/her own linguistic and non-linguistic knowledge and resources and, finally, the goal needs to represent more than simply *using* the target language (Mackey 2012: 57). Regarding the last point, this feature implies that the learner is expected to show a development when using the target language, assuming that tasks and activities can help him/her to achieve

such a progress during task performance. Mackey (2012: 58) further argues that even though conveying meaning plays a crucial part in TBLT, the learner's attention needs to be guided and directed towards formal features, i.e. linguistic features of the L2. What is important is that all those processes, i.e. drawing attention to form, focus on communication with another person and using own linguistic resources, are happening in a meaningful context that provides the learner with opportunities to employ his/her knowledge of the L2 appropriately. What is important is the fact that an interactional process needs to be created "since it is through [this] interactional [process] that learning opportunities can occur" (Mackey 2012: 59). In other words, one of the main features of TBLT is the aspect of *action*, which means in order to learn and acquire second language *skills*, students need to be actively involved in the *tasks* that have been the main interest of recent interactional research (Mackey 2012: 58). Van den Branden (2006: 6) states that TBLT can be described as more or less learning-*by*-doing and not only learning *how* to do it.

Van den Branden's (2006: 4) definition of a task as "an activity in which a person engages in order to attain an objective, and which necessitates the use of language" correlates perfectly with the above described objectives of task based language teaching as active student involvement into the classroom discourse. A task includes both, it is goal-driven and simultaneously the objective of language learning (and teaching) as the task is what "language learners need to be able to perform" (Van den Branden, Bygate & Norris 2009: 6). Another important criterion of a task is represented by the fact that the activities used in the language classroom are not some sort of artificial invention of the teacher's mind, but need to correlate to real life situations that language students will need to cope with outside of school (Van den Branden 2006: 6). In order to match the task criteria, Ellis and Shintani (2014: 135f.) argue that a task needs to primarily focus on meaning, i.e. semantic and pragmatic meaning, a need for communication, i.e. a "gap" (ibid.) should be present, the resources used during task performance come from the learners themselves not from the teacher and, finally, a goal or outcome needs to be described before starting the task. Interestingly and contrasting to teacher-centred classrooms, the teacher does not serve as source of information or monitor, but the task itself acts as a stimulus for natural language use when learners are "performing real-life language tasks" (Van den Branden 2006: 9). In contrast, an exercise only represents language practice, where no focus on

meaning is required (Ellis & Shintani 2014: 136). Examples are grammar exercises with incoherent sentences for practice.

Different task designs include numerous ways of providing students with opportunities for L2 production, as well as modification when negotiation of meaning occurs (Mackey 2012: 22). Typically, when talking about classroom discourse and task performance, one assumes that non-native speaker interactions are more common than native and non-native interactions. Mackey (2012: 22), however, claims that task designs include both situations, either interactions with another non-native speaker, i.e. learner of the target language or interaction with native speakers. Taking into consideration the context of an Austrian classroom, it seems to be rather unlikely that students have as many interactions with a native speaker as with a "comparably skilled or competent learner" (Mackey 2012: 22). In contrast, including native-like speakers, i.e. the teachers of the language, this assumption seems to become more likely. Given that tasks and activities are more likely to elicit *learner* talk, even the assumption that conversations with native-like speakers, the teachers, seems to apply more for whole classroom discussions than for task performance. Since classroom activities present a very wide range to choose from, some of the most frequently used task types in second language and interaction research will be described in more detail. Three of the most common categories for task type differentiation are: (1) open vs. closed tasks, (2) one-way vs. two-way tasks and (3) required information exchange vs. optional information exchange (cf. Mackey 2012: 22ff.).

The main distinction between an open and a closed task is that a closed task provides the student with a tighter structure and the teacher's expectation of finite answers, whereas an open task leaves room for creative language use which anxious learners might experience as a more threatening situation than closed tasks. A typical example for closed tasks is a fill-the-gaps/blanks activity with provided words, whereas an open task might ask the student to come up with an own ending for a story or to talk about his/her own experience with a certain topic. Reciprocity is another category for task description, as the task is either one-way, i.e. non-reciprocal, or two-way, i.e. interaction is needed to achieve a goal. A monologue or holding a presentation is an example for the one-way tasks, in contrast, dialogues, discussions or role plays represent two or even three-way tasks. There is much discussion going on regarding the beneficial effects of both one-way and two-way tasks. It is argued that while two-way

tasks seem to engender more evidence of negotiation of meaning among learners, one-way tasks provide the student with more possibilities of output production and modification (Mackey 2012: 24). A similar bipolar view is found in the required information exchange and optional information exchange distinction, given that in required information exchange negotiation of meaning seems to occur more often than in optional exchange tasks (Mackey 2012: 24f.). Concluding, one could say that different tasks provide the student with opportunities to focus on an improvement of different language skills and competences. These different hypotheses about tasks and their influence on language learning show that there is not *one* right way of teaching a second language meaning that there is not *one* task to learn all the different aspects of a new language. It seems to be obvious that task variety and diversity of activity types are not only a welcome change but a necessity for successful second language acquisition.

Regarding task performance studies, some popular tasks types, typically found in interaction research, are spot-the difference tasks, story completion or picture sequencing tasks, jigsaw tasks, consensus tasks and dictogloss tasks (Gass & Mackey 2007; Mackey 2012: 19, 23f.). Even though these task types are frequently used in interactional research, one might raise the question if they are as commonly used in the everyday language classroom. Lynch (1997: 323f.) emphasises the fact that the development of particular classroom activities might be hindered through the attempt of general pedagogical objectives achievement. In other words, some situations may include communication difficulties which are not seen (by some teachers) as a possibility for negotiation of meaning but more as a problematic situation which needs to be solved to successfully fulfil the task. Assuming that some teachers show a tendency to interfere rather than let the student try to solve interactional problems his/herself, teacher involvement might be seen as an interruptive event in students' conversation or negotiation by intervening and providing a *proper* way of doing the activity. Hence, such teacher involvement represents a hindrance to the development of the student's interactional and problem solving skills.

Robinson (2007a: 9f.) provides three approaches to task classification, behaviour descriptive approaches, information-theoretic approaches and ability requirements approaches. The first approach, the behaviour descriptive approach, focuses on observing "what people actually do while performing a task" (Robinson 2007a:9), the second, the information-theoretic approach, emphasises cognitive processes and

information processing stages which are employed to fulfil the task successfully (ibid.) Finally, the third approach, the ability requirements approach, includes individual differences and their effects on task outcome (Robinson 2007a:10). A more detailed way of task type classification is, on the one hand, the distinction between listing, ordering and sorting, comparing, problem-solving, sharing personal experiences and creative tasks (Willis 1996). This way of classification can be used in a typical task sequence, starting with listing some information for a certain topic, followed by ordering and sorting, i.e. categorise the information in certain groups. Another type of task according to Willis' (1996) categorisation is the comparing task, here the learner either compares or contrasts the already acquired knowledge to new information. During problem-solving tasks, the learner should be reasoning or making decisions in order to solve a problem, whereas during personal experience tasks the student shares own opinions or experiences with his/her classmates. The last type of task, the creative task, can involve larger projects or creative ways to employ the already learned and practised language or language features (cf. Willis 1996; Müller-Hartmann & Schocker-von Ditfurth 2011: 90; Breyer 2000: 48-51).

On the other hand, a broader approach of differentiation between information-gap, opinion-gap and reasoning-gap tasks can be made (Ellis & Shintani 2014: 137). What is important when regarding information, opinion or reasoning-gap tasks, one needs to consider that those tasks are all meaning-focused activities (cf. Breyer 2000: 9). Initially, Prabhu (1987) proposed this triadic distinction, where during the information-gap activities the student transfers given information, during the reasoning-gap task new information is obtained through analysing given information, and during opinion-gap activities the student includes personal and emotional attitudes (cf. Breyer 2000: 9). Another important distinction is the difference between focused and unfocused tasks. Focused tasks refer to tasks which use specific language that might be provided prior to actual task performance, whereas during unfocused tasks learners can use their own knowledge of the L2 freely (Ellis & Shintani 2014: 138). Regarding the use of complex language one might argue that a focused task provides opportunities to practice and automatise complex structures which, ideally, are then used during unfocused tasks to complexify produced output. Foster and Skehan (2009: 276) emphasise that tasks provide the perfect means for a beneficial IL development, as "tasks engage the very processes that lead to acquisition". Negotiation of meaning,

which is assumed to occur more frequently during agreement tasks, may support those acquisition processes (cf. Foster & Skehan 2009: 276). In summary of the above description of different task types as well as their implications for the language classroom, a task is:

> "an activity in which meaning is primary, there is some sort of relationship to the real world, task completion has some priority, and the assessment of task performance is in terms of task outcome." (Skehan 2009b: 83).

According to this description, one might assume that a task is one *single* item during the learning process, or in particular, during one lesson. This assumption might apply to the actual performance during the task. Pre- and post-task activities, however, show that a task entails more than one activity (Skehan 2009b: 99ff.). In order to successfully complete the task and to have a beneficial effect on the interlanguage system, it seems to be obvious that during pre-tasks it is helpful to provide students with sufficient language skills, as well as to reduce task complexity to promote complex language production (Skehan 2009b: 99). With regard to post-task activities, it has been argued that the knowledge of what will be asked after task completion can have a great influence on how students approach the task (Skehan 2009b: 101). Therefore, some post-task activities may precipitate the chance that the learner's main focus lies on the meaning and not on the production of accurate and/or complex language during the actual task (Skehan 2009b: 101). Skehan and Foster (1997: 199f.), however, show that post-task activities have hardly any influence on fluency and complexity of language production and that the impact on accuracy was also rather insignificant, though slightly higher than on complexity and fluency. Generally speaking, one could argue that post-task activities have little influence on actual task performance (cf. Skehan and Foster 1997: 202). In order to confirm the different hypotheses regarding post-task activities more research into that field would be appreciated (cf. Skehan and Foster 1997). Due to the limited scope of this research and due to the fact that recent studies did not show an enormous influence of post-task activities on task performance, post-task activities are not in the focus of this study. Considering these theories and hypotheses, it is implied that the main focus of language teachers should lie on the implementation and use of *appropriate* tasks, i.e. neither too challenging nor to simple ones, given that both might hinder the development of learners' interlanguage system and the production of complex language. The simplification of the cognitive load of the task can be achieved

through activation of prior knowledge or through reduction of communicative pressure and stress (Skehan 2009b: 99f.). Regarding task sequencing, Robinson (2007b: 210) proposes that a task sequence should start with rather simple tasks, involving more and more complex tasks, providing the learner with a wide range of tasks with different complexity levels. A similar tendency can be found in Willis' (1996) task sequence.

According to Skehan (2009b: 98) some of the main factors that might produce communicative pressure or stress during task performance are time pressure, task modality, stakes and control. It is argued that active tasks tend to create a more stressful environment than passive tasks, for example, active production – speaking – creates more pressure than passive reception – reading (ibid.). Regarding stakes, it seems to be obvious that the learner is more pressured if task performance is assessed or unsuccessful performance entails negative consequences for the learner than if unsuccessful task completion has no impact at all. Control, however, can also be used in order to simplify or complexify a task, for example greater influence of the student on task design or performance is perceived as simpler than a strict design where no change is possible (Skehan 2009b: 98). In order to assess task performance, the three domains complexity, accuracy and fluency represent not only the "three dimensions of production" (Ellis & Shintani 2014: 148), but also a means of measurement for task performance (ibid.).

5.3   Task complexity

Regarding task complexity, the Comprehensible Input Hypothesis and the Comprehensible Output Hypothesis propose that a task should be slightly challenging in order to improve the interlanguage system and to keep the learner attentive during the learning process. Therefore, it seems to be clear that task complexity plays a rather important role in the language classroom and for the development of the interlanguage system. A common differentiation between simple and complex tasks is described by means of how conceptually demanding a task is (Robinson, Cadiero & Shirai 2009: 534). Hence, the so-called Here-and-Now tasks, typically using the present tense and offering the student to use personal information, are seen to be simpler than the more complex There-and-Then tasks, i.e. usage of past tense and/or fictional information (ibid.). According to Robinson (2001: 31) it is important to differentiate task complexity

from task difficulty. Task complexity appears to be a more objective and broader way of describing the cognitive demands of a task, whereas task difficulty differs from student to student, thus, is influenced through internal, i.e. individual factors. Additionally, external factors, such as motivation, can alter the learner's perception of task difficulty even though the complexity of the task remains unchanged (cf. Robinson 2001:31). Similarly, Schleppegrell (2004: 15) mentions those impacts on the perception of task complexity in the sense of difficulty, by including learner context and individual prior knowledge, aspects which can impact the subjectively perceived notion of task difficulty. In other words, task complexity comprises task difficulty, i.e. linguistic and cognitive demands of the task, as well as more individual and situational factors, i.e. communicative stress and individual differences (Robinson 2001: 28; Robinson 2007a: 12). In further detail, individual differences such as working memory or reasoning capacities can influence the perception of task complexity.

Table 1: The Triadic Componential Framework for task classification – categories, criteria, analytic procedures, and design characteristics (Robinson & Gilabert 2007: 164)

Task Complexity (Cognitive factors)	Task Condition (Interactive factors)	Task Difficulty (Learner factors)
(Classification criteria: cognitive demands) (Classification procedure: information-theoretic analyses)	(Classification criteria: interactional demands) (Classification procedure: behavior-descriptive analyses)	(Classification criteria: ability requirements) (Classification procedure: ability assessment analyses)
(a) Resource-directing variables making cognitive/conceptual demands	*(a) Participation variables* making interactional demands	*(a) Ability variables* and task-relevant resource differentials
+/− here and now +/− few elements −/+ spatial reasoning −/+ causal reasoning −/+ intentional reasoning −/+ perspective-taking	+/− open solution +/− one-way flow +/− convergent solution +/− few participants +/− few contributions needed +/− negotiation not needed	h/l working memory h/l reasoning h/l task-switching h/l aptitude h/l field independence h/l mind/intention-reading
(b) Resource-dispersing variables making performative/procedural demands	*(b) Participant variables* making interactant demands	*(b) Affective variables* and task-relevant state-trait differentials
+/− planning time +/− single task +/− task structure +/− few steps +/− independency of steps +/− prior knowledge	+/− same proficiency +/− same gender +/− familiar +/− shared content knowledge +/− equal status and role +/− shared cultural knowledge	h/l opennes to experience h/l control of emotion h/l task motivation h/l processing anxiety h/l willingness to communicate h/l self-efficacy

The Triadic Componential Framework (Robinson 2007a: 15f.) provides categories for task classification and criteria to categorize tasks. Task qualities or task features, i.e. cognitive factors like resource-directing or resource-depleting influences may be employed to vary task difficulty, i.e. to enhance or weaken the complexity of the task (cf. Robinson 2001: 29f.; Robinson et al 2014; Robinson, Cadiero & Shirai 2009: 536). Robinson (2007a) differs between three main categories: (1) task complexity, (2) task condition and (3) task difficulty. The first, task complexity, refers to cognitive factors, i.e. factors which effect the learner's attention, memory, reasoning et cetera (Robinson 2007a: 17), and can be divided into two subcategories, resource-directing and resource-dispersing variables. The resource-directing category involves variables that make cognitive demands and which are used to express and understand

52

the task (Robinson 2007a: 17). Those variables should direct the learner's attention towards certain features of the L2, thus supporting and enabling the noticing process (ibid.). Resource-dispersing variables, in contrast, make performative/procedural demands without guidance towards certain features of the L2 (Robinson 2007a: 18). Robinson (2007a: 18) states that those resource-dispersing variables require the learner to access automatised chunks and features of the L2, i.e. the "already established interlanguage system" (ibid.). The second main category focuses on interactional criteria and task conditions and its subcategories refer to (a) participation variables and (b) participant variables. The first regards interactive demands of the task and the nature of the task, whereas the latter refers to differences of the individual learners and their effect on task performance (Robinson 2007a: 14). The last main category, task difficulty and learner factors, is subclassified into learner abilities and affective variables (Robinson 2007a: 19). Those variables are mainly concerned with individual learner differences and learner centred criteria, thus showing that task difficulty can be seen as a more subjective category, depending on the individual participant. The complete list of variables can be seen in Table 1.

6 Purpose of research and Hypotheses

While examining learners' complexity of spoken language, this research will answer some significant questions concerning complex language production. The main purpose of research is depicted by the question of how learners achieve complexity in spoken language and which aspects of task complexity and classroom discourse can influence this complex language production. Regarding the importance of tasks and activities, the main focus lies on learner language during task performance and the impact of tasks on output production. The influence of task complexity as well as task condition on language production will be analysed. Additionally, the correlation between accuracy, fluency and complexity will be observed and compared between the different task performances. Alongside recent research in this field of study, some hypotheses about language complexity in second language classroom discourse will be taken into consideration.

6.1 Hypothesis 1

Different tasks and activity types influence complexity in learner talk. In particular, more complex tasks engender more complex language production, whereas simple tasks are more likely to evoke more simple language.

6.2 Hypothesis 2

Planning time has a positive impact on the learner's complexity and fluency. Pre-task activities provide the student with more time to think about the content and to formulate more complex ideas. Accuracy, however, will not be significantly affected by planning time.

6.3 Hypothesis 3

A trade-off effect appears when task complexity is too high. This hypothesis implies that during a task that is cognitively too challenging, accuracy and complexity will

suffer, or only one domain will show an increase at the expense of the other domain and *vice versa*. However, an adequately challenging task will engender more complex as well as more accurate language. In other words, a co-occurrence of complexity and accuracy is possible, when performing a task which is not too demanding and not too simple.

6.4   Hypothesis 4

Interactive tasks which promote negotiation of meaning can be beneficial for complexity. More time and more opportunities for speech production and modification cause more complex language and provide the student with more time to edit previously uttered phrases, therefore, offering means of complexification.

7   Methods and Design of Study

7.1   Participants, design and procedure

The participants of this study visit a business school in Vienna and attend a language class with a level of B1+, i.e. an intermediate course. The class consists of 22 students, 6 male and 16 female students who agreed to participate in this study; however, they did not receive detailed information on the subject of the research. Even though the general level at this stage is B1, some students excelled this level whereas others showed slightly weaker performances, rendering some tasks rather difficult as language levels varied. The students have already been in school for twelve years, thus a general exposure to the English classroom of ten years can be assumed. Given that some students come from other countries and have different language backgrounds no clear statement can be made. However, a minimum of six to seven years of English lessons can be assumed. In this particular school and in this grade, the number of English lessons is two hours per week.

Six lessons in the course of three weeks were audio recorded. In particular, these six lessons had a special focus on oral production and interaction in the classroom and on decision making and reaching compromises. In lesson 1, the students were confronted with the topic "Making Decisions" for the first time, the main objective and the main focus of this lesson lay on familiarisation with the topic, acquisition of new vocabulary as well as including personal experiences with decision making. Lesson 2 focused on past decisions and grammatical structures, especially how to form and use the conditional type two. The main interest of lesson 3 was on narrative task production, in particular a mini-presentation about the problems in the life of teenagers. In lesson 4, the students read a text about problems, revised and practised the conditional, followed by a listening activity about having arguments. For the following lesson, lesson 5, the students had to prepare a role for a role play and as a pre-task activity useful vocabulary and phrases were revised and learned. In the last lesson that was recorded, the main focus lay on having discussions and arguments and, finally, the students were asked to debate certain topics, which will be listed in detail below. The teacher explicitly asked the students to have *fiery arguments.*

A selection of five tasks, ranging from simple to more complex tasks, including narrative as well as dialogic tasks was made, transcribed and the measurements for complexity, accuracy and fluency were applied. The selected tasks represent the most typically used tasks in the language classroom and provided the students with different task conditions and task demands, thus, representing a varied range of task types.

Tasks and activities where students were grouped in pairs or groups of three were recorded in order to compare task performance during different tasks and by different students. The whole-class interaction was recorded during those six lessons and four additional recording devices were used to record the individual groups during task performance. Those recording devices where used during the individual tasks and the groups were chosen randomly during the different lessons. Thus, there are several group recordings for the individual tasks which provide more varied data. The recording of the tasks and lessons did not interfere with normal classroom routine and the tasks were not recorded additionally to normal classroom interaction but during normal lesson procedure.

7.2   Tasks and activities

From the six lessons, five tasks were chosen for transcription and detailed analysis. The chosen tasks represent a variety of common task types in the English classroom and where chosen according to task demands and task condition criteria which render the task more complex. Detailed information about those criteria can be found in Table 2 at the end of this subchapter. In order to categorize the tasks into simple and complex tasks, the *Triadic Componential Framework for task classification* (cf. Robinson 2007a; Robinson & Gilabert 2007; Robinson, Cadierno & Shirai 2009) was applied (cf. Chapter 5.3 Task complexity). This classifying system distinguishes between task complexity and task condition, i.e. cognitive and interactive factors of the task and task difficulty, i.e. individual learner differences. Given that no information about individual differences of the participants was provided, task difficulty will not be included into the analysis.

Per task, three to four recordings were done in either groups of three or dyads. The tasks and activities done in class range from simple tasks to more complex ones,

with or without planning time, including interactive as well as monologic tasks as well as one role play. Task prompts are taken from the students' course book *English Unlimited 3, Coursebook HAK/HUM B1+*.

7.2.1 The Simple Task

The Simple Task is a simple jump-in task where the students are provided with two to three minutes planning time. It is an interactive task where the students deal with personal and more general information about their behaviour and attitudes towards decision making. The students were asked to discuss the following questions in pairs:

(1) when and how to revise before exams?
(2) where to spend your holidays?
(3) what to buy, when and where? and
(4) what to spend your pocket money on? (English Unlimited 3 Coursebook 2011: 46)

Given the few task elements, i.e. discuss about making decisions, no demand of spatial, intentional and perspective reasoning, resource-directing variables of this task facilitate the cognitive demands of the task. Similarly, resource-dispersing variables do not add to the cognitive load, therefore, making the task simple. Before actual task performance, students were provided with two minutes planning time, providing them with enough time to read through the questions and think about some answers.

7.2.2 The Monologue

The Monologue is a narrative task where no interaction is requested. The learner prepares a mini-presentation about his or her problems in life. Planning time consists of five minutes written preparation for a one minute monologue, however, the students were asked to write down notes but no sentences in order to prevent them from reading their presentation instead of speaking freely. During task performance students were reminded that no questions and interruption is requested as questions will be asked and answered in a follow-up task. An example presentation of what these mini-presentations should look like was presented by the teacher at the beginning of the lesson.

Regarding the cognitive factors of task complexity, it is a rather simple task as neither resource-directing nor resource-dispersing factors add additional load to the cognitive demands. Considering that students were asked to hold a monologue, interactive factors do not complexify task demands. Generally, students hold presentations in front of the class, however, during this activity the presentation was held in a dyadic situation which rendered the situation more unfamiliar to the students. This uncommon situation might contribute to a more unfamiliar, hence complex context.

7.2.3 The Role Play

In the Role Play students were asked to prepare their roles at home, additionally, useful vocabulary was practised as a sort of pre-task. The planning time of two to five minutes was directed towards using the previously learned phrases for reaching a compromise. Additionally, the objective of the task was to reach an agreement and to negotiate different opinions. Regarding task complexity, resource-dispersing variables, such as inclusion of more elements, spatial, causal and intentional reasoning and change of perspective, and resource-dispersing variables, like +/- single task, inclusion of several steps as well as no immediate prior knowledge about the task, mostly rendered the task more complex, thus the role play is considered a complex task, including more than two participants. Interestingly, it is also the only task where the students had to present not their own view but another person's perspective.

The prompt for the role play reads as follows:

Work in A/B/C groups. You're a family, and you need to discuss some problems. Student A, you're the oldest sibling, look on p. 171; Student B, you're the middle one, look on p. 174; Student C, you're the youngest sibling, look on p. 181.

Example: A: *How about if we do the cooking together?*

B: *But that would mean coming home earlier.* (English Unlimited 3 Coursebook 2011: 51)

7.2.4   The Complex Tasks

The first Complex Task is actually a post-reading task and is done in a dyad. Before the actual task performance the students read a dialogue using if-sentences and were asked to discuss what they would do if they were in the position of Ann and Mike, the couple in the dialogue who talked about a problem in their relationship. Further, they were asked to give their opinion on partnerships. The reading activity can be found in the appendix (Picture 1). Regarding task classification criteria, the There-and-Then condition, perspective taking and reasoning, the task classifies as complex as both resource-dispersing and resource-directing factors increase task complexity. The first Complex Task provides the students with prior knowledge and a reading of a dialogue immediately before the discussion offers information about the content. However, no actual planning time was provided.

The second Complex Task is a speaking activity prompted by the book *English Unlimited 3* where the learners, in a group of three, were asked to discuss and argue about how they behave in (1) a conflict, (2) resolve an argument, (3) what their opinion about public anger display is and (4) cultural differences (cf. English Unlimited 3 Coursebook 2011: 53).

The prompt in the book reads as follows:

Discuss your ideas in groups.

1 What do you do when you're angry with:
□ a classmate?   □ a relative?   □ a friend?   □ a teacher?

2 How do you think conflict should be resolved at school or at home? Give an example from your experience.

3 How do you feel about getting angry in public? Is it acceptable in Austria?

4 Is it the same or different in other cultures you know of?
(English Unlimited 3 Coursebook 2011: 53)

Resource-dispersing and resource-directing variables render the task as complex as the first Complex Task. The second Complex Tasks includes no planning time and the students have to start with the discussion without any preparation.

Table 2 presents all four task types and their complexity level according to Robinson's (2007a: 15f.) *Triadic Componential Framework*. The criteria rendering the

task more complex can be found in Table 2. The complete framework can be found in Table 1, in Chapter 5.3 Task complexity. According to this categorisation, the tasks were named, selected and analysed. Given that planning time is of importance to this study, its absence or presence will be indicated for all four task types. The presence of planning time facilitates the tasks whereas its absence renders it more complex.

Table 2: Task complexity of the recorded tasks

Task type	Criteria for task complexity Cognitive factors	Criteria for task complexity Interactive factors
The Simple Task	+ causal reasoning + *planning time*	- one-way flow
The Monologue	+ causal reasoning - task structure + *planning time*	none
The Role Play	- few elements + spatial reasoning + causal reasoning + intentional reasoning + perspective taking - single task - few steps - no immediate prior knowledge + *planning time*	- open solution - one-way flow - few participants - few contributions needed - negotiation not needed
The Complex Tasks	- here-and-now - few elements + causal reasoning + intentional reasoning + perspective-taking - single task - few steps - independency of steps - *planning time*	- one-way flow - convergent solution - few participants - few contributions needed - negotiation not needed

Table 2 shows that the Role Play is also considered as a more complex task, however, here the main difference between the Role Play and the Complex Tasks is the aspect of planning time. The Role Play provides the students with sufficient and guided planning time, whereas the Complex Tasks do not include planning time.

7.3   Data analysis

7.3.1   *Dependent and independent variables*

Independent variables such as planning time, task complexity, and interactivity of the task play an important role during the analysis and show how complex language production might be influenced through external factors. A complete list of those factors is presented in Table 1, and Table 2 lists the relevant factors for the individual tasks.

The dependent variables, accuracy, fluency and complexity, are measured in several ways that will be explained in detail in the following section. This study will consider all three domains of performance; however, a special focus will lie on complexity scores (cf. Skehan & Foster 2007: 201). Additionally, the analysis will examine how the dependent variables interact with each other in different tasks and if a trade-off effect happens and or will be influenced through task complexity during task performance.

7.3.2   *Measures*

Accuracy

Accuracy is defined as the absence of errors. Therefore, the number of errors per clause will be the indicator for accuracy during task performance. However, elliptical single-word answers will not be included as those might distort actual accuracy scores.

Fluency

Fluency will be measured by mid-clause pauses, given the fact that end-clause pauses are native like and appear in everyday speech in order to add emphasis or a short break. Repair fluency will be taken into consideration if the student has to rephrase the

utterance completely in order to convey the meaning correctly, i.e. false starts. However, self-initiated repair can be considered as a part of spontaneous speech production and is not included in this analysis when the general structure of the utterance remains the same. Additionally, not only pauses that result in complete silence but also common fillers like "aah, mmm" et cetera will be considered as speech pauses.

Complexity S

Complexity S is measured by the mean number of subordinate clauses. Assuming that coordination is more likely to be an indicator for beginning learners of a second language, the main focus of this analysis lies on subordination without the inclusion of coordinated clauses.

Complexity L

Complexity L, also called Syntactic Complexity, is measured through mean length of clause which has been argued to present a more adequate and "predictive" means of measurement for more advanced learners (cf. Norris & Ortega 2009: 564, 574). Norris and Ortega (2009) argued that more means of complexity measurements may present a more precise picture of how complexity can be influenced through task complexity and other performance domains.

8 Findings and Results

The results section falls into three main parts: the first part provides a detailed description of quantitative analysis of the data-set, the second section is examining the scores of the performance domains with a series of ANOVAs and T-tests across tasks. The last part focuses on the interdependent relationship between the dependent variables accuracy, complexity and fluency in the individual tasks.

8.1 Description of the data-set

Table 3 gives an overview of the groups recorded per task. Due to unintelligibility, only three groups are analysed of the Simple Task and only two groups of the first Complex Task. In a qualitative analysis of the first Complex Task, one can observe that there is a significant difference between the two groups which is based on the fact that group two did not show much willingness to participate and transformed the speaking activity into a writing activity. Therefore, the results of this group will not be considered in the analysis. Group 1 of the first complex task will be regarded as group 5 of the Complex Tasks in the analysis. Table 3 shows the length of the different tasks by the number of clauses per task.

	Simple Task	Monologue	Role Play	The Complex Tasks	
Group 1	102	69	126	105	96
Group 2	140	42	101	67	162
Group 3	106	82	99		191
Group 4		116	78		190
Total	*348*	*309*	*404*	*172*	*639*
used in the analysis	**348**	**309**	**404**	**744**	

Table 3: clauses per task

The following section will present the quantitative analysis of the data per task. A presentation of the absolute numbers will provide an insight into the data-set, whereas the numbers in relation to task length will build the basis for the quantitative analysis that is presented in the following subchapters.

The Simple Task

	Fluency	Accuracy	Complexity S	Complexity L	words per task	clauses per task
Group 1	17	11	34	5,63	574	102
Group 2	28	10	53	5,55	777	140
Group 3	12	10	47	4,90	519	106

The Monologue

	Fluency	Accuracy	Complexity S	Complexity L	words per task	clauses per task
Group 1	31	10	27	5,55	383	69
Group 2	5	2	16	5,45	229	42
Group 3	37	6	35	5,96	489	82
Group 4	42	22	46	5,83	676	116

The Role Play

	Fluency	Accuracy	Complexity S	Complexity L	words per task	clauses per task
Group 1	30	12	52	4,75	598	126
Group 2	28	11	39	5,63	569	101
Group 3	23	7	31	5,45	540	99
Group 4	22	8	38	5,54	432	78

The Complex Tasks

	Fluency	Accuracy	Complexity S	Complexity L	words per task	clauses per task
Group 1	27	15	40	5,47	525	96
Group 2	38	8	63	5,41	876	162
Group 3	33	20	87	5,17	987	191
Group 4	23	16	71	5,16	981	190
Group 5	23	3	61	5,08	533	105

Comparison of mean values and in relation to task length

	Simple Task	Monologue	Role Play	Complex Tasks
Fluency	16,00	34,54	25,74	23,49
Accuracy	9,12	11,38	9,44	8,46
Complexity S	38,51	39,89	39,98	44,31
Complexity L	5,36	5,66	5,28	5,26

Table 4: Quantitative analysis of the complete data-set

8.2 The individual performance domains compared across tasks

Regarding the performance areas accuracy, fluency and complexity, the results in Table 5 show that there were highly significant scores in the domain of complexity S (sig. ,000) and complexity L (sig. ,000) and that there is a tendency for significance considering the accuracy scores (sig. ,001). However, fluency scores do not show significant results (sig. ,007). When analysing the different domains in more detail (cf. Table 4), those results were confirmed and will be discussed below. The ANOVAs also present more significant outcomes for accuracy and complexity, whereas fluency results remain mostly insignificant. Relevant tables of the ANOVAs will be presented in the following section, a complete set of tables can be found in the Appendix (Table 13-23). A detailed description regarding the individual domains will be discussed in the following sections.

Table 5: One-Sample Test – all dependent variables

	Test Value = .005				95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Accuracy	15,276	3	,001	9,59568203925E0	7,5966655691E0	1,1594698509E1
Fluency	6,532	3	,007	2,49379359600E1	1,2788201318E1	3,7087670602E1
Complexity S	32,297	3	,000	4,06683655550E1	3,6660983049E1	4,4675748061E1
Complexity L	58,863	3	,000	5,38388	5,0928	5,6750

8.2.1 Fluency

In Table 5 one can observe that the significance of the fluency scores is rather low (sig. ,007). Linear regression ANOVA tests as well as a detailed analysis of Table 4 showed that in the domain of fluency, the Monologue presented the most significant outcomes. The Monologue presented the most occurrences of pauses and false starts with a mean value of 34.54 (cf. Table 4), thus was significantly more dysfluent than the other tasks. The comparison of the mean values in Table 4 illustrates that during task performance of the Simple Task, learner speech was at its highest fluency with only 16 pauses and false starts, followed by the Complex Tasks with only 23.49 pauses/false starts. During the Role Play, similar to the Monologue, the learners tended to pause

more often mid-clause and/or initiated repair through false starts (25.74 pauses/false starts). Hence, during the Simple Task learners tended to speak more fluently than during the Monologue and the more complex tasks, i.e. the Role Play and the Complex Tasks. However, fluency scores during the Complex Tasks were still at a medium level, though not as high as during the Simple Task. An illustration of the results can be found in Diagram 1.
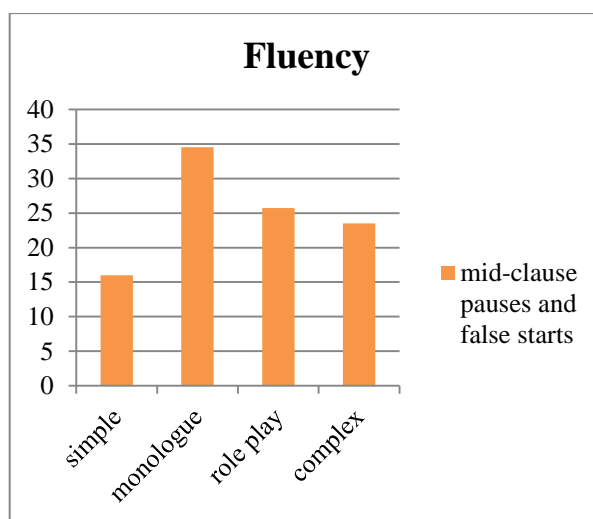

Diagram 1: Fluency

8.2.2 Accuracy

Turning to accuracy, in Table 5 an indication for a tendency towards significant results (sig. ,001) is presented. Regarding the different tasks, even though the Monologue, similar to fluency, presented the most significant results and highest number with 11.38 errors per clause (Table 4), the learner language became more accurate when task complexity increased (cf. Diagram 2). The scores of the Monologue can be interpreted in so far that the most inaccurate speech was performed during the Monologue, whereas the other tasks, whether Simple or Complex, stimulated the students to produce more accurate speech. Table 4 shows that during task performance of the Complex Tasks, the lowest number of errors (8.46 errors per clause) occurred followed by the Simple Task with 9.12 errors per clause. The Role Play presented the second highest number of errors per clause (9.44), thus remains at a medium level. Diagram 2 illustrates the presented results.
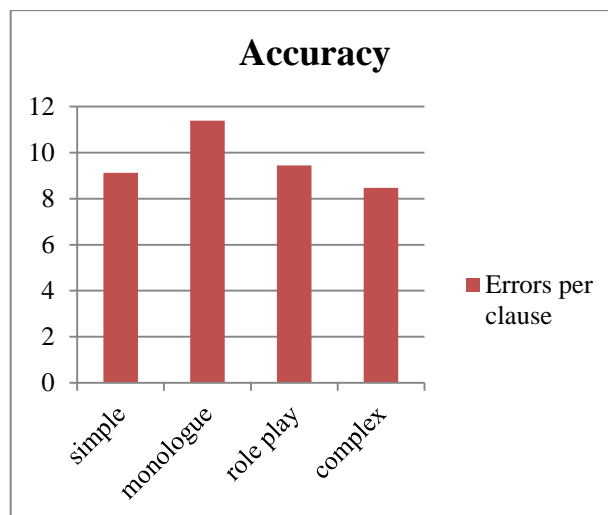
Diagram 2: Accuracy

8.2.3 *Complexity S and Complexity L*

As Table 5 indicates, both complexity S and complexity L scores showed the most significance (sig. ,000). Table 4 shows that complexity L measures, analysed by mean length of clause, were significantly higher during task performance of the Monologue (5.65) as during the Simple (5.36) and the Complex Tasks (5.26) or the Role Play (5.28). The scale of Complexity L scores is thus rather narrow, ranging from 5.65 words per clause in the Monologue to 5.26 words per clause in the Complex Tasks (see Table 4). An explanation for this outcome might be that the learners tended to produce longer clauses when no fear of interruption endangered their speech and when the threat of losing their turn to another interlocutor was not present. These results, however, propose that students show a tendency of producing shorter clauses in an interactional task than during narrative task performance. This finding is parallel to Michel's, Kuiken's and Vedder's (2007) study where similar results were obtained. Additionally, during Simple Task performance the learners' utterances were slightly longer than during Complex Task performance.

In contrast, Table 4 shows that complexity S measured by mean number of subordinate clauses was rather low during the Monologue with 39.89 subordinate clauses per clauses. Simple Task performance showed the lowest number of subordinate clauses with a mean number of 38.51. During the Role Play, a mean value of subordinate clauses of 39.98 was analysed, which represents a mediocre level. The

Complex Tasks showed the highest number of subordination with a mean value of subordinate clauses of 44.31. Hence, the mean number of subordination increased along task complexity, rendering learner speech more complex. A comparison of both complexity domains is presented in Diagrams 3 and 4.
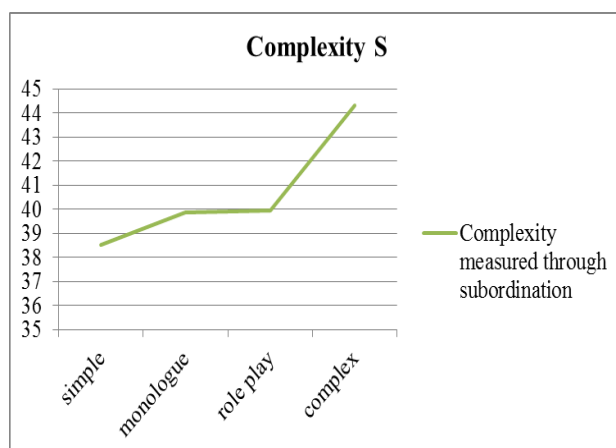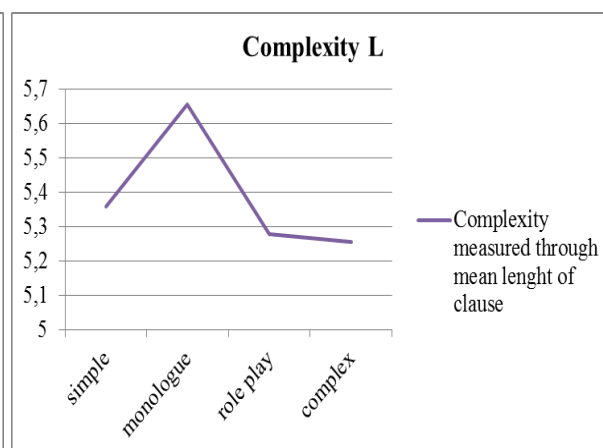

Diagram 3: Complexity S

Diagram 4: Complexity L

Table 6: Paired Samples Test – Complexity S and task complexity

	t	df	Sig. (2-tailed)
Pair 1 Complexity S - task_complexity	51,754	3	,000

Paired Differences				
			95% Confidence Interval of the Difference	
Mean	Std. Deviation	Std. Error Mean	Lower	Upper
3,81733655550E1	1,47517547685E0	7,37587738426E-1	3,58260321820E1	4,05206989280E1

Table 4 as well as Diagram 3 show that complexity S scores measured by subordination showed a considerable rise with increasing task complexity. The mean number of subordination in the Simple Task is 38.51, whereas with increasing complexity the mean number of subordination rose to 44.31 in the Complex Tasks. In Table 6, a T-test with paired samples of the variables complexity and task complexity presented significant results (sig. ,000). In comparison, the paired sample T-Tests of the other tasks and task complexity, which can be found in the appendix and which mostly show rather insignificant results with sig. ranging from ,007 when comparing accuracy

and task complexity to ,025 when comparing complexity L and task complexity. It can thus be argued that complexity and task complexity are more likely to act as interdependent variables in classroom discourse. An ANOVA test (Table 7) with the independent variables Complex and Simple Task indicated a high significance (sig. ,014) in relation with the dependent variable complexity S.

Table 7: ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	19,024	2	9,512	2470,429	,014[a]
	Residual	,004	1	,004		
	Total	19,027	3			

a. Predictors: (Constant), complex, simple

b. Dependent Variable: Complexity S

Similarly, as is presented in Table 8, a coefficient test showed even higher significant results for the dependent variable complexity S in the Complex and the Simple Tasks with a significance of ,003. In other words, there was significant difference between Complexity S scores during task performance of the Simple Task and Complex Tasks. Considering the rise in scores (cf. Table 4; Diagram 3), it can be argued that the Complex Tasks "pushed" (cf. Swain 1985) the students to produce more complex language than the Simple Task. Besides the fact that the most significant results are presented during task performance of the Complex Task or in comparison of the Simple and the Complex Tasks, the results of the Role Play and the Monologue were also marginally significant. In other words, complexity scores of the Monologue or the Role Play were not as high as during task performance of the Complex Task, tough higher than during task performance of the Simple Task (cf. Table 4)

Table 8: Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	45,843	,219		208,966	,003
	simple	1,425	,076	,283	18,748	,034
	complex	-4,379	,076	-,869	-57,621	,011

a. Predictors: (Constant), complex, simple

8.2.4   *Planning time and the performance domains*

Regarding planning time and hypothesis 2 of this study, significant differences between the tasks that provided planning time, i.e. the Simple Task, the Monologue and the Role Play, and the tasks that did not provide planning time, i.e. the Complex Tasks, could be observed. A series of paired samples T-tests (cf. Appendix Table 14) supports this finding, showing that complexity S scores and planning time present highly significant results (sig. ,000), followed by complexity L scores in combination with planning time (sig. ,001) and accuracy scores and planning time (sig. ,002), whereas fluency scores where only show a marginal significance in combination with planning time (sig. ,009). An ANOVA (cf. Appendix Tables 19-23) test (sig. 0,036) and a coefficient test with sig. ,001 confirmed the significance of planning time on the variable complexity. However, regarding fluency and planning time an ANOVA test confirmed that the results were insignificant (sig. ,873). The complete set of tables can be found in the Appendix, Table 14 and Tables 19-23. A more detailed analysis of Table 4 confirmed the significant results of the paired T-Tests. In particular, against an assumed beneficial effect of planning time on complexity and fluency scores, a tendency for an opposing direction could be observed. Table 4 shows that the highest number of subordination and the fewest number of errors could be found in the Complex Task which did not provide any planning time. Only fluency scores benefited from planning time in the Simple Task. However, during the Monologue which provided the students with guided planning time the most dysfluent speech was produced, thus implying that planning time and fluency only present a marginally significant relationship. In Contrast, complexity L scores showed an improvement when planning time was provided, especially in the Monologue and decreased in the Complex Task where no planning time was provided. An interpretation of these results will be provided in the discussion in Chapter 9.

Concluding it can be argued that the Simple Task was the most fluent task, with medium accuracy and complexity L scores, but low complexity S scores. The Monologue represented the most dysfluent and inaccurate task, with medium complexity S scores and the highest complexity L scores. The Role play scored

mediocrely on all performance domains. The Complex Tasks elicited the most accurate language and the highest number of subordinate clauses (complexity S scores), fluency scores remained at a medium level, with the lowest complexity L (mean length of clause) scores.

8.3 The inter-relationship of the dependent variables

Taking into consideration various hypotheses about trade-off effects (cf. Skehan & Foster 2001; Foster & Skehan 2009) or concurrent stimulation of the performance domains (cf. Robinson 2001, 2003, 2007) and regarding hypothesis 3 of this study, the inter-relationship of the three variables accuracy, fluency and complexity was observed via several paired T-tests, some are illustrated in the following section (Table 9 to Table 12), the complete set can be found in the Appendix.

Table 9: Paired Samples Test – Complexity S and Fluency

	t	df	Sig. (2-tailed)
Pair 1 Complexity S - Fluency	4,040	3	,027

Paired Differences				
			95% Confidence Interval of the Difference	
Mean	Std. Deviation	Std. Error Mean	Lower	Upper
1,57304295950E1	7,78677787220E0	3,89338893610E0	3,33992836028E0	2,81209308297E1

Several paired T-tests showed that there were no significant results for an interdependent relationship between fluency and accuracy, fluency and both complexity measures, complexity S measured via subordination and complexity L measured via mean length of clause. Table 9 represents one example of such a paired T-test including the variables complexity S and fluency and indicates that no significance (sig. ,027) was observed, other T-tests showed similar results and can be found in the Appendix. Given that the least fluent task was at the same time the most inaccurate task (the Monologue), there seems to be no obvious trade-off effect between fluency and accuracy. Regarding complexity and fluency, during The Simple Task performance student spoke fluently,

while complexity S scores remained rather low. In contrast, during the Complex Task performance complexity S scores were at their highest, whereas fluency scores remained at an acceptable level. However, there is no highly significant relationship between the performance domains accuracy and fluency or complexity S and fluency. Comparing the paired T-tests (e.g. Table 9) with the single T-tests (Table 5), one can argue that fluency scores where insignificantly influenced through task complexity. Hence, the results can somehow confirm the above mentioned arguments that it is more likely that the student focuses on either form or meaning, i.e. on accuracy and complexity *or* on fluency. Nevertheless, a trade-off effect seems to be more likely to occur between complexity and fluency rather than between accuracy and fluency. Especially complexity L, measured through the mean length of clause and fluency showed a tendency for a trade-off effect (cf. Table 4). Particularly, during the Monologue the most occurrences of pauses and false starts were presented, simultaneously, clause length was at its highest (see Table 4). Therefore, a trade-off effect between fluency and complexity L seems to be legitimate.

Table 10: Paired Samples Test – Complexity L and Accuracy

	t	df	Sig. (2-tailed)
Pair 1 Complexity L - Accuracy	-7,769	3	,004

Paired Differences				
			95% Confidence Interval of the Difference	
Mean	Std. Deviation	Std. Error Mean	Lower	Upper
-4,21180535400E0	1,08426230349E0	5,42131151743E-1	-5,93710863484E0	-2,48650207316E0

Turning to complexity L, in Table 10 it can be observed that there is a tendency for significance (sig. ,004) when pairing accuracy and complexity L. In particular, accuracy scores defined by errors per clause, where concurrently lower with a rise in complexity L scores measured through mean length of clause (cf. Table 4). During the Monologue, for example, the students tended to produce more inaccurate utterances as well as longer clauses. Whereas during the Complex Tasks the learners spoke more accurately, however, due to the interactive aspect of the tasks, the clauses showed a tendency to be shorter. It therefore seems to be legitimate to argue *for* a trade-off effect between accuracy and complexity L along increasing task complexity.

As shown in Table 11, the paired T-test of the dependent variables complexity S and complexity L indicated highly significant results (sig. ,000). Strikingly, when examining the results in more detail (cf. Table 4; Diagram 3 and 4), one can observe that the different tasks and different task complexity have significantly different effects on complexity S and complexity L scores. Complexity L, for example benefited most from the Monologue which on the contrary only scored mediocrely on complexity S measured by subordination. Taking into consideration the different means of measures it seems legitimate to argue that when the mean number of subordinate clauses per clause increases, mean length of clause decreases. Hence, it could be argued for the possibility of some sort of trade-off effect between the different levels of language complexity.

Table 11: Paired Samples Test – Complexity S and Complexity L

		t	df	Sig. (2-tailed)
Pair 1	Complexity S – Complexity L	27,191	3	,000

Paired Differences				
			95% Confidence Interval of the Difference	
Mean	Std. Deviation	Std. Error Mean	Lower	Upper
3,52844888698E1	2,59532929760E0	1,29766464880E0	3,11547408027E1	3,94142369368E1

The most significant results, however, are presented in the paired T-test between accuracy and complexity S with a significance of ,000 (see Table 12). In other words, an interdependent relationship between the two form-based performance domains is highly probable and significant. Regarding errors per clause and subordination per clause, it can be observed that the students tended to produce fewer errors and more subordinate clauses when task complexity increased (cf. Table 4). During the Monologue, however, the learners produced the highest number of errors per clause while subordinate clauses remained at an acceptable level. Similarly, in comparison to the Simple Task, during the Role Play accuracy scores were lower with slightly higher complexity S scores.

Therefore, there is some tendency towards a trade-off effect between accuracy and complexity S when task complexity or difficulty is either too high, i.e. the Role

Play, or too low, i.e. the Simple Task. Additionally, narrative task performance, i.e. the Monologue can produce a trade-off effect, given that complexity L scores were at its highest, whereas accuracy and fluency scores were at their lowest (cf. Diagram 1, Diagram 2, Diagram 4, Table 4). However, during complex task performance both accuracy and complexity S scores showed the best results, thus the most beneficial effects for the learners. In other words, a trade-off effect between accuracy and complexity S cannot be completely confirmed, as it seems to be more likely that both performance domains profit simultaneously from more complex tasks. In contrast, a trade-off effect between accuracy and complexity L scores seems to be more probable.

Table 12: Paired Samples Test – Complexity S and Accuracy

	t	df	Sig. (2-tailed)
Pair 1 Complexity S - Accuracy	18,872	3	,000

Paired Differences				
			95% Confidence Interval of the Difference	
Mean	Std. Deviation	Std. Error Mean	Lower	Upper
3,10726835158E1	3,29303964385E0	1,64651982193E0	2,58327225919E1	3,63126444396E1

8.4   Individual differences and CAF

As the discussion of this chapter will show, individual differences might have the most impact on the development of task performance. According to Skehan (2009b: 279) "cutting edge language" as well as negotiation of meaning depend on the learner's willingness to use the target language and test own hypotheses. A safety-first approach, however, hinders the student to apply the target language during task performance and to restructure previously produced output (cf. Skehan & Foster 1997). Comparing Extract 1 and Extract 2, one can observe that some students are willing to make an effort and negotiate meaning, whereas others just tend to employ their first language (L1) without trying to restructure their output in the target language. What is of importance is that task complexity does not seem to stimulate the student to negotiate meaning, but that the more social aspect of individual differences has a greater influence on the willingness to negotiate. It seems to be the case that willingness to

negotiate is more likely to be influenced through the interlocutors of the conversation and nature of their interaction, i.e. if the partners can talk freely with each other or if the presence of one partner feels threatening to the speaker. The first example, taken from the Simple Task performance shows a dyad of students who try to find the right phrase in the target language collaboratively, thus providing a beneficial influence on the interlanguage system.

Extract 1: The Simple Task, Group 2

S2: yes I compare very very very products very much <1> products </1> and and I have a high <L1de> wie sagt man </L1de>

S1: <1> okay </1>, <L1de> sags auf Deutsch </L1de>

S2: <L1de> ich achte sehr auf das Preis-Leistungs-Verhältnis </L1de>

S1: <L1de> aso ahm ahm oh Gott, </L1de> ahm I take <L1de> na warte, </L1de>

S2:ha so <2> for me it's not </2>

S1:<2> I I I I make sure </2> that the prices ah the price and the ahm

S2: the quality of the good

S1: the quality of the good is very ahm comparable

Taking into consideration the notion of adequacy and appropriateness, it seems to be legitimate to argue that even though the levels of the performance domains did not increase in such situations, yet the communicative goal was achieved. Regarding Extract 1, taken from the Simple Task, one can observe that negotiation of meaning took place; however, complexity levels did not rise significantly. Considering the proficiency level of the students according to the CEFR, B1, it can be argued that a negotiation like this can benefit the language acquisition process. The students did not only return to their L1, but tried to articulate the phrase in the target language. Additionally, communicative competence descriptors demand the students to develop interactional skills and to negotiate meaning, for example, at the level B1, regarding the skill speaking interactively, a descriptor which enforces negotiation of meaning is "Ich kann Alltagssituationen sprachlich bewältigen […], auch wenn ich hin und wieder um Wiederholung oder Erklärung einzelner Wörter bitten muss" (European Language Portfolio 2005: 38). Even if language accuracy as well as fluency might have suffered, the students achieved a communicative goal and could perform successfully. Communicative situations which support the student in achieving task objectives even at the expense of CAF *in* the situation can still benefit a progress in the performance domains in the long run. Similar to the concept of priming (cf. McDonough 2006),

complexity, accuracy and fluency can be positively influenced, even though it they are not affected immediately in the communicative situation. Adequate and appropriate reactions in particular situations thus can serve not only a communicative purpose but also benefit complexity, accuracy and fluency. Including the question of adequacy and appropriateness, thus contributes to combine a progress in the performance domains while achieving scholastic standards of the Austrian curriculum as students need to be able to react appropriately in certain situations (cf. CEFR 2001). Pallotti (2009: 598) supports this view by stating that language proficiency levels include sociolinguistic and discursive aspects as well as syntactic complexification. Hence, appropriate reactions in interactive situations, like negotiation of meaning in the example, might be seen as an increase in language proficiency during task performance (cf. Pallotti 2009: 598).

Regarding another group of learners during task performance of the Role Play, which can be considered more complex than the Simple Task, one student (S1) showed no willingness to negotiate or restructure what he/she intended to say. Moreover, the partners (S2 and S3) did not support S1 in reformulating the phrase in the target language or in using another expression. During the entire task performance of this role play, S1 remained more or less quiet with some mixed English-German phrases. Most of the discussion took place between S2 and S3 who were more fluent in the target language. This example shows that in a typical language classroom, proficiency levels often present a wide gap between individual students which can have either positive or negative effects on task performance and willingness to negotiate or support other students while using the target language.

Extract 2: The Role Play, Group 3
S3: S1 what's you saying about this situation?
S1: mmm it would be <L1de> umständlich?</L1de>
S3: aahm
S2: ahm soo the next problem? we have a new TV .

Another important influence on willingness to negotiate might be represented by personal and impersonal prompts. There seemed to be a tendency for the students to show more enthusiasm during task performance when talking about their own experience or their own ideas than during the Role Play. The Role Play provided the students with fictional characters and fixed ideas which were not adaptable to the students' own principles and views. However, in the Complex Task, a discussion about

problem solving and own attitudes towards conflict solving, the students tended to take turns more enthusiastically, frequently spoke simultaneously and generally showed more interest in their interlocutors. The first example, Extract 3, is taken from the Role Play where the students decided on the distribution of different rooms. Generally, one might argue that the own room is of importance to the students and that they would fight for their own room. Extract 3, however, shows that the students are not really personally involved and just try to finish the task as quickly as possible and without much arguing. Another indicator of a lack of involvement is the "and now?" ending, which is even uttered in the L1, and shows that the students only tried to finish as fast as possible without becoming personally involved. In other groups, expressions like "but (.) we shouldn't (.) ah ma- make this decision ahm so complicated" (Role Play, Group 1) support this argument and show that some students were solely willing to discuss as much as necessary to finish the task rather quickly.

Extract 3, Role Play, Group 4

S2: yeah that's (a brilliant idea?) to swap the rooms but there will be a problem because I want the biggest room and the biggest room is from (.) ah eh from S3 and we must try to (.)=

S3: = convince

S2: ja to convince about it

S1: aahm how about if we swap our rooms only for maybe 3 months and then <1> just swap </1> again yes

S3: <1> swap it again </1>

S2: ja I think it it COULD be a good <2> idea </2>but we must ask S3 because she have the biggest room @

S3: <2> I suppose </2> @ I suppose that's a good idea to swap rooms.

S1: okay

S3: okay

S2: <L1de> und jetzt? </L1de>

In contrast, Extract 4, taken from the more personal discussion about their own experience and attitudes towards conflicts in the Complex Task shows that the students are eager to explain their own ways of reacting and their feelings in situations of conflict. Moreover, the group partners are interested in the feelings of their interlocutors and support them by presenting their own explanation for and experiences of such a situation.

Extract 4, Complex Task, Group 4

S2: jaa. ahm I'm never (.) angry

S1: ja I hate it when I have conflicts with my mum. I'm always

S2: ja ja I'm I'm always crying. I'm always crying

S3: but why crying?

S2: because she shouted like I <2> I'm the worst kid in the world </2>

S1: <2> I haven't done anything and she's screaming at me </2>

S3: aah it's not solution when you cry

S2: ja but she doesn't she doesn't see when I cry (.) I don't show it to her.

Additionally, individual speaking style has an enormous effect on complexity and fluency measures. On the one hand, students showed a tendency to pause more often and to stop mid-clause for thinking about a way to formulate a phrase. On the other hand, some learners may prefer to take time to plan what they intend to say before uttering a phrase, thus speaking more fluently according to fluency measures. Comparing now both speaking styles, however, the achievement of communicative goals as well as task performance can be seen similarly successful in both forms. Comparing Bygate's (2009: 264) hypothesis that narrative tasks enhance fluency levels to the results of this study, it can be observed that there is a tendency in the opposite direction, implying that during the Monologue the students spoke least fluently. In the example, Extract 5, below taken from the Monologue one can observe that this student speaks rather slowly including a lot of mid-clause pauses. In order to illustrate the number of pauses, the symbol for silenced pauses (.) and the fillers ("ah") are bold and in italics.

Extract 5, the Monologue, Group 4

S1: okay I can start so *(.)* two there are two things with *ah ah* them I'm not happy in my life. *ahm* the first problem is my time manag- time management *ah* I'm *ah* very bad at it? and I do always things in the last two seconds like to do homework or to learn for *ah* a test? and but in school it doesn't affect me in school <1> because </1> I have good marks but at home *ah* my parents are very unhappy with *ah* my *ah (.)* slow motion moves, you know, I'm very slow and I take a lo*ooooo*t of time in the bathroom to making my hair or something like that and that's that's *ah* what I want to change in my life to be better at *ah* to control my *ah* time and to manage it *ah*. my second problem is *ah(.)* sport? and *ah* because to combine sport school and religion? you know?

S2: <1> mhm aha </1> ja

Given that one group contrasted the general tendency and made very few pauses/false starts during the narrative task, the influence of task condition might also

be rather student dependent and no general statement can be made. One example of more fluent speech is taken from The Complex Task, group 1. Considering individual differences, this group stood out from the other groups as language proficiency, as well as individual factors like motivation and cooperation between the group members was significantly higher than in other groups that were recorded. Again, pauses are marked in bold and are in italics.

Extract 6, The Complex Task, Group 1

S1: ja because when you *ahm* like have a relationship you would look for someone who is like kinda similar to your *(.)* personality. like if someone is outgoing you don't want *ah* a couch potato.

S3: but also when two different types meet each other they could rub off on the other person. so then the couch potato is more of an active person and wants to do something which can be good too so. I think it's all about the balance.

S2: ja. ja. ja

S3: and what you do with the partnerships. because if you just accept how it is and one of *ahm* them has to do what the other one has-ah wants then it's bad but if you find a compromise how both of them get happy then everything's fine.

Comparing Extract 6 and Extract 5, one can see that in the Monologue, there were significantly more occurrences of pauses than in the interactive task, the Complex Task. Hence, fluency measures might show that during some tasks speakers tend to pause more often, however, sometimes speaking slowly or pausing mid-clause can also be seen as individual choice of style. In contrast, accuracy measures might not be as influenced through such stylistic choices since it does not seem to be likely that learners consciously decide to include errors. Robinson (2003: 662f.) stated that individual differences might exert a greater influence on explicit knowledge learning than on implicit learning. Working memory, intelligence and aptitude, as examples for individual differences, seem to be more likely to affect explicit language learning. However, considering the results of the study and the individual performances, it can be argued that individual differences such as willingness to take risks, intro and extroversion as well as openness for new things or new languages may have an impact on the learning of implicit knowledge as well. Particularly language complexity might be affected by the learner's willingness to restructure and complexify the output is highly dependent on risk taking.

9　Discussion

Comparing the results to the formulated hypotheses (see p. 50f.), it can be argued that some hypotheses were supported, whereas others did not apply completely in this particular study. The discussion falls into two main parts, in the following section previously made hypotheses will be discussed, followed by a comparison to the results of other recent studies.

9.1　Hypotheses 1 to 4

Hypothesis 1	confirmed
Hypothesis 2	cannot be confirmed in all cases
Hypothesis 3	confirmed
Hypothesis 4	can be partly confirmed

Hypothesis 1: Different tasks and activity types influence complexity in learner talk. In particular, more complex tasks engender more complex language production, whereas simple tasks are more likely to evoke more simple language.

Considering the results and students' performance during different tasks with different complexity levels, this hypothesis can be confirmed. Skehan's and Foster's (cf. 2009a, 2001, 1997) assumption that simple tasks provide the students with more time to focus on accuracy since the task is not cognitively challenging, did not prove right in this case. Robinson's (cf. 2001, 2003) cognition hypothesis, thus, can be confirmed to some extent in this study. However, it applies rather to Complexity S measured via subordinate clauses than Complexity L measured via mean length of clause. The results showed that Complexity L was more likely to benefit from the narrative task, i.e. the Monologue as from the Complex Tasks.

Hypothesis 2: Planning time has a positive impact on the learner's complexity and fluency. Pre-task activities provide the student with more time to think about the content and to formulate more complex ideas. Accuracy scores, however, will not be influenced by planning time.

Hypothesis 2 cannot be confirmed due to lack of evidence and the limited scope of this study. The most complex language was produced during tasks which did not provide planning time for the students, i.e. the Complex Tasks. Even if it seems legitimate that planning time provides the students with enough preparation, thus resulting in more fluent and more complex speech, it is very student as well as situation dependent if such planning time is really used appropriately. It might occur in some cases that the learner focuses solely on the meaning, i.e. the content, without thinking about the structure and/or how to articulate what he/she intends to say. Particularly complexity S did not benefit significantly from planning time but from higher task complexity. Similarly, fluency suffered most during the narrative tasks where the students were specifically asked to formulate their ideas, i.e. their problems in life, to take notes and to use specific vocabulary and phrases which were practised in a pre-task activity. However, most students tended to make more pauses during this task than during the Complex Tasks where no planning time was provided. It can thus be argued that spontaneous speech might provide the student with more opportunities to make the output more complex as well as more fluent. Given that fluency activities are argued to provide the students with more native-like language use (cf. Skehan and Foster 2009: 280), spontaneity and a lack of planning time might enhance a real-life-like aspect of the activity and pushes the learner to utter more complex as well as more fluent output (cf. Mackey 2012; Robinson 2003, 2001). Regarding accuracy scores, the hypothesis can be confirmed to some extent as planning time did not benefit accuracy. However, it might be argued that students did not focus on accurate speech, even though they were asked to focus on form during planning time of some tasks, for example the Role Play and the Monologue.

Hypothesis 3: A trade-off effect especially between accuracy and complexity appears when task complexity is too high; however, an adequately challenging task will engender more complex as well as more accurate language.

Regarding the results and findings, hypothesis 3 can only be partly confirmed. Due to a lack of tasks with extremely challenging task complexity, one can solely suppose that a tendency towards a trade-off effect exists when tasks are too challenging. Considering the analysed tasks, the Role Play presented itself as the most challenging though not the most complex task. The assumption that the Role Play might have been the most challenging task for the students is based on rather dysfluent speech, low accuracy and acceptable complexity S and complexity L scores. In contrast, Skehan's and Foster's (2009a, 2001, 1997) proposal that the less complex a task, the more accurate the output, does not find support in the present data. However, the assumption that an adequately challenging task stimulates the student to produce accurate and complex speech simultaneously proved right in this study. What is of importance is that the results of this study present the mean number of complexity and accuracy scores and can vary significantly between individual students. Considering individual differences of the students, some tasks might present themselves more complex to some students than to others, thus can easily become too challenging (or too simple) and hinder beneficial complexity and accuracy development.

Hypothesis 4: Interactive tasks which promote negotiation of meaning can be beneficial for complexity.

The results of this study can confirm hypothesis 4 to some extent. It shows in some cases that interaction can have beneficial effects on language complexity. In contrast to Mackey's (2012: 24) hypothesis that one-way tasks provide more possibilities for output production, regarding the length of the different tasks (cf. Table 2: clauses per task) the results of this study showed that generally more language (measured by the number of clauses) was produced during two-way tasks, i.e. interactive tasks than during the Monologue. Regarding Complexity L (measured by mean length of clause), the Monologue elicited longer clauses than interactive tasks.

However, more output production in general as well as a higher number of subordinate clauses was found in interactive tasks. So, the assumption that interactive tasks are more likely to elicit more complex language proved right in most cases (cf. Bygate 2009). Contradicting Bygate's (2009: 264) assumption, monologic tasks did *not* enhance fluent speech, but the highest number of mid-clause pauses and false starts was found during the Monologue. An explanation for this result could be that the students were neither properly prepared nor used to hold a presentation just in front of one person. However, comparing the individual groups, there was one group where the students spoke very fluently. The exact numbers can be found in section 8.1. Hence, it can be argued that for some students the Monologue was beneficial for fluency, however, no general conclusion can be drawn. Comparing Complexity S scores of the Monologue to the more interactive tasks, strikingly only the tasks with higher task complexity elicited more complex language. The Simple Task, however, could not stimulate the students to produce more complex language, thus the Monologue *can* be beneficial for language complexity when compared to more simple tasks. Additionally, regarding Complexity L, the Monologue stimulated the students to produce longer clauses than during interactive tasks. It can be argued that in interactive tasks clauses tend to be shorter due to turn taking, interruption through questions, confirmation checks et cetera and, finally, due to negotiation of meaning. A narrative task, which does not include such disruptive features, provides the learner with more time to utter longer clauses, thus to increase Complexity L. Concluding, one might argue that the interactive aspect of the task is not as influential than task complexity for the complexity of the produced output. What is important, however, is that interactive tasks elicited more output in general than the narrative task and prompted the students to modify and restructure their output thanks to negotiation of meaning.

On the whole, the hypotheses were supported, however, some areas, such as planning time and a trade-off effect due to overwhelming task complexity, require further research to be completely understood. It seems to be legitimate to argue that task complexity has a great impact on a speaker's complexity, accuracy and fluency level, and it should therefore be considered during task sequencing and classroom design. Interactive as well as monologic tasks elicit CAF and can benefit different areas of language proficiency. However, dialogic tasks elicit a larger amount of output and

provide the students with more opportunities to negotiate meaning (cf. Bygate 2009). Considering that most students who participated in the study are at an intermediate level, subordination turned out to be the most appropriate means of measurement for language complexity. What is interesting is that Complexity L scores showed a different tendency than complexity S scores. However, at this level it might be argued that it is very student dependent if phrasal elaboration or length of clause indicates a complexification of the output.

9.2 Comparison to recent studies

In order to compare the results to other studies, the statistics and findings section present mostly average scores of all groups during task performance. Regarding the individual students, however, various individual differences as well as different proficiency levels play an enormous role in language production and can influence the scores and results of all three performance domains.

9.2.1 *The impact of planning time and time pressure*

Recent studies supported the hypothesis that planning time might be beneficial for fluency, complexity and in some cases even accuracy (cf. R. Ellis 2009; Yuan & Ellis 2003; Ellis & Yuan 2004; Skehan 2009b among others). Considering cognitive factors, for example attention allocation or working memory capacities, it seems to be legitimate to assume that planning time provides the students with opportunities to prepare and structure output production during task performance, thus improving complexity, fluency and accuracy scores (cf. Schmidt 2001: 13f.). One criticism, which has already been mentioned, refers to the fact that even though planning time might be provided, one can never assure that this time is used for actual task preparation (cf. Schmidt 2001: 18). Gilabert (2007: 62) as well as Yuan and Ellis (2003: 7) argue that pre-task planning does not seem to improve accuracy scores, which can be supported by the results of this study, as the most accurate speech was produced when no planning time was provided. Although it seems to be legitimate that pre-task planning directs the student's attention to form, this study, as well as other studies (Gilabert 2007; Yuan and Ellis 2003 among others) did not show a significant influence of planning on accuracy scores. Ellis' and

Yuan's (2004) study showed that complexity and fluency scores improved when planning time was provided. Similarly, Gilabert (2007: 62) found that pre-task planning lead to higher fluency and a lexically enriched language production. Lexical variety, next to fluency, seems to be more likely to benefit from pre-task planning time by providing the students with enough time to choose appropriate terms and to vary the used vocabulary by "[mobilizing] less frequent words" (Skehan 2009b: 515). Regarding complexity, this study contrasts those results as more complex as well as more accurate language was produced during the complex task without planning time. The role play, which included previous preparation at home additionally to directed planning time in class, however, only resulted in an average score for complexity, accuracy as well as fluency. Therefore, planning time did not prove to be as influential on complexity scores as task complexity. The most fluent speech was observed during the combination of simple task and inclusion of some planning time, thus supporting R. Ellis (2009) assumption that it seems to be more likely that higher fluency scores are obtained through planning time in simpler tasks than in complex tasks. An explanation for the outcome of this study might be that the students were already familiar with the topic and task design, thus planning time was not essential for a rather fluent but very accurate and complex output production during the more complex tasks.

According to Larsen-Freeman (2009: 584) task repetition may have beneficial effects on language production, even if those effects do not occur immediately. Similar to priming, task repetition can positively influence task performance, thus more familiar tasks may evoke more complex, fluent as well as accurate language. Regarding this study, the Complex Task, a discussion about certain topics, seemed to be more familiar to the students than the Monologue where the students were asked to hold a mini presentation in pairs. Bygate (2009: 264) argued that task repetition causes a loss in accuracy and fluency but an increase in complexity scores. In contrast, Ahmadian & Tavakoli (2010: 20) stated that task repetition has a beneficial influence on fluency *and* complexity. Considering that the Complex Task which is likely to be the most familiar task and most complex task for the learners, elicited more complex *and* more accurate speech, however with a slight decrease in fluency scores, those arguments can only be partly confirmed. Concluding, the positive effect of task repetition in combination with task complexity might have a beneficial impact on the performance domains, however, in order to confirm this hypothesis more research in this area would be required.

Additionally to pre-task planning time, students can plan on-line, i.e. during actual task performance. It is suggested that without time pressure students have enough time to carefully structure their output while speaking, "leading to increased complexity and accuracy" (R. Ellis 2009: 500). Obviously, fluency scores will suffer due to with-in task planning given that the student needs more time *during* speaking to articulate his/her ideas than when having enough preparation *before* the actual task performance. Interestingly, in R. Ellis' (2009) study, dysfluency and false starts tended to decrease in the course of task performance, thus R. Ellis (2009: 501) proposed the idea that unpressured on-line planning might cause the most beneficial impact at the beginning of a task. Regarding the Monologue and the most dysfluent task, it can be observed that the students throughout the different groups did not produce fewer pauses or false starts in the course of the mini presentation. It could be argued that the students had enough time for pre-task planning, thus the amount of on-line planning was already reduced. However, there is not enough evidence to support this assumption. Similar results can be observed for the Simple Taskand the most fluent task. In contrast, some groups of the Role Play showed that there is a tendency towards a reduction of pauses and false starts when the task progresses, however, in some groups the number of pauses and false starts increased again towards the end of the task, implying that pauses are more likely to follow a normal distribution curve. R. Ellis' (2009) implication for the language classroom that time pressure should not be emphasised at the beginning of the task, but that it might be more appropriate to be emphasised towards the end of task performance, still seems to be a rather practical approach, as students often need to be reminded that the task needs to be finished in time.

Regarding time pressure and its influence on language production, it has been argued that when feeling pressured the students are more likely to rely on simple language (cf. Skehan 2009b). Particularly the achievement of a communicative goal can have an enormous influence on task performance, which can be seen in a previously stated example of this research, where students tried to achieve a communicative goal at the expense of certain performance domains. Additionally, Skehan (2009b: 93) argued that limited time can reduce modification, thus the students prioritize conveying the meaning successfully even though language complexity, accuracy and fluency might suffer. Interestingly, Foster, Tonkyn and Wigglesworth (2000: 356) found out that non-native speakers showed a tendency for word-by-word processing when time pressure

was reduced, implying that time pressure can push the students to use more idioms and more native like language. However, when time pressure was weakened a significant rise in complexity, accuracy and fluency scores was observed (ibid.). An example for a negative as well as a positive influence of time pressure can be seen in Extract 7, taken from the Complex Task, group 2. Before this episode, the students tended to make a lot of pauses which are significantly reduced towards the end of the task which benefited fluency. In contrast, the speaker mostly produces coordination without a great number of subordination, even though the teacher directed the student's attention towards the use of subordinate clauses – the use of conditionals. Again, the pauses are marked in bold and are in italics.

> **Extract 7, The Complex Task, Group 2**
>
> T: try to come to a conclusion in a minute, ja check the vocabulary again ja? see if you can use more if sentences
>
> S2: yeah I'm not very often in [Country] so there <un> xx </un> to talk about them if they don't work they put their guns down (.) and use that and solve their problem at another way. and that's not a joke that's the problem (.) they don't talk to another they if that don't work? also don't solve the problem they take their guns and shoot at each other. (.) but ***ah*** more the *(.)* crazy people will do that normal people ja they talk about it (.) and ignore mostly their problem.

Strikingly, the student pauses rather frequently, however, not mid-clause but end-clause. Those pauses do not affect fluency, as end-clause pauses are more natural and frequently used in native-speaker talk.

9.2.2 *CAF - an interdependent relationship*

Complexity, accuracy and fluency are three language performance domains which can be measured separately, however, an interdependent connection between those three factors cannot be denied. It has been argued that complexity, accuracy and fluency compete with each other during task performance (cf. Skehan & Foster 1997; Housen & Kuiken 2009), and the results of this study support this view to some extent. Particularly fluency seems to stand in a more competitive position towards complexity and accuracy, thus the assumption that the student focuses either on meaning, i.e. fluency or on form i.e. accuracy and complexity might be true in some cases (cf. Skehan 2009a). It has been argued that the student might tend to speak less fluently when the focus lies on

form, especially on complexity, as more attentive capacities are used to structure, modify and complexify the output (cf. Bygate 2009). In contrast, during this study the students spoke rather fluently and used more complex language during the task performance of the complex tasks, only the narrative task showed a slight increase in complexity in connection with dysfluent language. A similar tendency can be observed when comparing fluency and accuracy scores, where the most accurate speech production occurred during the complex task, again accompanied by fewer pauses and false starts than during the monologue. However, during simple task performance the least number of pauses and false starts was found while complexity scores remained rather low, thus showing that students tend to speak more fluently when task demands are not too cognitively challenging. On the contrary, complexity scores reached a peak with considerably acceptable fluency scores during complex task performance. A trade-off effect between fluency and task complexity can be confirmed, however, fluency levels can still be at an adequate level when complex and accurate language is produced.

Turning to the relationship of accuracy and complexity, a more significant result was obtained. The results of this study draw a parallel to Skehan's and Foster's (2007) study which showed that the decision making task, i.e. a complex task, elicited both more complex and more accurate language production, assuming that attention is allocated not separately towards accuracy or complexity but that the focus may lie on form in general. An additional study (Foster & Skehan 2009) resulted in similar outcomes which support the hypothesis that a cognitively challenging task, however not too demanding, can promote an improvement of both accuracy and complexity scores. In contrast to this study where accuracy and complexity scores were at their highest during complex task performance, Foster's and Skehan's (2009) results show that the highest complexity scores were achieved during the narrative task and the highest accuracy scores during the simple task. Hence, a trade-off effect seems to be more student dependent than assumed; however, there seems to be a tendency for a concurrent improvement of accuracy and complexity along increasing task complexity. Additionally, Gilabert (2007: 64) stated that complexity and accuracy show a stronger relationship than complexity and fluency or accuracy and fluency. It is further argued that the allocation of attention seems more likely to influence accuracy and complexity, whereas high fluency scores might have a stronger relation to planning and structuring

of the output before actual task performance (Gilabert 2007: 64). This might present a plausible explanation for the more significant relationship between accuracy and complexity scores in this study. Accuracy and complexity may need to be seen to stand in a closer relationship, thus changing the triadic distinction back to the broader differentiation between form, i.e. accuracy and complexity, and meaning, i.e. fluency. Contrary to a similar previous distinction, it seems to be more likely that when students focus on form, not only one form-based performance domain improves at the expense of the other, but both can increase simultaneously, supporting Robinson's (2003; 2001) Cognition Hypothesis.

An interesting development of Complexity L can be observed when comparing the monologic task to the more complex task. During the narrative task, Complexity L was at its highest, whereas the complex task only seemed to elicit shorter clauses. Given that the students were not disrupted during the monologue, they had time to elaborate their phrases during speech production without the danger of losing their position as a speaker. In contrast, during the complex task, the students had to discuss several topics, give their own opinion and react to the other interlocutors, thus quick turn-taking appeared to shorten clause length but enhance the use of subordinate clauses. Similarly, Michel, Kuiken and Vedder (2007) observed that interactive tasks with quick turn-taking tended to result in lower Complexity L scores than narrative tasks. Furthermore, Complexity L measured through mean length of clause might be a more adequate measure for advanced learners (Norris & Ortega 2009: 563f.) which might be another explanation for a decrease of Complexity L along increasing task complexity. Regarding Accuracy and Complexity L and language learner proficiency, those two performance domains seem to compete for cognitive capacities at an intermediate level. Larsen-Freeman (2006: 582f.) supports this assumption by arguing that trade-off effects may change in the course of the learning process and that proficiency levels can have a supportive as well as a more competitive impact on the development of the interlanguage system.

Tasks and activities proved to be very influential on proficiency and performance. As this study in combination with other recent studies (e.g. Skehan & Foster 2007; Larsen-Freeman 2006; Gilabert 2007; Michel, Kuiken & Vedder 2007; Foster & Skehan 2009; Bygate 2009 et cetera) has shown, task complexity and the performance domains are highly inter-dependent. The hypothesis that a slightly

linguistically as well as cognitively challenging task may promote a development in the performance domains can thus be reinforced. Individual differences as well as planning time appear to effect language production and classroom discourse significantly. However, to completely understand how individual differences and planning time interact with task performance and the performance domains, more research is required.

Conclusion

The main objective of this thesis was to explore how classroom discourse, especially tasks and activities can influence the three performance domains, complexity, accuracy and fluency. Additionally, cognitive factors proved to play an influential role during task performance.

The presentation of several theoretical approaches regarding the language acquisition and the language learning process as well as the influence of the competence-based classroom indicated that the process of learning a new language is closely connected to the active use of the language in the classroom. However, not only the input provided by the teacher is of importance, comprehensible output and negotiation of meaning enable the learner to actively use the target language and to consciously notice certain features and particularities while performing a task. The second chapter emphasised the fact that cognition and individual differences exert a huge impact on how language is acquired. Additionally to the importance of explicit and implicit knowledge while learning a new language, attentional resources and individual differences, like working memory capacities can facilitate learning a new language and support the student to perform more cognitively challenging and demanding tasks.

The three performance domains and their measurements were presented in the following chapter. Given the fact that complexity, accuracy and fluency are more likely to benefit when the student has the possibility to produce output more freely and without teacher interference, a definition of task-based language teaching and a general introduction to typical second language classroom discourse was provided. Additionally, I described task complexity and how it might influence task performance based on Robinson's (2007a) Triadic Componential Framework. Regarding the interaction of the three performance domains, two main opposing theories can be found in the literature. On the one hand, the trade-off hypothesis which provides a more competitive view of complexity, accuracy and fluency, and on the other hand, the cognition hypothesis which argues for a simultaneous improvement of the performance domains along increasing task complexity. In addition to the influence of task complexity, planning time has been argued to positively affect fluency, whereas there is mixed evidence for a beneficial impact of planning time on accuracy and complexity.

Based on the conceptual background and findings of recent studies, I formulated four main hypotheses which regard (1) the influence of task complexity on the performance domains, (2) the beneficial effects of planning time, (3) the question if cognitively too challenging tasks promote a trade-off effect and, finally, (4) the influence of task condition, i.e. interactive or narrative task design, on complexity, accuracy and fluency.

The empirical part of this study was conducted in a high school setting in a "Bundeshandelsakademie", i.e. a business school in Vienna. The participants frequented the third grade of this school, they were aged between 16 and 18 years and the general level was at B1+ (cf. CEFR), in other words, at an intermediate proficiency level. A selection of tasks, ranging from simple to more complex tasks, was recorded and analysed. Statistical tests confirmed the significance of the results and supported some of the previously made hypotheses.

Particularly Complexity S and Complexity L scores were highly significant, followed by accuracy scores which showed a tendency for significance. In contrast, fluency scores did not show such significant results. The Simple Task elicited the most fluent speech with mediocre Accuracy and Complexity L levels, however, complexity S scores remained rather low. The Monologue was the most dysfluent and the most inaccurate task, with medium Complexity S levels, however, Complexity L scores reached a peak during narrative task performance. Almost all measurements remained at a medium level during task performance of the Role Play, except slightly low Complexity L scores. The Complex Tasks turned out to elicit the most accurate and complex speech with normal fluency scores, however, at the expense of Complexity L.

Based on the results of this study, some of the previously made hypotheses could be confirmed, however, in order to formulate a clear argument about the effects of planning and task complexity, more research in this particular field would be required and desired. What is of importance regarding the measurements of complexity is that different types of complexity might benefit from different task types. This study showed that Complexity L (measured through mean length of clause) and complexity S (measured through subordination) can stand in an opposing position, i.e. presenting a trade-off effect. Regarding the outcomes in general, the cognition hypothesis could be confirmed to some extent; similarly, in some cases a trade-off effect appeared to take

place. Hence, it might be argued that both hypotheses seem to have legitimate arguments depending on the particular situation, context and the individual student. Strikingly, individual differences seemed to exert an even greater influence on task performance than task complexity, whereas due to the lack of sufficient evidence no clear statement about planning time can be made. In order to provide a more detailed picture of the importance of planning time and the influence of overwhelming task complexity on second language acquisition, more research in this area would be required.

As regards the interplay of complexity, accuracy and fluency, the results suggest that there seems to be a tendency towards a form and meaning distinction. In other words, accuracy and complexity show a stronger relationship (Gilabert 2007), implying that a differentiation of form, i.e. complexity *and* accuracy, and meaning, i.e. fluency might be more appropriate than a triadic distinction. Regarding classroom discourse and task sequencing, however, it appears to be important that a varied range of tasks with different complexity levels are included to assure a balanced development of all three performance domains. What is of importance for teaching and learning is the proposition that cognitive processes and individual differences are highly influential on output production and that the teacher as well as task design can only support the student in testing new hypotheses, using "cutting edge language" (Skehan 2009b) and improving their interlanguage system.

References

Coursebook and Portfolio

Tilbury, Rea, Clementson, Hendra, Donotha Mayr, Mhakar and Pope-Hoffmann. 2011. *English Unlimited 3. Coursebook BHAK/HUM. B1+*. Cambridge: CUP.

Abuja, Annau, Ganster, Keiper, Mayer-Taushitz, Mittendorfer, Nezbeda, Steinhuber and Winkler. 2005. *Das europäische Sprachenportfolio 15+ für junge Erwachsene. Sekundarstufe II – European Language Portfolio.* Graz/Salzburg: Österreichische Sprachen-Kopmetenz-Zentrum/ÖSZ & Cetner für Berufsbezogene Sprachen/CEBS.

CEFR. Council of Europe. 2001. *Common European Frame of Reference for Languages*. Cambridge: CUP.

Lehrplan Englisch
https://www.bmbf.gv.at/schulen/unterricht/lp/lp_ahs_os_lebende_fs_11854.pdf?4dzgm2 [25.4.2015]

Secondary Literature

Ahmadian & Tavakoli. 2010. "The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production." *Language teaching research* 15(1), 35-39.

Atkinson, D. (Ed.). 2011. *Alternative Approaches to Second Language Acquisition*. Routledge: New York.

Baddeley, A.D. & G.J. Hitch. 1974. Working memory. In G. Bower (ed.) *The psychology of learning and motivation* 8, 47-90. (New York: Academic Press)

Baddeley, Alan. 1992. "Working memory." *Science Biography in Context*. Web. [26.11.2014]
[http://ic.galegroup.com/ic/bic1/AcademicJournalsDetailsPage/AcademicJournalsDetailsWindow?failOverType=&query=&prodId=BIC1&windowstate=normal&contentModules=&displayquery=&mode=view&displayGroupName=Journals&limiter=&u=43wien&currPage=&disableHighlighting=false&displayGroups=&sortBy=&source=&search_within_results=&p=BIC1&action=e&catId=&activityType=&scanId=&documentId=GALE%7CA11973025]

Baddeley, A.D. 2003. "Working memory and language: an overview". *Journal of Communication Disorders* 36, 189-208.

Bialstock, E. 1994. Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition* 16, 157-168.

Braidi, S.M. 2002. "Reexamining the role of recasts in native-speaker/nonnative-speaker interactions." *Language Learning* 52(1), 1-42.

Breyer, K.M. 2000. "Task-based language teaching: theory, research, practice." Diplomarbeit, University of Vienna.

Bygate, M. 2009. "Effects of task repetition on the structure and control of oral language". In K. Van den Branden, M. Bygate, and J. Norris (eds.). *Task-based Language Teaching.* Philadelphia: John Benjamins.

Dahl, Ö. 2004. *The Growth and Maintenance of Linguistic Complexity.* Amsterdam: John Benjamins Publishing Company.

DeJong, N. 2009. "Second Language Learning of Grammar: Output Matters too." In Piske, T. & M. Young-Scholten. *Input Matters in SLA.* Bristol: Multilingual matters, 95-115.

DeKeyser, R. 2008. "The complexities of defining complexity." Paper presented at AAAL 2008, April 1st 2008, Washington D.C.

DeKeyser, R. 1998. "Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar." In C. Doughty & J. Williams (eds.). *Focus on form in classroom language acquisition.* New York: Cambridge University Press, 42-63.

Deutscher, G. 2009. "Nominalization and the origin of subordination". In T. Givón and Masayoshi Shibatani (eds.). *Syntactic Complexity. Diachrony, acquisition, neuro-cognition, evolution.* Amsterdam/Philadelphia: John Benjamins Publishing Companiy, 199-214.

Dewaele, J.M. 2009. " Individual differences in second language acquisition" In W.C. Ritchie and T.K. Bhatia (eds.). *The new handbook of second language acquisition.* Bingley: Emerald Group Publishing Limited.

Ellis, R. 1994. *The study of second language acquisition.* New York: OUP

Ellis, R. & Barkhuizen, G. 2005. *Analysing learner language.* Oxford: OUP.

Ellis, R. 2009. "The differential effects of three types of task planning on the fluency, complexity and accuracy in L2 oral production." *Applied Linguistics* 30(4), 474-509.

Ellis, R. 2003. *Task-based Language Learning and Teaching.* Oxford University Press.

Ellis, R. 2008. *The Study of Second Language Acquisition (2nd edition).* Oxford: Oxford University Press.

Ellis, R. & F. Yuan. 2004. "The effects of planning on fluency, complexity, and accuracy in second language narrative writing". *SSLA* 26, 59-84.

Ellis, R & Shintani. 2014. *Exploring language pedagogy through second language acquisition* research. London: Routledge.

Ellis, N. C. (2005). "At the interface: Dynamic interactions of explicit and implicit language knowledge." *Studies in Second Language Acquisition* 27, 305-352.

Ellis, N.C. 2001. "Memory for language." In P. Robinson (ed.). *Cognition and Second Language Instruction*. Cambridge: CUP, 33-69

Ellis, N.C. 1994. *Implicit and Explicit learning of languages*. San Diego: Academic Press Inc.

Ellis, N.C. & S. Sinclair. 1996. "Working memory in the acquisition of vocabulary and syntax : Putting language in good order." *Quarterly Journal of Experimental Psychology* 49A, 234-250.

Foster, P., A. Tonkyn and G. Wigglesworth. 2000. "Measuring spoken language: A unit for all reasons." *Applied Linguistics* 21, 354-75.

Foster, P. and P. Skehan. 2009. "The influence of planning and task type on second language performance. In K. Van den Branden, M. Bygate, and J. Norris (eds.). *Task-based Language Teaching*. Philadelphia: John Benjamins.

Gass, S. 1988. "Integrating research areas: A framework for second language studies." *Applied Linguistics* 9, 198-217.

Gass, S. 1984. "The empirical basis for the universal hypothesis of interlanguage studies." In Davies, A., Criper, C.and Howatt, A.P.R (eds.). *Interlanguage.* Edinburgh: Edinburgh University Press.

Gass, Susan and Alison Mackey. 2006. "Input, interaction and output. An overview". *AILA Review* 19, 3-17.

Gass, S., J. Behney & L. Plonsky. 2013. *Second Language Acquisition: An Introductory Course.* New York: Routledge.

Gass, S. & A. Mackey. 2007. *Data elicitation for second and foreign language research.* New Jersey: Lawrence Erlbaum Associates.

Gilabert, R. 2007. "The simultaneous manipulation of task complexity along planning and (+/- Here-and-Now): effects on L2 oral production." In Maria del Pilar García-Mayo (ed.). *Investing tasks in formal language learning.* Clevedon: Multilingual Matters, 44-68.

Givón, T. & M. Shibatani (Eds.). 2009. *Syntactic Complexity. Diachrony, acquisition, neuro-cognition, evolution.* Amsterdam: John Benjamins Publishing Company.

Greenbaum, S. and R. Quirk. 1990. *A Student's Grammar of the English Language.* London: Longman Group UK Limited.

Hilton, H. 2008. "The link between vocabulary knowledge and spoken L2 fluency." *Language Learning Journal* 36/2: 153-166.

Housen, A. & F. Kuiken. 2009. "Complexity, accuracy and fluency in second language acquisition." *Applied Linguistics* 30(4), 461-473.

Housen, A., S. Van Daele and M. Pierrard. 2005. "Rule complexity and the effectiveness of explicit grammar instruction" In Housen, A. and M. Pierrard (eds.): *Investigations in Instructed Second Language Acquistion.* Berlin: Mouton de Gruyter.

Iwashita, N., T. McNamara, and C. Elder. 2001. "Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information processing approach to task design." *Language Learning* 21(3), 401–36.

Krashen, S. 2009. "The Comprehension Hypothesis Extended." In Piske, T. & M. Young-Scholten. *Input Matters in SLA.* Bristol: Multilingual Matters, 81-94.

Krashen, S. D. 1984 [1981]. *Second language acquisition and second language learning.* Oxford: Pergamon.

Krashen, S. D. 1982. *Principles and practice in second language acquisition.* Oxford: Pergamon.

Krashen, S. 1985. *The Input Hypothesis: Issues and Implications.* London and New York: Longman.

Lantolf, J. 2000. "Introducing sociocultural theory." In J. Lantolf (ed.) *Sociocultural Theory and Second Language Learning.* Oxford: OUP.

Larsen-Freeman, D. 2006. "The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English." *Applied Linguistics* 27(4), 590-619.

Larsen-Freeman, D. 2009. "Adjusting expectations: the study of complexity, accuracy, and fluency in second language acquisition". *Applied Linguistics* 30(4), 579-589.

Long, M. H. 1996. "The role of the linguistic environment in second language acquisition." In Ritchie, W.C.; Bhatia, T.K. (eds). *Handbook of language acquisition* 2. New York NY: Academic Press, 413-468.

Levelt, W. 1991 [1989]. *Speaking: From Intention to Articulation.* Cambridge, MA: MIT Press.

Ludden, D. and P. Gupta. 2000. "Zen in the art of language acquisition: statistical learning and the less is more hypothesis." In L., R. Gleitman and A. K. Joshi (eds.). *Proceedings of the 22nd Annual Conference of the Cognitive Science Society.* Hillsdale, NJ: Erlbaum.

Lyster, R. and L. Ranta. 2013. "Counterpoint piece: the case for variety in corrective feedback research". *Studies in Second Language Acquisition* 35, 167-184.

Mackey, Alison. 1999. "Input, interaction, and second language development. An empirical study of question formation in ESL". *SSLA* 21, 557-587.

Mackey, A. 2012. *Input, Interaction, and Corrective Feedback in L2 Learning.* Oxford: Oxford University Press.

Mackey, A. and J. Philp. 1998. "Conversational interaction and second language development: recasts, responses, and red herrings?". *Modern Language Journal* 82, 338-56.

McDonough. 2006. "Interaction and syntactic priming: English L2 speaker's production of dative constructions". *Studies in second language acquisition* 28, 179-207.

Müller-Hartmann, A. & M. Schocker-von Ditfurth. 2011. *Task-supported language learning.* Parderborn, München, Wien, Zürich: Ferdinand Schöningh.

Newell, A. 1990. *Unified theories of cognition.* Cambridge, MA: Harvard University Press.

Newport, E. 1990. "Maturational constraints on language learning." *Cognitive Science* 14, 11-28.

Norris, J. & L. Ortega. 2009. "Towards an organic approach to investigating CAF in instructed SLA: The case of complexity." *Applied Linguistics* 30, 555-578.

Norris, J. and L. Ortega. 2003. "Defining and measuring SLA" In Doughty, C. and M. Long (eds). *Handbook of Second Language Acquisition.* Oxford: Blackwell, 717-761.

Ortega, L (ed.). 2011. *Second language acquisition: critical concepts in linguistics.* London: Routledge.

Ortega, L. 2003. "Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing." *Applied Linguistics* 24: 492-518.

Ortega, L. 2009. *Understanding Second Language Acquisition.* London: Hodder Education.

Ortega, L. (ed.). 2005. "Methodology, epistemology, and ethics in instructed SLA research". *Special Issue of the modern language journal* 89(3),

Pallotti, G. 2009. "CAF: Defining, Refining and Differentiating Constructs". *Applied Linguistics* 30(4). Oxford: OUP, 590-601.

Paradis, M. 1994. "Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA." In N.C. Ellis (ed.). *Implicit and explicit learning of languages.* San Diego, CA: Academic Press, 393-420.

Peterwagner, R. 2005. *What is the matter with communicative competence? An analysis to encourage teachers of English to assess the very basis of their teaching.* Wien: Lit Verlag.

Pica, Teresa. 2005. "Classroom learning, teaching, and research: a task-based perspective". *The Modern Language Journal* 89(3), 339-352.

Pilar García Maya, M., M. Junkal Gutierrez Mangado & M. Martínez Adrían (Eds.). 2013. *Contemporary approaches to Second Language Acquisition.* Amsterdam: John Benjamins Publishing Company.

Prabhu, N.S. 1987. *Second language pedagogy.* Oxford: OUP.

Robinson, P. 2001. 'Task complexity, task difficulty, and task production: exploring interactions in a componential framework.' *Applied Linguistics* 22, 27-57.

 Robinson, P. 2003. 'Attention and memory during SLA' In Doughty, C. and M. Long (eds). *The handbook of second language acquisition.*

Robinson, P. 2007a. "Criteria for grading and sequencing pedagogic tasks". In Maria del Pilar Garcia Mayo (ed.). *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters, 7-27.

Robinson, P. 2007b. "Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty." *IRAL* 45,193-213.

Robinsnon, P. and R. Gilabert. 2007. "Task complexity, the Cognition Hypothesis and second language learning and performance". *IRAL* 45, 161-176.

Robinson, P., T. Cadierno and Y. Shirai. 2009. "Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production." *Applied Linguistics* 30(4), 533-554.

Robinson, P., A. Mackey, S. M. Gass and R. Schmidt. 2014. "Attention and awareness in second language acquisition" in S. Gass and A. Mackey (eds.). *The Routledge Handbook of Second Language Acquisition.* New York: Roudledge.

Robinson, P. and N. C. Ellis. 2008. "Conclusion: Cognitive linguistics, second language acquisition and instruction: Issues for research." In Robinson, P. and N. C. Ellis (eds.): *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Routledge.

Schleppegrell, M. (2004). *The language of schooling: A functional linguistics perspective.* Mahwah, NJ: Lawrence Erlbaum Associates.

Schmidt, R. 2001. "Attention." In P. Robinson (ed.) *Cognition and Second Language Instruction*. Cambridge: CUP, 3-32.

Seedhouse P, Walsh S. 2010. "Learning a Second Language through Classroom Interaction." In Seedhouse, P., Walsh, S., Jenks, C (eds.). *Conceptualising 'Learning' in Applied Linguistics*. Basingstoke: Palgrave Macmillan, 127-146.

Selinker, L. 1972. "Interlanguage." *International Review of Applied Linguistics* 10(3), 209-231.

Skehan, P. 2009a. "Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis." *Applied Linguistics* 30(4), 510-532.

Skehan, P. 2009b. "A framework for the implementation of task-based instruction". In K. Van den Branden, M. Bygate and J. Norris (eds.). *Task-based language teaching. A reader.* Amsterdam/Philadelphia: John Benjamins Publishing Company, 83-108.

Skehan, P. 2001. "Tasks and language performance." In Bygate M., P. Skehan and M. Swain (eds.): *Research Pedagogic Tasks: Second Language Learning, Teaching, and Testing*. London: Longman.

Skehan, P. 1998. *A cognitive approach to language learning.* Oxford: OUP.

Skehan, P. 1992. Strategies in second language acquisition. *Thames Valley University Working Papers in English Language Teaching* 1

Skehan, P. and P. Foster. 2007. "Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research." In S. Van Deale, A. Housen, F. Kuiken, M. Pierrard & I. Veder (Eds.). *Complexity, accuracy, and fluency in second language use, learning, and teaching.* Brussels: University of Brussels Press.

Skehan, P. and P. Foster. 2001. "Cognition and tasks" In P. Robinson (ed.) *Cognition and Second Language Instruction.* Cambridge: CUP.

Skehan, P. and P. Foster. 1997. "Task type and task processing conditions as influences on foreign language performance." *Language Teaching Research* 1, 185-211.

Swain, Merrill. 2005. "The output hypothesis: theory and research". In Hinkel, Eli (ed). *Handbook of research in second language teaching and learning.* New Jersey: Lawrence Erlbaum Associates, 471-483.

Swain, Merrill. 1985. "Communicative competence: some roles of comprehensible input and comprehensible output in its development". In Gass, Susan and Madden, C. (eds). *Talking to learn: conversation in second language acquisition.* Rowley, MA: Newbury House, 235-253.

Tonkyn, A. 2007. "Short-term changes in complexity, accuracy and fluency: Developing progress-sensitive proficiency tests" In Van Daele et al. (eds.). *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching*. Brussels: Contactforum.

Towell, R. 2007. "Complexity, accuracy and fluency in second language acquisition research." In Van Daele et al. (eds.). *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching*. Brussels: Contactforum.

Van Daele, S., A. Housen, F. Kui ken, M Pierrard and I. Vedder. (eds). 2007. *Complexity, Accuracy and Fluency in Second Language Use, Learning and Teaching*. Brussels: Contactforum.

Van den Branden, K. 2006. "Introduction: Task-based language teaching in a nutshell". In K. Van den Branden (ed.). *Task-based language education: From theory to practice.* Cambridge: CUP.

Van den Branden, K., M. Bygate and J. Norris. 2009. "Task-based language teaching: Introducing the reader". In K. Van den Branden, M. Bygate and J. Norris (eds.). *Task-based language teaching. A reader.* Amsterdam/Philadelphia: John Benjamins Publishing Company, 1-15.

Walsh, S. 2012. "Conceptualising classroom interactional competence." *Novitas-ROYAL* 6(1), 1-14.

Walsh, S. 2011. *Exploring classroom discourse: Language in action.* London: Routlege

Walsh, S. 2006a. *Investigating Classroom Discourse.* London: Routledge.

Walsh, S. 2006b. "Talking the talk of the TESOL classroom." *ELT Journal* 60(2), 133-141.

Williams, J. and J. Evans. 1998. "What kind of focus and on which forms?" In Doughty, C. and J. Williams (eds.): *Focus on Form in Classroom Second Language Acquisition.* Cambridge University Press.

Willis, J. 1996. *A framework for task-based learning.* Harlow: Longman.

Widdowson, H.G. 2003. *Defining issues in English language teaching.* Oxford: OUP.

Widdowson, H.G. 1989. "Knowledge of language and ability for use". *Applied Linguistics* 10(2), 128-137.

Wolfe-Quintero, K., S. Inagaki & H.Y. Kim. 1998. *Second language development in writing: Measures of fluency, accuracy & complexity.* Honolulu: University of Hawaii Press.

Yuan, F. and R. Ellis. 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.
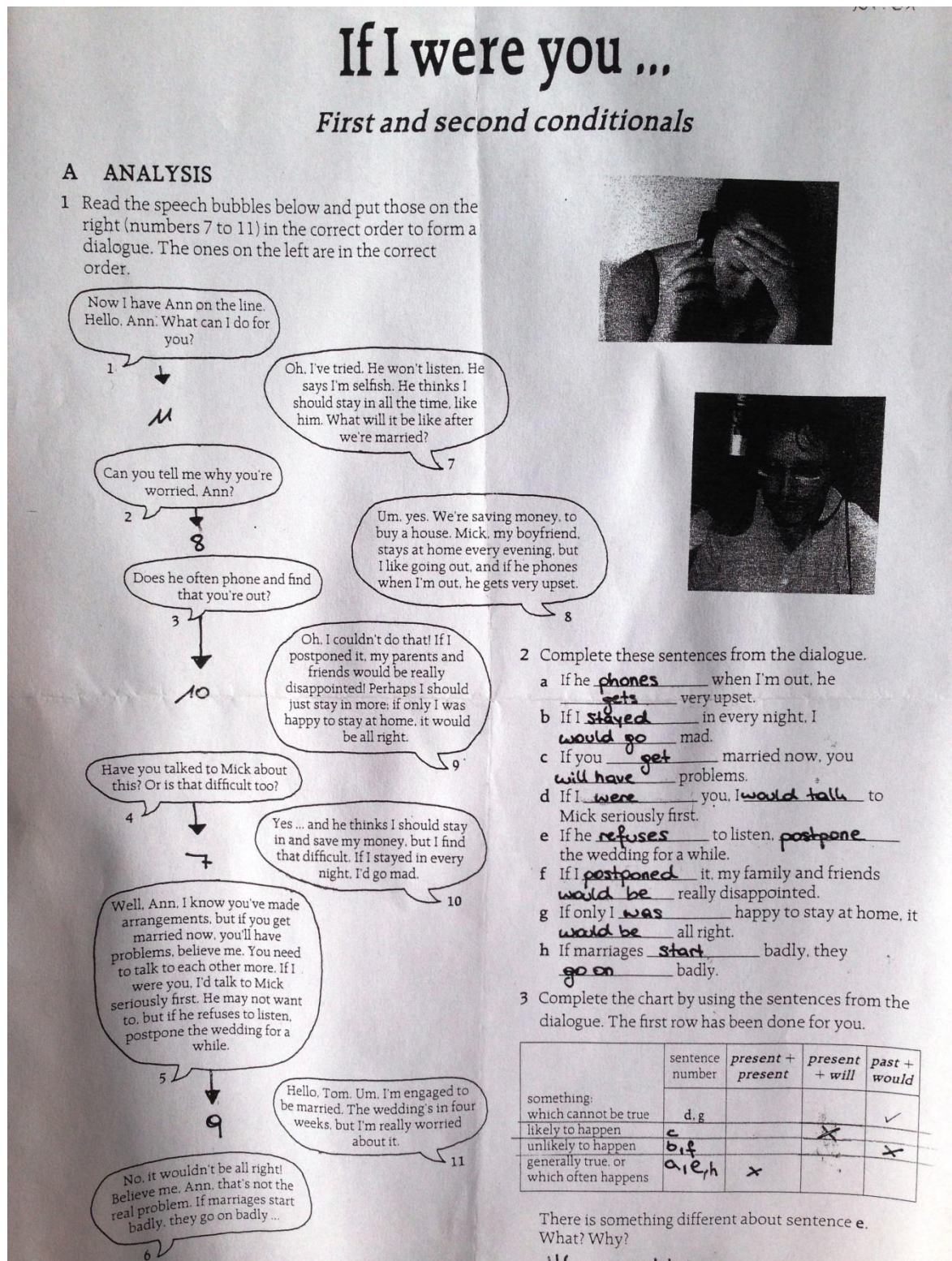
Appendix

Handout

Tables

Paired Sample T-Tests

Table 13: Paired Sample tests: task complexity

Paired Samples Test

		Paired Differences							
				Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
		Mean	Std. Deviation		Lower	Upper			
Pair 1	Accuracy - task_complexity	7,100682039	2,133622236	1,066811118	3,705612938	10,49575114	6,656	3	,007
Pair 2	Fluency - task_complexity	22,44293596	7,443269434	3,721634717	10,59903330	34,28683862	6,030	3	,009
Pair 3	SyntacticComplexity - task_complexity	2,88888	1,38755	,69378	,68097	5,09678	4,164	3	,025

Table 14: Paired Sample tests: planning time

Paired Samples Test

		Paired Differences							
				Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
		Mean	Std. Deviation		Lower	Upper			
Pair 1	Accuracy - planning_time	8,350682039	1,608414181	,8042070903	5,791336156	10,91002792	10,384	3	,002
Pair 2	Fluency - planning_time	23,69293596	7,714717404	3,857358702	11,41709901	35,96877291	6,142	3	,009
Pair 3	SyntacticComplexity - planning_time	4,13888	,60777	,30388	3,17178	5,10597	13,620	3	,001
Pair 4	Complexity - planning_time	39,42336556	2,040964116	1,020482058	36,17573620	42,67099491	38,632	3	,000

Table 15: Paired Sample Tests: Accuracy

Paired Samples Test

		Paired Differences							
				Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
		Mean	Std. Deviation		Lower	Upper			
Pair 1	Accuracy - Fluency	-15,3422539	6,677550134	3,338775067	-25,9677263	-4,71678154	-4,595	3	,019
Pair 2	Accuracy - Complexity	-31,0726835	3,293039644	1,646519822	-36,3126444	-25,8327226	-18,872	3	,000
Pair 3	Accuracy - SyntacticComplexity	4,211805354	1,084262303	,5421311517	2,486502073	5,937108635	7,769	3	,004

Table 16: Paired Sample Tests: Complexity S

Paired Samples Test

		Paired Differences							
				Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
		Mean	Std. Deviation		Lower	Upper			
Pair 1	Complexity - Accuracy	31,07268352	3,293039644	1,646519822	25,83272259	36,31264444	18,872	3	,000
Pair 2	Complexity - Fluency	15,73042960	7,786777872	3,893388936	3,339928360	28,12093083	4,040	3	,027
Pair 3	Complexity - SyntacticComplexity	35,28448887	2,595329298	1,297664649	31,15474080	39,41423694	27,191	3	,000

Table 17: Paired Sample Tests: Fluency

Paired Samples Test

		Paired Differences							
				Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
		Mean	Std. Deviation		Lower	Upper			
Pair 1	Fluency - Complexity	-15,7304296	7,786777872	3,893388936	-28,1209308	-3,33992836	-4,040	3	,027
Pair 2	Fluency - SyntacticComplexity	19,55405927	7,509268707	3,754634353	7,605137049	31,50298150	5,208	3	,014
Pair 3	Fluency - Accuracy	15,34225392	6,677550134	3,338775067	4,716781544	25,96772630	4,595	3	,019

Table 18: Paired Sample Tests: Complexity L

Paired Samples Test

		Paired Differences							
				Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
		Mean	Std. Deviation		Lower	Upper			
Pair 1	SyntacticComplexity - Accuracy	-4,21180535	1,084262303	,5421311517	-5,93710863	-2,48650207	-7,769	3	,004
Pair 2	SyntacticComplexity - Fluency	-19,5540593	7,509268707	3,754634353	-31,5029815	-7,60513705	-5,208	3	,014
Pair 3	SyntacticComplexity - Complexity	-35,2844889	2,595329298	1,297664649	-39,4142369	-31,1547408	-27,191	3	,000

ANOVAs

Table 19: ANOVA Accuracy and planning

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1,727	1	1,727	1,149	,396[b]
	Residual	3,008	2	1,504		
	Total	4,735	3			

a. Dependent Variable: Accuracy

b. Predictors: (Constant), planning_time

Table 20: ANOVA Complexity S and planning

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	17,670	1	17,670	26,039	,036[b]
	Residual	1,357	2	,679		
	Total	19,027	3			

a. Dependent Variable: Complexity

b. Predictors: (Constant), planning_time

Table 21: ANOVA Complexity L and planning

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,022	1	,022	,566	,530[b]
	Residual	,078	2	,039		
	Total	,100	3			

a. Dependent Variable: SyntacticComplexity

b. Predictors: (Constant), planning_time

Table 22: ANOVA Fluency and planning

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2,802	1	2,802	,033	,873[b]
	Residual	172,099	2	86,049		
	Total	174,901	3			

a. Dependent Variable: Fluency

b. Predictors: (Constant), planning_time

Table 23: Coefficient test Complexity S and planning

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	34,606	1,258		27,502	,001
	planning_time	4,854	,951	,964	5,103	,036

a. Dependent Variable: Complexity

English abstract

Classroom discourse provides the learner with opportunities to improve the language performance domains: complexity, accuracy and fluency (CAF). Tasks and activities prompt the student to use the target language and express his/her ideas through either simple or more complex language. Task complexity is said to have an influential effect on language complexity, in more concrete words cognitively challenging tasks might stimulate the learner to produce more complex and more accurate language. Considering the trade-off hypothesis and the cognition hypothesis, the extent to which task complexity influences the development of the three domains and the question if these domains compete with each other during task performance, were observed. Additionally, the paper focuses on the impact of pre-task planning on CAF and the difference between narrative and interactive tasks. The findings have implications for classroom and task design as task complexity tends to have significant effects on language complexity, fluency of speech as well as accuracy.

German abstract

Der Diskurs im Klassenzimmer bietet den LernerInnen viele Möglichkeiten Komplexität, Genauigkeit und Flüssigkeit der Sprache, die Domänen der Sprachperformanz (abgekürzt CAF), zu verbessern. Aufgabenstellungen und Übungen veranlassen den Lerner/die Lernerin die Zielsprache zu verwenden und eigene Ideen zu formulieren, entweder durch den Einsatz von simpler oder komplexer Sprache. Man nimmt an, dass die Komplexität dieser Aufgabenstellung einen wirkungsvollen Einfluss auf die Sprachkomplexität ausübt. Folglich wird in dieser Arbeit davon ausgegangen, dass komplexere und kognitiv forderndere Aufgabenstellungen den Lerner/die Lernerin dazu stimulieren, komplexere und akkuratere Sprache zu produzieren. Das Ausmaß dieses Einflusses der Aufgabenstellungskomplexität wurde im Bezug auf die „Trade-off" Hypothese und der „Cognition" Hypothese analysiert. Weiters wurde die Frage, wie die Domänen der Sprachperformanz in Verbindung stehen, betrachtet und es wurde analysiert, ob diese Verbindung von konkurrierender Natur ist. Ein Teilfokus dieser Arbeit lag auf dem Einfluss von Planungszeit, entweder vor oder während der Aufgabenerfüllung, auf die Komplexität, Genauigkeit und Flüssigkeit der Sprache,

sowie ob Unterschiede zwischen narrativen und interaktiven Aufgaben vorhanden sind. Die Resultate zeigen Implikationen für das Klassenzimmer und für die Aufgabengestaltung auf, da eine Tendenz für eine signifikante Auswirkung der Aufgabenstellungskomplexität auf die Komplexität, Flüssigkeit und Genauigkeit der Sprache besteht.

CURRICULUM VITAE

Persönliche Angaben

Name   Julia Stromer

Geburtsdatum   25.05.1989

E-Mail   j.stromer@gmx.at

Ausbildung

WS 2011/12   Erasmus-Semester an der Université Catholique de l'Ouest
                        (Angers, F)

2008-2015   Lehramtsstudium Englisch Französisch an der Universität
                        Wien

2003-2008   Handelsakademie Retz

Besondere Berufserfahrung

SS 2014   Studienassistentin am Centre for English Language Teaching,
                       Department of English, Universität Wien

Seit Dezember 2014   Englischlehrerin an der Bundeshandelsakademie und
                       Handelsschule Wien 11

2013-2015   Nachhilfelehrerin bei Christiane Humer, die Nachhilfe
                      (Englisch und Französisch)

2008-2013   Touristenführerin durch Österreichs größten Weinkeller, Retz

2009-2012   Lernhilfe für Volksschüler, Wiener Kindergärten

Vorträge

Mai 2015   "Complexity in second language classroom discourse"
                     (*7. Nachwuchstagung der Österreichischen Gesellschaft für
                     Sprachendidaktik – ÖGSD,* PH Burgenland, Eisenstadt)