



UvA-DARE (Digital Academic Repository)

Focus on Form in task-based language teaching

van de Guchte, M.

Publication date

2015

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

van de Guchte, M. (2015). *Focus on Form in task-based language teaching*. [Thesis, fully internal, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Focus on Form in Task-Based Language Teaching

Marrit van de Guchte



Marrit van de Guchte Focus on Form in Task-Based Language Teaching

FOCUS ON FORM
IN TASK-BASED LANGUAGE TEACHING



UNIVERSITY OF AMSTERDAM

Cover Design: Monique van Hootegem
Cover Picture: termcoord.eu
Printed by Ipskamp Drukkers B.V., Amsterdam
ISBN: 978-94-6259-921-5

This publication was supported by grants from the Graduate School of Teaching and Learning, University of Amsterdam, the Gooise Scholenfederatie, Bussum and the A. Roland Holst College, Hilversum.

**FOCUS ON FORM
IN TASK-BASED LANGUAGE TEACHING**

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde
commissie, in het openbaar te verdedigen in de Aula der Universiteit

op donderdag 17 december 2015, te 13.00 uur

door

Marrit van de Guchte

Geboren te Goes

Promotiecommissie

Promotores: Prof. dr. G.C.W. Rijlaarsdam
Universiteit van Amsterdam

Dr. M.A.H. Braaksma
Onderwijsraad

Dr. P. E. Bimmel

Overige leden: Prof. dr. C.A.M. van Boxtel
Universiteit van Amsterdam

Prof. dr. K. van den Branden
Katholieke Universiteit Leuven

Dr. N.H. de Jong
Universiteit Utrecht

Prof. dr. F. Kuiken
Universiteit van Amsterdam

Prof. dr. M.L.L. Volman
Universiteit van Amsterdam

Faculteit: Faculteit der Maatschappij- en Gedragwetenschappen

TABLE OF CONTENTS

DANKWOORD	viii
CHAPTER 1 Introduction	1
CHAPTER 2 Focus on form in task-based language teaching	5
CHAPTER 3 Focusing on meaning and form in pre-task video model observations: Effects on planning processes and task performance	25
CHAPTER 4 Learning new grammatical structures in task-based language learning: The effects of recasts and prompts	49
CHAPTER 5 Focus on form through task repetition	73
CHAPTER 6 Discussion and conclusions	95
REFERENCES	111
AUTHOR INDEX	119
SUMMARY	123
SUMMARY (DUTCH)	129

DANKWOORD

Na verschillende jaren in zowel het voorgezet als universitair onderwijs te hebben gewerkt, kwam ik in 2011 terecht in het onderwijsonderzoek. Ik had de behoefte om de praktijk meer met theorie te verbinden en had me op dit pad geen betere begeleiders kunnen wensen dan Gert Rijlaardam, mijn promotor en Martine Braaksma en Peter Bimmel, mijn co-promotoren.

Ik wil beginnen met het bedanken van Gert Rijlaarsdam. Gert, vanaf het eerste moment heb je me zoveel vertrouwen gegeven in de afronding van dit proefschrift dat ik er daardoor zelf ook altijd in geloofd heb. Ik heb genoten van onze inhoudelijke discussies, het samen analyseren en interpreteren van de data, onze gezamenlijke presentatie bij mij op school en onze mailuitwisselingen. Je was niet alleen betrokken bij de inhoudelijke kant van het proefschrift, maar ook bij de menselijke kant en dat heb ik enorm gewaardeerd.

Lieve Martine, jij was als mijn dagelijks begeleider goud waard. Ik heb zoveel van je geleerd op het gebied van data-analyses, artikelen schrijven, congrespresentaties en zoveel meer. Nadat je me had geïntroduceerd in de wondere wereld van SPSS, hebben we vele momenten samen achter de computer doorgebracht en dat vonden we allebei heerlijk om te doen. Elk stuk dat ik je stuurde, had je in no-time van feedback voorzien waardoor ik weer snel verder kon. Naast alle proefschrift gerelateerde zaken hadden we ook heel fijn persoonlijk contact met als essentieel onderdeel 'goede koffie'. Ook denk ik met veel plezier terug aan het EARLI congres in München en onze Starbucks bezoeken.

Peter, wij hebben samen een mooi traject afgelegd. In 2001 was je mijn docent binnen de lerarenopleiding, in 2007 werden we collega's en in 2011 werd je mijn co-promotor. Ik verheugde me altijd op onze besprekingen in Hilversum bij jou en Annie 'aan de keukentafel'. Waar onze afspraken begonnen als mini-colleges waarin jij nog eens nauwgezet al je kennis met me deelde, gingen die later over in mooie discussies tussen ons twee over grammatica en taakgericht taalonderwijs.

Mijn speciale dank gaat uit naar Kees Elsinga, voorzitter van het bestuur van de Gooise Scholenfederatie, Loes Lauteslager en Louis van Kaam, beide oud-rectoren van het A. Roland Holst College. Naast de Universiteit van Amsterdam, hebben zowel de GSF als het ARHC mijn promotie-traject gefinancierd en binnen de school ondersteund. Ik dank jullie voor het vertrouwen en de ondersteuning in de afgelopen vier jaar.

Praktijkgericht onderzoek doen op je eigen school is onmogelijk zonder de hulp van vele collega's en leerlingen. Ik wil mijn collega's van de sectie Duits bedanken voor hun interesse en het testen van toetsen voor mijn onderzoek in hun eigen klassen. Ook veel dank voor de collega's van het Quest-team die de leerlingtaken voor me hebben beoordeeld en hun interesse hebben getoond. Dank ook aan alle leerlingen die de lessen en de toetsen hebben gemaakt. En bedankt Bea, voor het meedenken en het gebruik van de bibliotheek. Een speciaal woord voor Paul Halma, die gedurende de afgelopen vier jaar mijn teamleider bij Quest was, maar ook degene met wie ik de experimenten in de klassen heb uitgevoerd. Bedankt Paul voor het

nauwkeurig volgen van al mijn protocollen. Daarnaast heb je me op school op alle mogelijke manieren ondersteund en toonde je veel interesse voor het taakgericht leren. Ik heb veel gehad aan je luisterend oor en je feedback.

Claudia, Martina, Melanie en Ingrid, bedankt voor het afnemen van de toetsen, het uitwerken van alle opnames en het geruststellen van vele leerlingen.

Huub van den Bergh wil ik bedanken voor zijn hulp bij de statistische analyses. Folkert Kuiken en Marije Michel hebben de leerlingtaken voor me beoordeeld en van waardevolle feedback voorzien. Nivja de Jong heeft me zeer geholpen bij de *fluency* beoordelingen. Daphne en Christien, wat was het fijn dat jullie mijn Engelse teksten hebben geredigeerd en mijn *academic writing skills* naar een hoger plan hebben getild.

Heel veel dank ook voor alle onderzoekscollega's van POWL. Mijn kamergenoten, Jantina, Mark, Rijkje, Daphne en Tanja bedank ik voor hun hulp, interesse en de lol om alle flexwerkperikelen. Gerhard, we zijn min of meer tegelijk aan dit promotie-avontuur begonnen en het was dan ook ontzettend fijn om een maatje te hebben die in hetzelfde schuitje zat. Verder bedank ik de taalgroepcollega's, die ik allemaal veel beter heb leren kennen door het lunchlab. Bedankt voor jullie scherpe analyses, alle kennis die ik heb opgedaan door jullie onderzoek en natuurlijk de gezelligheid. Marloes en Marjolein, ik vond het altijd een fijn vooruitzicht om rond achten al een kop koffie met jullie te kunnen drinken en bij te kletsen. Lieve Natas, onze afspraken tussen jouw colleges en mijn artikelen door waren een heerlijke afleiding. Peter en Carla, door jullie aanmoediging heb ik me erg gesteund gevoeld binnen POWL. Ook bedank ik de collega-vakdidactici van de ILO en Hermine en Doris met wie ik de afgelopen vier jaar altijd een goed contact heb gehouden. Eline, bedankt voor alle TBLT- en fijne andere gesprekken.

Een promotietraject doe je niet alleen, daar heb je de steun van je familie en vrienden hard bij nodig. Daarom wil ik mijn ouders graag bedanken die mij altijd enorm gesteund hebben en die Anna en Ties vaak te logeren hebben gehad wanneer ik moest schrijven. Dank je lieve mam, voor alle wekelijkse telefoontjes waarin niets je ooit teveel was. Lieve zus, jij die er altijd voor me bent en die leuke dingen met Anna en Ties deed wanneer ik weer eens in de UB zat. Bedankt ook lieve schoonfamilie voor de support. Lieve Yvette, jij weet met je positieve en nuchtere kijk mij altijd van de juiste tips te voorzien. Lieve Eva, ook al woon je nu op Aruba, in mijn hart ben je er bij. Lieve Brit, samen de ILO gedaan, samen promoveren, jij moest wel mijn paranimf worden. Dank voor alle telefoongesprekken en koffiedates waarbij we konden spuien over ons proefschrift en onze successen konden delen.

Lieve Rik, zonder jou was dit hele traject niet mogelijk geweest. Je hebt me altijd gesteund en aangemoedigd ook al kwam er de afgelopen jaren veel op jouw schouders terecht. Je weet me als geen ander aan het lachen te maken en er is niets fijners dan thuiskomen en plezier maken met jou en de kinderen. Lieve Anna en Ties, mama's 'boekje' is af en nu kunnen we weer heel veel leuke dingen gaan doen.

CHAPTER 1

INTRODUCTION

1. AIM AND SCOPE OF THIS THESIS

In 2009 I started working as a teacher of German language and literature at the ‘Quest’ department of a Dutch secondary school. In the ‘Quest’ department the language curriculum for German, English, and French is based on the principles of task-based language teaching (TBLT). TBLT is an educational framework for the theory and practice of teaching second and foreign languages, which places meaning-based, communicative tasks at the heart of language procedures in the classroom (Van den Branden, Bygate & Norris, 2009). Communicative tasks are tasks such as:

You are having a party. Tell your partner how to get from the school to your home.
(p. 29)

Listen to the automated ticketing service for ‘What’s on around town this weekend’.
Make a list of movies, and concerts and how much they cost. Work with three other students and decide where to go. (p. 27)

You are in a clothing store and have \$150 to spend. Look at the clothing items on the worksheet. Find out the prices, and decide what to buy. (p. 27)

(Nunan, 2004)

Having learners perform communicative tasks enables them to acquire new linguistic knowledge and to proceduralize their existing knowledge (Ellis, 2009a). An important difference to more traditional approaches to language teaching is that in TBLT language is treated as a tool for making meaning rather than as an object to study (Ellis, 2003).

In addition to the ‘Quest’ department, the school where I teach German also has a regular department in which students follow a more grammar-focused textbook-based language curriculum. Some teachers of this regular language department expressed their doubts about the task-based language teaching approach used in the Quest department. They feared that language curriculum focused mainly on content and pleasure and not on the acquisition of specific grammatical structures (see also criticism on TBLT by Sato, 2010). Because I was familiar with the theoretical rationale for TBLT through my other professional role as a teacher educator at the University of Amsterdam, I felt the need to underpin the effectiveness of task-based language teaching with empirical research. Both the former director of the A. Roland Holst College, the president of the board of the Gooise Scholen Federatie, and

the director of the Research Institute of Child Development and Education of the University of Amsterdam supported this idea and offered me a grant to conduct this research.

The current school dilemma, focus on content (meaning) versus focus on grammar (form) became the starting point of my PhD research project. Reviewing literature on TBLT principles, I concluded that working with tasks as a central unit of the curriculum did not exclude a focus on grammar. Moreover, according to Van den Branden, Bygate and Norris (2009) TBLT ‘even encourages a focus on form in view of optimizing the learning potential of task-based educational activities’ (p. 6).

Following Skehan (1998a; Foster & Skehan, 2013), we defined a focus on grammar within a communicative task, also known as *Focus on Form* (FonF), as learners directing their attentional resources to language in order to avoid error or to use more complex language. A focus on form, especially during meaning-based activities, has proven to promote both the efficiency and effectiveness of language acquisition processes (see review by Norris & Ortega, 2000).

Several frameworks for creating a task-based lesson have been proposed (Skehan, 1996a; Willis, 1996). These frameworks consist of a pre-task, during-task, and post-task. During the pre-task, the teacher explores the topic with the students, highlights useful words and phrases, and helps students understand task instructions to prepare for the task. In the during-task stage students prepare and carry out the task, whereas the post-task stage may be dedicated to a public task performance, reflection on the task performance or redoing the same or a similar task. In this PhD project we examined the effects of different focus on form strategies during each stage of the task-based teaching framework. The focus on form strategies can be considered instructional strategies in educational design. Throughout the book we will refer to them as FonF strategies.

Study 1 focused on the *pre-task stage*. In this study we examined the effects of having learners focus on either form or meaning during the observation of peer model videos on the accurate use of the German grammatical structure, dative case after a two-way preposition, and in addition on several measures of complexity. Study 2 focused on the *during-task stage*. In this study we examined the effects of two different feedback types, prompts and recasts, on the accurate use of two different German grammatical structures. These structures concerned more simple German comparative forms such as *toller* (nicer) and *größer* (bigger) and the more complex dative case after a two-way preposition such as *An der Wand hängt ein Poster* (On the wall hangs a poster).

Study 3 investigated a focus on form in the *post-task*. For this study we examined the effects of task repetition on the accurate use of the same grammatical structures as in Study 2, the comparatives and the dative case after a two-way preposition.

In addition, we investigated whether the effectiveness of the FonF strategies corrective feedback and task repetition depended on the grammatical structure involved. Furthermore, we examined whether learners’ focus on an accurate task performance had negative consequences for the fluency or complexity of that same performance. With this last question we investigated Skehan’s (1996a, 1998a) *Trade-off Hypothesis* which claims that learners have limited attentional resources. As a con-

sequence, learners' attention to either an accurate, complex or fluent task performance would entail less attention to the other dimensions and thus diminish the performance therein.

We investigated three major research questions:

- RQ1 Does the FonF strategy positively affect learners' declarative knowledge of the new grammatical structures and their oral task performance in terms of accuracy (Study 1, 2, 3) and complexity (Study1)?
- RQ2 Do the effects depend on the grammatical structure involved? (Studies 2, 3)
- RQ3 Does learners' focus on accuracy come at the expense of oral fluency (Study 1, 2, 3) or oral complexity (Study 1)?

2. ORGANIZATION OF THE THESIS

The rest of this thesis consists of five chapters and presents three studies. In chapter 2 we set up the theoretical framework on focus on form in TLBT which forms the theoretical basis for our three experimental studies. Chapters 3, 4, and 5 each address one experimental study. In chapter 6 we present a general discussion.

We will present and discuss the studies in the natural and logical order of the stages of the task-based framework, respectively the pre-task, during-task, and post-task stage. It must be mentioned though that the experiments were actually carried out in a different order: Study 2, 3 and 1. Chapter 3 answers the question on the effects of observational learning in the *pre-task* stage (Study 3). Chapter 4 answers the question on the effects of corrective feedback at the *during-task* stage (Study 1). Chapter 5 answers the question on the effects of task repetition at the *post-task* stage (Study 2). Chapters 3, 4 and 5 have been submitted as articles to international journals: one has been accepted for publication (chapter 4) and two (chapter 3,5) are currently under review*. The results of all three studies have been presented at several conferences including EARLI (Van de Guchte, Braaksma, Rijlaarsdam & Bimmel, 2013a), TBLT (Van de Guchte et al., 2013b, 2015b), and EUROSLA (Van de Guchte et al., 2014)

Chapter 2 presents the theoretical framework on focus on form in TBLT. Successively we review the shift from the traditional form-focused PPP method to more communicative approaches such as TBLT; the definition of tasks; the learning processes claimed to occur in TBLT; the implementation of FonF in the TBLT lesson; the effects of FonF strategies on different grammatical structures, and finally implicit and explicit measures of language learning.

In Chapter 3 we report on the study that investigated the effects of directing learners' attention to either form or meaning in the *pre-task* through guided observation of peer model videos. Two conditions were compared: one focused on form (FonF) during the observations, whereas the other focused on meaning (FonM).

* Writing a thesis in articles has both advantages and disadvantages. Clear advantage is that each chapter can be read separately. For this, first and subsequent citations are also formatted per chapter. A disadvantage is that there is sometimes overlap between chapters.

Think-aloud methods and a communicative oral task were used to measure the effects on both planning processes and task performance. Results of task performance showed that the FonF condition generated more (accurate) use of the target structure than the FonM condition, whereas the FonM condition outperformed the FonF condition on three complexity measures. Although we found a clear effect of the focus on task performance, this effect was not observed in the think aloud planning protocols.

In Chapter 4 we report the study that examined the effects of a focus on form *during-task* by comparing two types of corrective feedback. Students were randomly assigned to two experimental conditions: one receiving prompts, the other recasts. These experimental conditions were compared to a control condition, which was an intact class. The study involved two subsequent interventions: The first targeted a complex structure, dative case after a two-way preposition; the second a simple structure, comparatives. Pre-tests, immediate and delayed post-tests included written and oral accuracy as well as oral fluency. Statistical comparisons on both written and oral post-tests showed that prompts and recasts were both effective when compared to the control group, with prompts being superior to recasts. Furthermore, the effect of recasts depended on the structure: Recasts, as compared to prompts, were more effective for the comparative than for the dative on written accuracy.

In Chapter 5 we report the study that investigated the effects of task repetition in the *post-task* after having directed learners' attention to form during the main task. Students were randomly assigned to two conditions: One group repeated a similar task (R); the other group did not (NR). The study comprised two interventions: The first intervention targeted the German dative case after a two-way preposition; the second German comparatives. Pretests, immediate and delayed post-tests included metalinguistic knowledge, written and oral accuracy, as well as oral fluency. Results showed that the R-condition outperformed the NR-condition both on written accuracy and metalinguistic knowledge, for both structures. No statistical significant differences between conditions were found for oral accuracy.

In Chapter 6 we bring together our main findings for focus on form strategies in the three stages of the TBLT framework. In addition, we discuss the methodological decisions we made and provide suggestions for future research. Since the research questions of this thesis stem from instructional dilemmas in real task-based classrooms and the execution of the experimental studies found place in these same classrooms, we finally discuss the relevance of our findings for educational practice and possibilities for implementation.

CHAPTER 2

FOCUS ON FORM IN TASK-BASED LANGUAGE TEACHING

Chapter 2 presents the theoretical framework on focus on form in TBLT. Successively we review the shift from the traditional form-focused PPP method to more communicative approaches such as TBLT; the definition of tasks; the learning processes claimed to occur in TBLT; the implementation of FonF in the TBLT lesson; the effects of FonF strategies on different linguistic structures, and finally implicit and explicit measures of language learning.

1. FROM PPP TO TBLT

Before the introduction of task-based language teaching in the early 1980s (Ellis, 2003; Long, 1985; Nunan, 1989; Prabhu, 1987; Skehan, 1996a, 1998a; Willis, 1996), the presentation-practice-produce model (PPP) was the dominant method in second and foreign language teaching, research and practice around the world. The PPP model draws on both structuralism (linguistic theory) and behaviorism (learning theory) which claims that skills can be acquired through frequent practice (Hartley, 1998). The PPP model entails the following method: First, teachers present the linguistic structure to the learners (P1), then the learners practice with this structure by means of controlled exercises (drills) (P2), and finally learners try to use the structure in a more unrestricted way, for example through role playing (P3) (Criado, 2013). Although there has been a shift towards more communicative language teaching, more recently accelerated by the introduction of the Common European Framework of References for Languages (CEFR, Council of Europe, 2001), the PPP model is still common practice in many second and foreign language classrooms.

This begs the question, why the PPP model is considered to be so attractive. First, it is based on the claim that practice makes perfect (Criado, 2013) and the belief that sufficient practice with certain structures will lead to automatization of these structures. Second, Skehan (1996b) suggests that the PPP model provides teachers with a sense of power because the teacher is the expert in this model and thus in full control of what happens in the classroom. Third, it triggers feelings of security for learners because they know what procedures can be expected. These feelings are acknowledged to stimulate a positive attitude towards learning (Criado, 2013).

The PPP model, however, has been criticized by several advocates of task-based language teaching (TBLT). Lewis (1996) argues that PPP focuses too much on isolated linguistic structures without any reference to communicative use. Another point that has been made is that PPP emphasizes accuracy and correctness to such an

extent that it may prevent learners from experimenting with language (Willis, 1993). Finally, as Long and Crookes (1992) point out ‘people do not learn isolated items in the L2 one at a time, in additive, linear fashion, but as parts of complex mappings of groups of form-function relationships’ (p. 31). Moreover, this linear view on language learning disregards learners’ readiness (Pienemann, 1984, 1998) to learn certain language structures. Textbooks based on PPP may decree a certain order in teaching grammatical structures that learners are not ready to learn (Criado, 2013).

With the introduction of Communicative Language Teaching (CLT) in the late seventies, which placed functional language at the heart of language pedagogy, a new approach to language learning was promoted (see Van den Branden, Bygate & Norris, 2009). As Van den Branden et al. point out, the new CLT approach did not immediately cause teachers, curriculum designers and publishers of teaching materials to move away from the traditional PPP method. What happened was that text books such as *Deutsch Aktiv*, *Neue Kontakte* and *Stepping Stones* contained less form-focused exercises and in addition provided learners with more communicative opportunities to practice the use of linguistic structures in the last phase of the PPP model, the Production phase. In other words, this was a sort of PPP^{plus} method.

From the 1980s, Task Based Language Teaching (TBLT), with tasks as the central unit of teaching and instruction, was a logical next step based on Communicative Language Teaching (Skehan, 1998a, 1998b; Willis, 1996). TBLT is supported by constructivist learning theory in which learners ‘create meaning, learn by doing, and work collaboratively in mixed groups on common projects’ (Richards & Rodgers, 2001, p. 109). TBLT, though, did not produce a ready-made method that was the same for everyone. Rather, TBLT provided an approach to language teaching and learning with stronger and weaker manifestations (Richards & Rodgers, 2001). The distinction between strong and weak versions has been suggested by both Skehan (1998a) and Ellis (2003). They argue that in stronger interpretations of TBLT, the primary focus is on meaning, and the relatedness to authenticity and the real-world. In the weaker versions of TBLT, tasks are used as opportunities to practice the linguistic structures which have been learnt at an earlier stage. Skehan (2001) argues that in this version, tasks are used to ‘clothe’ the language structures (p. 1).

Several researchers (Ellis, 2009; Samuda & Bygate, 2008; Skehan, 2008a; Van den Branden et al., 2009) regard weak versions of TBLT as *task-supported* instead of *task-based* language learning. Müller-Hartmann and Schocker-von Ditfurth Schöningh (2011) disagree with this view, maintaining that tasks can still be the focus for teachers as part of a task-supported syllabus, as an alternative to established form-centered syllabi. Both Carless (2012) and Littlewood (2007) acknowledge the concerns in implementing strong meaning-focused versions of TBLT in foreign language teaching. They argue that teachers find it hard to control large groups of learners carrying out tasks. Furthermore, teachers comment that during planning and task performance learners tend to avoid the use of the target language. Finally, strong versions of TBLT may conflict with educational traditions (see Littlewood, 2007 on East-Asian foreign language contexts). According to Carless (2012) the solution to these concerns asks for a ‘non-doctrinaire’ approach to TBLT, which allows for a mixture of local approaches and main principles of TBLT. In this line of thinking, Littlewood (2004) proposed a continuum of commu-

nicativeness of activities (see Figure 1) in which both strong and weak versions of TBLT have their place. The left side of the continuum involves activities that focus more on the teaching of grammatical forms, whereas the right side demonstrates more ‘authentic’ communication focused on meaning. The three categories in between have features of both such as ‘communicative language practice’ or ‘structured communication’.

The three studies in this dissertation were situated in a TBLT-context, in which we used focused tasks that attempt to elicit the use of a particular grammatical structure but primarily focus on meaning.

Focus on forms		←	→	Focus on meaning	
Non-communicative learning	Pre-communicative language practice	Communicative language practice	Structured communication	Authentic communication	
Focusing on the structures of language, how they are formed and what they mean, e.g. substitution exercises, ‘discovery’ and awareness-raising activities	Practising language with some attention to meaning but not communicating new messages to others, e.g. ‘question-and-answer’ practice	Practising pre-taught language in a context where it communicates new information, e.g. information-gap activities or ‘personalized’ questions	Using language to communicate in situations which elicit pre-learned language, but with some unpredictability, e.g. structured role-play and simple problem-solving	Using language to communicate in situations where the meanings are unpredictable, e.g. creative role-play, more complex problem-solving and discussion	

Figure 1. Continuum from focus on form to focus on meaning (Littlewood, 2004, p.322).

2. DEFINING A TASK

Over the last 20 years, tasks have been defined from various standpoints and frames of reference. Long (1985) must have been one of the first who introduced the concept of a task.

A piece of work undertaken for oneself or for others, freely or for some reward. Thus, examples of tasks include painting a fence, dressing a child, filling out a form, buying a pair of shoes, making an airline reservation, borrowing a library book, taking a driving test, typing a letter, weighing a patient, sorting letters, making a hotel reservation, writing a check, finding a street destination and helping someone cross the road. In other words, by ‘task’ is meant the hundred and one things people do in everyday life, at work, at play and in between. (Long, 1985, p. 89)

Several researchers (see Ellis, 2003; Nunan, 1989, 2004; Richards, Platt & Weber, 1985) argued that not all of these tasks in Long’s definition demand language pro-

duction, for example, painting a fence. They therefore suggested that the successful completion of tasks should at least include language.

Another point then is, what kind of activities that require language use, constitute a task? Is the term task reserved only for activities that require learners to convey their message or do tasks also include activities that are aimed at accurate language use (Ellis, 2003)? Both Long (2015) and Skehan (1998a) argue that tasks should only involve meaning-focused activities. Breen's (1987) definition also allows the use of exercises: 'A range of work plans which have the overall purpose of facilitating language learning- from the simple and brief exercise type to more complex and lengthy activities such as group problem-solving or simulations and decision-making' (Breen, 1987, p. 23). In Ellis' (2003) definition, as well, a task is not restricted to a focus on meaning but also allows for a focus on language forms:

'A work plan that requires learners to process language pragmatically in order to achieve an outcome that can be evaluated in terms of whether the correct or appropriate propositional content has been conveyed. To this end, it requires them to give primary attention to meaning and to make use of their own linguistic resources, although the design of the task may predispose them to choose particular forms. A task is intended to result in language use that bears a resemblance direct or indirect to the way language is used in the real world. Like other language activities, a task can engage productive or receptive, and oral or written skills and also various cognitive processes' (Ellis, 2003, p. 16).

For example, tasks in Ellis's definition may predispose learners to choose particular forms, so called focused tasks. Ellis stresses however that focused tasks should not be confounded with exercises, because focused tasks primarily focus on meaning and exercises 'display correct use of the target feature' (Ellis, 2014, p. 107).

Differences in definitions of a task also refer to how 'real' tasks should be. In Long's definition, tasks are strongly connected to language that is used in the real world, also known as *target* tasks (Long 2015). Nunan (1989, 2004), approaching tasks from a pedagogical perspective, does not stress the relation to the real-world in his definition but focuses more on the pedagogical role of the task.

[A task is] a piece of classroom work which involves learners in comprehending, manipulating, producing, or interacting in the target language while their attention is focused on mobilizing their grammatical knowledge in order to express meaning, and in which the intention is to convey meaning rather than to manipulate form. The task should also have a sense of completeness, being able to stand alone as a communicative act in its own right. (Nunan, 2004, p.4)

Unlike Long, Nunan (2004) argues that not all tasks should involve real-world activities, because learners may experience semantic, pragmatic, lexical and syntactic difficulties. A connection between the task and the real-world would then be sufficient. For that reason, Nunan (2004) introduces the term *pedagogical* tasks which refer to tasks that occur in the second or foreign language classroom. These pedagogical tasks, may be derived from target tasks and are often simplified by breaking target tasks up into sub-tasks or include a pre-task that may provide learners with background information (Long, 2015).

Skehan's definition (1998a) only implicitly mentions the pedagogic role of the task (Norris, Brown, Hudson & Yoshioka, 1998) and focuses more on the cognitive processes that are involved in task performance. Based on several definitions of other researchers, Skehan (1998a) selected five essential characteristics of a task:

- Meaning is primary;
- There is some communication problem to solve;
- There is some sort of relationship to comparable real-world activities;
- Task completion has some priority;
- The assessment of the task is in terms of outcome (Skehan, 1998a, p. 95)

With this definition, Skehan focuses on the outcome of the task, another difference in how tasks are defined. Like Skehan, Willis's (1996) definition is also aimed at the achievement of a real outcome: 'A task is a goal-oriented activity in which learners use language to achieve a real outcome' (Willis, 1996, p. 53)

It remains unclear however, what exactly is meant by outcome. Skehan, for example, does not describe the outcome in terms of language. In the light of the idea that tasks should focus primarily on meaning and be clearly related to real-world activities, one could argue that it is not important how the learner handles the task, as long as the communicative goal of the task is achieved (Rodgers, 2009). In a similar vein, Widdowson (2003) argues that successful task completion may result in minimal language use, which does not lead to any learning. Ellis (2003) therefore suggests that a distinction should be made between the *outcome* and the *aim* of a task. The outcome then refers to what learners achieve at the end of the task, such as the completion of a family survey or family tree, whereas the aim of the task would refer to the elicitation of 'meaning-focused language use' (Ellis, 2003, p. 8).

Bygate, Skehan and Swain (2001) conclude that different definitions are needed for all the different purposes that tasks can have. To that extent they suggest a basic 'all-purpose' definition, that can be adapted for pedagogical and testing purposes: 'A task is an activity which requires learners to use language, with emphasis on meaning, to attain an objective' (Bygate et al, 2011, p. 11).

In all three studies in this dissertation we used tasks that 1) require an outcome in terms of content (meaning); 2) predispose learners to use particular grammatical structure; 3) are connected to real-world situations; and 4) engage learners in cognitive processing. Because Ellis' (2003) definition includes all four criteria, this will be the one that the studies in this dissertation will be based on.

3. LEARNING PROCESSES IN TASK-BASED LANGUAGE LEARNING

Proposals for task-based language teaching draw on a variety of claims about and research into the cognitive processes that may stimulate second and foreign language acquisition (Robinson & Gilabert, 2013). The cognitive processes that promote learning in TBLT will be reviewed in this section.

Tasks in foreign language acquisition (FLA) and second language acquisition (SLA) provide learners with all sorts of input. Initially it was Krashen (1981) who

stated that for learners to acquire implicit knowledge of a second or foreign language they do not need extensive rule explanations, but need to be exposed to *communicative and comprehensible input*. This claim is based on L1 acquisition theories and research which show that learners extract rules from the input without conscious learning. This comprehensible input would be best provided in 'low anxiety situations' and contain 'messages students really want to hear' (Krashen, 1981, p. 6). Furthermore, the input should not be too easy and not too complex. According to Krashen's Input Hypothesis (1981), learners should be exposed to a more complex input than their current level of language proficiency, also known as level 'i (input) plus 1'.

Besides the provision of input, the use of tasks can also be supported by theories of language output. Based on her observations in Canadian immersion classrooms, Swain (1985) argued that comprehensible input is not the only factor that could contribute to language acquisition. That is to say, learners in these immersion classrooms students were exposed to large amounts of comprehensible input. Although these learners became quite fluent in the second language, their progress on syntactical and morphological language proficiency, was lacking in comparison to native-speaking peers. From these findings Swain concluded that for acquisition to happen, learners should not only be provided with comprehensible input, but should also be 'forced' to produce language themselves (*see Output Hypothesis*, Swain, 1985, 2000, 2005). Swain reasoned that, when learners are given opportunities to speak or write in the second or foreign language, they may notice that they are not able to say what they want to say in the target language. Noticing this problem may push learners to modify their output and thus promote language learning.

Other researchers consider the interaction between learners as a crucial element of task-based language acquisition. For instance, when one of the participants in a conversation says something that the other does not understand, they may keep the conversation going by means of confirmation checks, clarification requests and comprehension checks (Pica, 1987). According to Long's (1996, 2015) *Interaction Hypothesis* it is this 'negotiation of meaning work that (...) facilitates acquisition because it connects input, internal learner capacities, particularly selective attention, and output in productive ways' (Long, 1996, pp. 451-452). Long's theory connects with both Krashen's Input and Swain's Output Hypothesis because the interaction between interlocutors (learner/learner or learner/teacher) may lead to comprehensible input and conversational feedback (Gass & McKey, 2006) and push learners to modify their output.

Other researchers justify the use of tasks from a cognitive or attention-driven perspective (Robinson, 2001; Schmidt, 1990; Skehan, 1998a; VanPatten, 1990). From this perspective, tasks not only give learners the opportunity to practice L2 speech but may also direct learners' attention to form-meaning relationships 'initiating learner analysis and restructuring of their interlanguage, improving their control of the language, and ultimately pushing the development of language knowledge and proficiency' (Norris, Bygate & Van den Branden, 2014). However, as Skehan (1998b) points out, during input processing and task performance, learners have limited attentional resources and must choose where to allocate their attention.

According to Skehan learners have three choices:

- 1) They pay attention to the content (meaning) of the task and focus on lexis, which may promote a more *fluent* performance;
- 2) They pay attention to form-control, which may result in a more *accurate* performance;
- 3) They take risks and experiment with new language, which may lead to a more *complex* performance. (Skehan, 1998b, pp. 269-270)

Drawing on theories on limited attention, VanPatten's (1990) study showed that, during the processing of instruction, learners prioritize meaning over form. This initial focus on meaning implies that a focus on form should be attracted (implicitly) or directed (explicitly) by either the task itself, the teacher or another student. This view connects with Schmidt's *Noticing Hypothesis* which claims that 'input does not become intake for language learning unless it is noticed, that is, consciously registered' (Schmidt, 2010, p. 1). In other words, language development can only be achieved when learners pay conscious attention to linguistic forms in the input, so that these can be processed in the short-term memory and incorporated into the learners' interlanguage (Schmidt, 1990).

Drawing on Schmidt's theories on noticing, Long (1991; Long & Robinson, 1998) introduced the term *focus on form* (FonF) as a reaction to the common practice of *focus on forms* (see Doughty & Williams, 1998). Unlike a focus on forms which entails the practice of linguistic features isolated from a communicative context, Long (1991) argued that learners should focus on linguistic elements 'as they arise incidentally in lessons whose overriding focus is on meaning or communication' (pp. 45-46). According to this theory, attention to form in both the input and/or during pushed output, would be beneficial to language development (Long, 2015).

To conclude, Skehan (2003) argues that almost all TBLT researchers, either from an interactionist or attention-driven perspective, agree on the generalizations that 1) interaction only is not sufficient to promote language learning; and 2) during interaction learners should, in some way, focus on form. However, there is, as yet, no consensus on how a focus on form should be implemented in TBLT instruction. The three studies in this dissertation (Chapter 3, 4, and 5) will be discussed in the light of SLA and FLA theories on input, output, noticing and attention, and finally limited attentional resources.

4. GRAMMAR IN TBLT

Several researchers (Sheen, 2003; Swan, 2005) claim that the role of grammar in task-based language teaching is often neglected (see Ellis, 2009a, 2014). Sheen (2003), for example, considers Long's focus on form a myth and argues that there is no grammar syllabus in TBLT since 'the treatment of grammar depends on unplanned problems in communication, arising during communicative activities' (p. 226). In the same line, Swan (2005) notes that TBLT 'bans' the use of a grammar syllabus, not realizing that it is very difficult for learners to shift from meaning to

form, when they have little knowledge about L2 grammar. Swan argues that ‘...focus on form is valuable once learners are ready to integrate the language elements they know into realistic communicative exchanges, but this will often need to be preceded by discrete presentation and practice’ (p. 394). Ellis (2009a, 2014) responds to this criticism by stating that, although TBLT may not have an explicit grammar syllabus, all supporters of TBLT see a role for grammar in terms of a focus on form during meaningful communication. However, how researchers interpret the term focus on form and how this can be implemented in TBLT instruction differs. In this section we will first review different approaches to focus on form (FonF), based on researchers’ different views on language learning. Then, we will discuss what these approaches mean for the implementation of FonF in task-based instruction.

4.1 Long: Reactive and Brief FonF

Michael Long (1991) was the first researcher who introduced the term focus on form. Evaluating both synthetic and analytic approaches to language learning, he concluded that synthetic approaches focused too much on forms, while analytic procedures were overly focused on meaning. Whereas analytic approaches discarded the potential of instruction, synthetic approaches assumed that once learners were exposed to a certain structure, sufficient practice would help them automatize it (Long, 2015). Based on the limitations of these two approaches, Long (1991; Robinson & Long, 1998) introduced an analytic approach that included a *focus on form*. According to his definition focus on form involves: ‘...an occasional shift of attention to linguistic code features by the teacher and/or one or more students triggered by perceived problems with comprehension or production’ (Long & Robinson, 1998, p. 31). In a recent definition, Long (2015) stresses that acquisition cannot be based solely on implicit learning from exposure to input, but also relies on explicit learning, through attention to grammar and lexis:

[The] *reactive* use of a wide variety of pedagogic procedures (PPs) to draw learners’ attention to linguistic problems in context, as they arise during communication..., thereby increasing the likelihood that attention to code features will be synchronized with the learner’s internal syllabus, developmental stage and processing ability. Focus on form capitalizes on a symbiotic relationship between explicit and implicit learning, instruction and knowledge. (Long, 2015, p. 27).

By not falling into the trap of focusing on explicit rule explanations and drill exercises, he suggests that during meaningful communication, learners’ attention should only be *briefly* drawn to form, and only *as a reaction* to a learner experiencing a problem (Long, 2015). Long considers focusing on form while learners experience a communicative problem as optimal because of learners’ ‘underlying psychological state’ (Long, 2015, p. 27). He explains that the fact that learners are producing the target structure could imply that they are developmentally ready for instruction (Long, 2015). Because of Long’s (1991) reactive stance, excluding any isolated teaching and practicing of forms, Doughty and Williams (1998) placed Long ‘at the most implicit end of the FonF continuum’ (p.5). As shown in Doughty and Williams (1998) volume on FonF, other researchers adopted broader interpretations of FonF,

that also include teaching and practice of predetermined structures through the use of communicative activities.

4.2 *DeKeyser and Lightbown: Explicit knowledge turns into implicit knowledge*

In contrast to Long, DeKeyser and Lightbown are at the most explicit end of the FonF continuum (Doughty & Williams, 1998). They even see a role for Long's 'dreaded' focus on forms, as long as there is a connection between form and meaning. DeKeyser's (2007, 2010) view on what kind of focus on form should be used in TBLT, draws on Anderson's (1993, 2000), skill-learning theories. In skill acquisition theories second language acquisition is considered the same as learning any other skill such as playing tennis or the piano. Learners start with declarative knowledge about a particular linguistic structure (describe the rule) and gradually, through practice, this knowledge turns into procedural knowledge, when learners apply this rule in real language use. In this regard, tasks are considered a means to turn declarative knowledge into procedural knowledge. DeKeyser (2007) argues that by engaging learners in communicative activities, they practice the use of language, while keeping 'the relevant declarative knowledge in working memory' (p. 52). In a later publication DeKeyser (2010) underlines that he does not refer to practice in terms of drills but to a broader concept of practice which involves guidance of meaning-form relationships. Citing Lightbown (2000, p. 443), he argues that '[w]hen practice is defined as opportunities for meaningful language use (both receptive and productive) and for thoughtful, effortful practice of difficult linguistic features, then the role of practice is clearly beneficial and even essential'. Furthermore, he suggests that role-play activities, task-based learning, and content-based learning lend themselves for systematic practice that combines meaning and form (DeKeyser, 2010).

With this view on language learning, DeKeyser advocates a strong-interface position which states that the instruction of explicit knowledge may contribute to SLA, in the sense that explicit knowledge is the start for what may become implicit knowledge (see also DeKeyser, 1995). This view is in contrast with the non-interface position which claims that explicit (declarative) and implicit (procedural) knowledge are independent and that there is no role for instruction in SLA (Krashen, 1985; Prabhu, 1987).

4.3 *Ellis: Explicit rule knowledge may facilitate implicit knowledge*

Drawing on Schmidt's (1990) Noticing Hypothesis, Ellis (2003) takes a different stance, also known as the *weak-interface* position. Ellis claims that implicit knowledge may not be directly developed through the instruction of explicit knowledge but may be facilitated by explicit knowledge. He explains that explicit knowledge about a linguistic structure may help learners notice the structure in the input they are exposed to and/or may help learners to notice-the-gap in the output they produce. According to Ellis, a focus on form can be achieved through either 1) the use of focused tasks, which have been designed to focus learners' attention to a

predetermined linguistic structure or 2) the implementation of methodologies which direct learners' attention to form while performing unfocused tasks which are not designed with a certain linguistic structure in mind (Ellis, 2003, 2009a).

4.4 Skehan: Rule-based and exemplar-based knowledge

Instead of dividing L2 knowledge in explicit/declarative or implicit/procedural knowledge, Skehan (1998a) proposed a dual coding system which distinguishes learners' knowledge of a language into exemplars (chunks-based language and idiom) and rule knowledge. In the former case, what learners have learned derives from language rules such as 'in German, a masculine noun changes into *'dem'* after a preposition that rules the dative case. In the latter case, learners use the dative form which they have encountered in the input, such as *'auf dem Tisch liegt ein Kugelschreiber [on the table lies a pen]'*. Skehan underlines that the two systems are not separate but coexist. He explains that learners can switch between these systems depending on their processing needs. There is also the possibility that structures that have been acquired through a rule become chunks and vice versa that acquired chunks may be analyzed at a later stadium (Skehan, 1996a, 1998a).

Regarding a focus on form, Skehan (1998a) claims that this is best achieved through the rule-based system 'since the precision and system which accounts for rule-organization will make the feedback more informative' (p. 88). Moreover, such a focus may restructure learners' L2 knowledge because the rules learners have learnt may be applied in different situations. According to Skehan (1998a) the exemplar-based system is not very accessible for a focus on form because 'exemplars are not part of a structured system...' (p. 89). In contrast, becoming more fluent in the L2 can be best obtained through the use of the exemplar-based system because this system gives learners direct access to ready-made chunks (Skehan, 1998a).

4.5 Willis: Focus on form not essential condition for learning

Willis (1996) allows for a focus on form in TBLT but does not consider it 'an essential condition for learning' (p. 16). In her view, tasks are never designed with a specific form in mind, because learners should be free to choose their own language structures and experiment with language. Willis reasons that during task performance the focus is on fluency and only after the performance, accuracy comes into play. After the task performance, learners may be engaged in consciousness raising activities that process specific language features and even practice activities that follow the analysis of language (Willis, 1996; Willis & Willis, 1996). Skehan (1998a) considers Willis's proposed focus on form activities very useful but criticizes the fact that the activities are not connected to theories on noticing or cognitive processes.

5. IMPLEMENTING FONF IN TASK-BASED INSTRUCTION

Two major ways of incorporating a focus on form in task-based language teaching can be distinguished: 1) by means of *focused tasks* which attempt to elicit the use of predetermined language structures; and 2) through the use of *methodological procedures* that focus on form during task performance of unfocused tasks (Ellis, 2003).

5.1 *FonF through the use of focused tasks*

Focused tasks are designed to make learners process specific linguistic structures which are ‘natural’, ‘useful’, or ‘necessary’ for successful task performance (see key article by Loschky & Bley-Vroman, 1993). In other words, the focus on form in such tasks is deliberate (Ellis et al., 2002). For example, a learner of German as a foreign language who needs to compare the locations of several Christmas presents in a spot-the-difference-task, makes it easier to conduct the task successfully when he uses prepositions such as *unter* [under], *auf* [on], *in* [in] (see example in Nunan, 2004, p. 96). In a way, focused tasks could be compared to a *focus on forms* because the linguistic structure is also predetermined. However, Ellis (2009a; Ellis & Shin-tani, 2013) stresses that there are two differences between focused tasks and a traditional focus on forms. First, focused tasks should adhere to the general criteria for tasks: 1) primary focus on meaning; 2) there is some kind of knowledge gap; 3) learners have to rely largely on their own resources; 4) there is a clearly defined outcome, other than the use of language. Second, learners are not explicitly told what the target structure of the task is.

Samuda (2001) provides an interesting example of a focused task that strikes a balance between form and meaning. For the task ‘Things in Pockets’ students were told the objects that came out of the pocket of a lost overcoat, and were asked to consider who the owner of the objects was. In a task chart, learners had to describe how certain they were (was it ‘possible’, ‘probable’ or ‘certain’?) about it. The task did not dictate the use of modal auxiliaries, but the use of them was encouraged through both task design and the teacher who shifted in attentional focus from meaning to form to meaning again, as the task unfolded.

A planned focus on form can be effective because it gives learners the opportunity to practice the same form repeatedly during meaningful communication (Ellis, Basturkmen & Loewen, 2002). Besides the practicing-claim which is based on skill-building theories, focused tasks also provide opportunities to communicate that may facilitate the learning of grammatical structures implicitly (Ellis, 2003).

Nevertheless, it should be noted that not all researchers are in favor of using focused tasks. Willis and Willis (2001), for example, argue that tasks that elicit certain language features are unnatural. They state that learners should be free to choose their own language structures for task completion and that these should not be planned by the teacher in advance. Similarly, Skehan (1998b) considers it ‘impossible to try to finesse the use of any particular structure in a task’ (p.279) and finds it rather optimistic to think that ‘structures can be reliably trapped inside a range of tasks without compromising the primacy of meaning’ (Skehan, 1998a, p.122). Skehan therefore suggests an approach that perpetuates the naturalness of the tasks

but, at the same time, offers systematic instruction that maximizes the chances for effective use of attention (see Skehan's framework for the implementation of task-based instruction, 1996a, 1998a).

5.2 *Implementing FonF methodologically*

Besides the use of focused tasks, a focus on form can also be implemented methodologically. These methods will be discussed in terms of the activities that can be conducted in the different stages of the TBLT lesson. In almost all proposals, the task has three stages, namely a pre-stage, a during-task stage, and a post-stage, introduced by Prabhu in 1987. In general, during the pre-task learners perform activities that prepare them for subsequent task performance. The during-task option is centered around learners' communicative task performance. Finally, the post-task includes activities that follow-up on the previous task performance (Ellis, 2003). Table 1 (adapted from Ellis, 2014, pp. 105-106) demonstrates the different positions of researchers regarding the place of FonF in the TBLT framework. We will discuss these positions in this section.

Table 1. Different approaches to the place of FonF in the TBLT framework

Willis (1996)	Long (1985; 1991; 2015)	Skehan (1998a; 2011)	Ellis (2003)
In the pre-task and post-task phases but not in the main task phase	In the main-task phase	Mainly in the pre-task phase	In all phases of a lesson

5.2.1 *FonF in the pre-task stage*

In Willis' (1996) model the teacher uses the pre-task to outline the task, explore the topic and highlight useful words and phrases. Although the introduction of language is not aimed at teaching particular grammatical structures, activating useful words and phrases can function as opportunities for a 'focus on form to be set in motion, and for noticing to occur' (Skehan, 1998a, p. 127). Activities that could stimulate learners' awareness of linguistic structures could include 'reading and talking about texts and picking out words and phrases' that may be highlighted at the post-task stage, 'teacher demonstration of the task with a good student' or 'audio or video recordings of fluent speakers doing the task' (for more examples see Willis 1996, pp. 43-45).

Ellis (2003, 2009a, 2014) argues that FonF can figure in all three phases of the TBLT framework. For the pre-task he suggests some of the same activities Willis does, but points out that there is a risk to providing models. After having seen the model, learners may approach the subsequent task as an exercise to practice the tar-

get structures they have noticed in the modeled task (Ellis, 2003). Ellis therefore suggests that teachers should be aware of the extent to which students are 'primed to attend to specific aspects of the model' (p. 246).

The last pre-task option, providing learners with time to plan their task, is suggested by Skehan (1996b; 1998a), Ellis (2003) and Willis (1996). Skehan (1996b) argues that planning time may reduce learners' cognitive load because it releases them from 'having to think of too many things at a time when both composing thoughts and producing speech' (p. 25). This freed-up attention may then be used by learners to focus their attention on form, resulting in a more accurate performance (Ellis, 1987; Foster & Skehan, 1996). Besides gains in accuracy, planning time may also lead to development in fluency or complexity (Crookes, 1989; Foster & Skehan, 1996). Planning time may be provided in two different ways, either guided or unguided. In the first manner learners are free to choose where to pay their attention to (Ortega, 1999) whereas in the latter, learners' attention is allocated to certain aspects of the task performance, for example to either form or content (Foster & Skehan, 1999) or to both form and content (Sangarun, 2005).

Finally, some remarks should be made regarding the effectiveness of FonF in the pre-task. Ellis (2003) notes that pre-task activities 'can help to create conditions that will make tasks work for acquisition', but cannot, 'fine-tune learners performance of a task...' (p.249). In the same vein, Skehan (1996a) argues that pre-task work creates *conditions* that may lead to learning, but may not *immediately* lead to learning. Skehan elaborates that we cannot be sure that the language, that learners were directed to in the pre-task, will occur during subsequent task performance.

5.2.2 *FonF in the during-task stage*

In Willis' (1996) model the during-task stage is called the *task cycle* in which learners first conduct the task and then go through a planning and reporting stage. Willis has strict ideas regarding a focus on form during task performance. She states that during task performance learners should focus on successful task completion and fluency development, not on accuracy (Willis, 1996). Moreover, she argues that learners are expected to use their own linguistic resources; the teacher only provides assistance when needed and does not focus on lexical or grammatical accuracy. In the second part of the task cycle learners plan to report on how the task went and finally report this to the class by means of a written or oral presentation. Willis and Willis (1996) point out that during the planning phase, before the public presentation, learners may also attend to accuracy by checking words and phrases they are not sure of.

Samuda (2001) comments that it may be a risk that Willis's model only allows for a focus on grammar, when the task has already been performed. She notes that 'certain aspects of certain forms may escape focus altogether if not in some way highlighted in task input, or nudged in task performance' (p. 122).

As opposed to Willis, Long (2007) regards focus on form primarily as a during-task option. Long reasons that such a FonF will be particularly effective because the linguistic focus is then connected to a meaning problem. In this way, new form-

meaning relationships are created; not by means of abstract language rules but through incidental learning which will be ‘perceived more quickly’ (Long, 2015, p. 317). Long (2007, 2015) suggests that providing learners with negative corrective feedback, in the form of recasts, is an ideal way of briefly directing learners’ attention to form. In the case of recasts, the interlocutor will repeat the error back to the learner in a corrected form. Recasts fit easily in Long’s conceptualization of focus on form because the teacher reacts ‘on the spot’ to the learner’s error in the target language, while keeping the focus on the content.

In addition to Long’s *conversational* FonF, Ellis et al. (2002) argue that learners may also be corrected when there is no problem in communication. The teacher may understand what the learner says, but still wants to focus learner’s attention to the error by correcting it. Ellis et al. (2002) call this switch in attention a ‘*didactic*’ FonF and suggest that it is aimed more at ‘negotiation of form’ than at ‘negotiation of meaning’ (p. 425). During these moments of negotiation of form learners are provided with feedback, which may vary from more implicit forms such as requests for clarification or recasts to more explicit forms. Explicit feedback can take the form of explicit correction in which learners are told, for example, ‘not gooder, but better’. Learners may also be exposed to metalinguistic information ‘you should use the dative case here, not the accusative case’. In addition, the teacher could attempt to elicit the answer again in the hope that the student will do it right, for example by saying ‘let’s try again’ (for more examples of corrective feedback see Lyster and Ranta, 1997).

According to Ellis et al. (2002), the advantage of more explicit forms of feedback is that linguistic target structures are more likely to be noticed. In the case of recasts, learners may not always be aware of their errors because they are not told that they have made a mistake or what the mistake was. Meta-analyses by Li (2010) and Lyster and Saito (2013) show that both implicit and explicit feedback are beneficial to SLA. In addition, Lyster and Saito (2013) found that the effects were larger for the more explicit prompts than the implicit recasts. Moreover, Li (2010) showed that the effect of implicit feedback was better maintained over time than that of explicit feedback.

5.2.3 *FonF in the post-task stage*

Based on the assumption that accuracy is acquired after fluency, Willis (1996) proposes that the post-task phase is the most suitable place for a focus on language form and use. Willis reasons that learners have already processed language for meaning in the task cycle and may then be encouraged to focus on certain structures in the post-task in order to promote language learning. The activities learners may carry out, involve both analysis and practice activities, such as reflection on linguistic structures, recycling of the task language, rechecking their text or audio or video or audio recording, or practicing useful phrases (Willis & Willis, 1996). According to Ellis (2003, pp. 258-262) there are three options in the post-task:

- 1) learners can repeat the same or a similar task. Research has shown that task repetition may improve learners' oral performance in terms of both accuracy, complexity and fluency (Ahmadian & Tavakoli, 2011; Bygate, 1996, 2001; Gass, MacKey, Alvarez-Torres, & Fernandez-Garcia, 1999; Lynch & McLean, 2000, 2001);
- 2) learners may reflect on how they performed their task; or
- 3) learners may focus on linguistic forms that appeared to be difficult during the main task performance by reviewing errors, performing consciousness raising tasks, or production-practice and noticing activities.

In addition to these options which promote consolidation and reflection, Skehan (1996a, 1998a) suggests that the use of post-tasks might change the way a task is conducted and how attention is allocated during the task. For example, when learners are aware of the fact that they are required to present the outcome of their task to a public (Skehan & Foster, 1997) or transcribe their own previous task performance (Foster & Skehan, 2013; Qian, 2014) they may not only focus on a fluent but also a more accurate performance, and may shift their attention from meaning to form in anticipation of a post-task activity.

5.2.4 *Implementation of FonF in the studies*

Doughty and Williams (1998) point out that it is perfectly possible to 'combine explicit and implicit FonF techniques, depending upon the particular acquisition circumstances' (p. 261). In line with this view, for the studies in this dissertation we used both design and methodological procedures to include a focus on form. First, the tasks were designed to elicit the use of predetermined grammatical structures because we wanted to increase the chance that the students would use the structures. Second, attention to these structures was drawn by means of implicit as well as more explicit FonF strategies.

In Study 1 learners were implicitly attracted to the target structures, while observing a video of other students performing a similar task. In Study 2, attention to the target structures was directed by different forms of corrective feedback (recasts and prompts). Finally, in Study 3 learners repeated a similar task, preceded by a focus on form in the during-task.

6. DIFFERENT FONF STRATEGIES FOR DIFFERENT STRUCTURES?

In the field of SLA it is commonly accepted that not all grammatical structures are acquired in the same manner (Larsen-Freeman, 1995). For that reason, Doughty and Williams (1998) argue that a focus on form 'should not be applied to all forms in the same way' (p. 211). This raises the question which FonF strategy should be used for which structure. To determine which form of instruction would be most effective for which structure, researchers started by making a distinction between easy and more complex structures. However, how can teachers know which grammatical structures cause their students more difficulties than others? To this extent, several criteria

have been proposed which vary from the complexity, scope, and reliability of the target structure to the developmental readiness of the learner and the contrast between L1 and L2.

First, there is the formal and functional *complexity* of the structure. Structures that involve several elements are supposed to be more complex than structures that involve only one element such as the plural-s (Ellis, 2003). A structure is functionally complex when the relation between form and function is opaque and less complex when this function is transparent (DeKeyser, 1998). Regarding complexity, Krashen (1982, 1985) claims that complex rules can only be learnt unconsciously (implicitly), whereas simple rules are open to explicit instruction. De Graaff (1997) operationalized complexity as the number of grammatical concepts that have to be taken into account in order to arrive at a correct form in language production and hypothesized that ‘when more criteria have to be applied, explicit instruction is more effective, as spontaneous noticing and processing would then be less likely.’ (p. 251).

Next to complexity, Hulstijn and De Graaff (1994) mention *scope* and *reliability* as factors that influence the effectiveness of instruction on grammatical structures. They suggest that rules of large scope and high reliability can be taught but rules that are unreliable or small in scope need not be taught. Either the learner cannot rely on applying the rule or the effort required to learn the rule is not justified. In reference to Skehan’s dual knowledge system they suggest that, ‘apart from memorization of fixed chunks, the acquisition of syntactic structures depends more heavily on abstract rule-based learning than the acquisition of morphological structure, for which exemplar based item learning is much likely to occur’ (p. 252).

Developmental readiness refers to the question, whether a learner is ready to process certain linguistic structures. According to Pienemann’s (1984,1998) Learnability (Teachability) theory, the acquisition of grammatical structures follows developmental sequences in ways which cannot easily be manipulated. According to this theory, teaching a grammatical structure will only be effective when the learner is developmentally ready for it.

Regarding *L1-L2 contrasts*, Ellis (2003) claims that L2 structures that correlate with the L1 are easy to acquire, whereas structures that do not have this correlation are more difficult. De Bot (1996) suggests that when learners learn a second or foreign language, they need to figure out which L1 procedures can be applied in the language to be learned. He explains that ‘some L1 rules can be used in L2 as they are; others may need to be adapted; some L2 rules are so different from anything in L1 that they have to be learned from scratch’ (pp. 548-549). De Bot argues that in all cases learners must acquire new declarative knowledge that can be turned into fast and automatic procedures.

To conclude, these criteria are important for two of the studies (2 and 3) in this dissertation because in these studies we investigated whether the effectiveness of the focus on form strategy depended on the grammatical structure involved. The designed tasks targeted either German comparatives or the dative case after a two-way preposition. The first we considered a simple structure because of the relatedness to the L1 and because most of the comparatives are formed by simply adding *-er* to the adjective or adverb. Besides rule-based elements the comparatives also include exemplar-based elements. The latter, we considered a more complex structure because

the use of the German dative case after a two-way preposition has no similarities to the L1. In addition, learners have to take several grammatical concepts into account in order to arrive at a correct form in language production.

7. MEASURING L2 KNOWLEDGE

Since it is assumed that language learning involves both implicit and explicit knowledge (Rebuschat, 2013), these two knowledge types should be reflected in measurements of learners' L2 knowledge. For that reason, in the studies on which this dissertation reports, both measures of implicit and explicit knowledge were used.

7.1 *Measuring learners' implicit knowledge*

Implicit knowledge is defined by Ellis (2003) as 'knowledge of language that a speaker manifests in performance but has no awareness of' (p. 105). To measure learners' implicit knowledge, Ellis (2005a) suggested, among other tests, the use of spontaneous production tasks, with the risk, that one cannot be totally certain that the student has no access at all to explicit knowledge.

In the three studies in this dissertation, learners' implicit knowledge was measured by means of meaning-based oral production tasks. These oral tasks were designed as focused tasks, which are often used in research because they provide a means of measuring whether learners have acquired a specific feature (Ellis, 2003; Doughty & Varela, 1998; Fotos & Ellis, 1991; Loschky & Bley Vroman, 1993). To measure learners' oral language production, Skehan (1996a, 1998a) proposed three indicators, complexity, accuracy and fluency, which are based on his dual-mode system and his theories on learners' limited attentional resources.

To assess oral *complexity*, several measurements have been suggested, varying from more general measures such as mean length of a chosen production unit, the amount of subordinated or coordinated sentences in a chosen production unit (Bulté & Housen, 2012) to more specific measures such as the frequency of use of linguistic structures (Ellis, 2003). In Study 1 of the dissertation, complexity was operationalized by means of 1) the use of the target structure (dative case of the article after a two-way preposition) per clause; 2) general complexity, defined as the total number of clauses per Analysis of Speech Unit (for AS-unit see Foster, Tonkyn & Wigglesworth, 2000); 3) complexity by subordination (number of subordinated clauses per AS-unit); and 4) complexity by coordination (number of coordinated clauses per AS-unit).

Oral *accuracy*, relates to a 'performance which is native-like through its rule governed nature' (Skehan, 1996a, p. 46). Some researchers used more global measures to assess accuracy, such as percentage of error-free clauses (Skehan & Foster, 1999) or error-free T-units, which is a dominant clause and its dependent clauses. Others measured accuracy in terms of target-like use of a particular grammatical structure, such as the article system or verb forms (Crookes, 1989; Juan & Ellis, 2003). In all three studies in this dissertation, learners' oral accuracy was measured

in terms of target-like use of the German dative case after a two-way preposition and target-like use of German comparatives.

Regarding *fluency*, Tavakoli and Skehan (2005) identified three key constructs of oral fluency in oral production: breakdown fluency (measured by silent and filled pausing), speed fluency (measured by articulation rate), and repair fluency (measured by number of false starts and repetitions). As De Jong and Wempe (2009) point out, in most speaking tests fluency is assessed by human raters who very likely use all constructs of fluency to judge the performance. De Jong et al. (2013) argue that ‘strong associations between utterance fluency and perceived fluency have been found’ (p. 895), regardless of the type of rater (e.g. L2 teacher or expert). It appeared that some objective measures of pausing and some measures of speech correlated with measures of perceived fluency by human raters. In addition, De Jong et al. (2013) explain that the aspect of fluency human raters judge, depends strongly on how these raters are instructed. For example, raters who are told to focus their assessment on speech rate and pausing, will pay attention to these aspects. Raters who are not instructed at all, tend to constitute their own definition of fluency. In contrast to human-rated fluency, De Jong and Wempe (2009) wrote a script in the software program PRAAT with which they could automatically measure speech rate, by counting syllables in speech samples. In Study 1 and 3 of this dissertation we used De Jong and Wempe’s written script to measure fluency, whereas in Study 2 we used human raters to assess learners’ fluency.

7.2 *Measuring learners’ explicit knowledge*

Explicit L2 knowledge is ‘that knowledge of rules and items that exist in an analyzed form so that learners are able to report what they know’ (Ellis, 1994, p. 702). Ellis’ (2005a) study demonstrated that a metalinguistic knowledge test and an untimed grammaticality judgment task are valid measures for tapping learners’ explicit knowledge.

Drawing on Ellis’s study, learners’ explicit knowledge was measured in the studies on which this dissertation reports by means of a meta-linguistic knowledge test. Considering that ‘having learners verbalize rules provides a rather conservative picture of what they know explicitly’ (Ellis, 2005a, p. 147), we added a second test to measure learners’ explicit knowledge by asking them to use these structures in a fill-in-the-gap exercise. In line with DeKeyser (1993), we opted for a fill-in-the-gap test because a) this format was most familiar to the students, and b) it forced them to provide a correct alternative rather than simply indicating the occurrence of an incorrect structure, which is a common limitation of many grammaticality judgment tests.

7.3 *Complexity, accuracy, fluency in competition?*

Regarding the three measures of oral production, complexity, accuracy, and fluency, Skehan (1996a, 1998a, 2009) claims that attentional capacity and working memory are limited. He argues that it is therefore not possible to achieve these aspects of

performance simultaneously. Learners must prioritize their attention to being accurate, fluent, or complex. According to Skehan's (1998a) Trade-off Hypothesis, attention to any one of these aspects of oral task performance will lead to less attention for the other aspects and may result in a diminished performance on those measures. In contrast, Robinson (2011) does not support the idea that accuracy and complexity are in direct competition with each other; moreover he argues that 'on some dimensions of task demands [...] increasing complexity [...] promotes more accurate, grammaticized production *and* more complex, syntacticized utterances' (p. 14). VanPatten (1990, 1996, 2004) affirmed Skehan's Trade-off Theory by demonstrating that learners prioritize meaning (fluency) over form (accuracy) during input processing.

In all three studies, we aimed to investigate Skehan's Trade-off Hypothesis. In Study 1, we examined trade-offs between learners' oral complexity, accuracy, and fluency. In Studies 2 and 3, we investigated trade-offs between accuracy and fluency.

8. CONCLUSION

In this chapter we discussed the shift from the traditional form-focused PPP method to more communicative approaches such as TBLT, the definition of tasks, the claimed learning processes in TBLT, the implementation of FonF in the TBLT lesson, the effects of instruction on different linguistic structures, and finally implicit and explicit measures of language learning. All of the elements discussed are important for the three experimental studies on which this dissertation reports.

First, the proposal for this dissertation stems from a discussion in the school where this research was conducted. There is a group of teachers whose overriding focus in the second and foreign language lessons is on the teaching of grammatical structures, and there are teachers whose lessons are primarily aimed at meaning (including a focus on form). Although a more communicative approach to language learning has been available since the introduction of the Common European Framework of References for Languages (CEFR, Council of Europe, 2001), the form-meaning discussion has still not been resolved in this school.

Second, we explored the definition of a task to support our decision to choose Ellis' (2003) definition as the one that would inform our research. Ellis' definition appeared most appropriate because it includes a focus on both meaning and form, authentic communication with a connection to the real world, linguistic outcome as an educational goal, and the involvement of cognitive processes.

Third, based on SLA and FLA theories, we used oral tasks that would trigger learning processes that are claimed to support language acquisition. To this extent, the tasks in our research provided learners with rich and comprehensible *input* in the pre-task and gave them opportunities for *output* in the during- and post-task stage. In addition, learners were directed to *form* through the use of focused tasks and different FonF strategies.

Fourth, we reviewed the implementation of FonF in the task-based lesson. Based on Skehan's (2006a) framework we designed tasks that involved three stages: a pre-

task, during-task, and post-task stage. Each study in this dissertation investigates a FonF strategy at a different stage, based on earlier FonF research and theory.

Fifth, in Study 2 and 3 we investigated whether the effectiveness of feedback and task repetition depended on the grammatical structure involved. To this extent we used tasks that targeted either the German comparative or the German dative case after a two-way preposition. Based on L1-L2 relatedness and the required number of cognitive steps, we regarded the comparative structure as simple and the dative structure as more complex for our Dutch participants.

Sixth, by using different FonF strategies we aimed to promote both implicit and explicit learning. We therefore used measures that enabled us to assess the extent of learners' implicit and explicit knowledge.

Finally, with these three studies we wanted to contribute to the literature on learners' limited attentional resources by investigating potential trade-offs between complexity, accuracy and fluency which were proposed in earlier research.

CHAPTER 3

FOCUSING ON MEANING AND FORM IN PRE-TASK VIDEO MODEL OBSERVATIONS: EFFECTS ON PLANNING PROCESSES AND TASK PERFORMANCE

Over the last decades several researchers have investigated the effects of focusing on form in the pre-task stage in task-based language teaching (TBLT). However, depending on the purposes of the lesson, the pre-task may not only be used to make learners focus on form but also on more complex interpretations of the task. The observation of models is a strategy that can be used effectively to direct learners' attention to either form or content in the pre-task. Therefore, the current study investigates the effects of directing learners' attention to either form or meaning in the pre-task through guided observation of peer model videos. Forty-eight ninth-grade students learning German as a foreign language were randomly assigned to two conditions: A focus on form (FonF) and a focus on meaning (FonM) condition. Think-aloud methods and a communicative oral task were used to measure the effects on planning processes and task performance. Results of task performance showed that the FonF condition generated more (accurate) use of the target structure than the FonM condition, whereas the FonM condition outperformed the FonF condition on three complexity measures. Although we thus found a clear effect of the focus on task performance, this effect was not observed in the think aloud planning protocols.

1. INTRODUCTION

Over the last decades, several researchers have investigated strategies that could effectively be used to direct learner's attention to form at the pre-task stage of the task based teaching cycle (Doughty & Williams, 1989; Ellis, Basturkmen & Loewen, 2002). Some of the focus on form research stems from the critique that task-based language teaching overemphasizes meaning and that grammar should play a more important role (see Ellis, 2009a; Sheen, 2003; Swan, 2005).

Most studies in focus on form research are based on theories on limited attentional resources (Skehan, 1996a, 1998a) which claim that when learners need to comprehend or produce language, they cannot pay attention to both content and form at the same time. Moreover, the VanPatten's (1990, 1996) studies revealed that when learners choose between meaning and form, they prioritize meaning. The findings of these studies imply that learners will not immediately focus on form by themselves, and that their attention needs to be directed to form by either task or teacher instructions. Strategies such as guided planning (Foster & Skehan, 1999; Sangarun, 2005), and modeling (Kim, 2013; Kim & McDonough, 2011) have been used to make learners focus on form during the pre-task. These strategies appeared, to a greater or lesser extent, successful in making linguistic structures salient and noticed (Schmidt, 1990) in order to promote language development.

However, the pre-task offers more possibilities than solely directing learners' attention to the use of certain linguistic structures (Skehan, 1996a). Besides focusing on form, the pre-task may also encourage learners to think actively about the content of the task and promote more complex, risk-taking language behavior (Skehan, 1998a). To be able to compare the effects of a pre-task in which learners focus on form or on content, we needed to design a pre-task strategy that could include two different foci. Guided observational learning from peer models was chosen as pre-task strategy. Peer model performances could provide rich input to the students and additional written, reflective questions on observation sheets could direct the students' attention to either meaning or form, while observing peer model videos.

Besides measuring the effects on task performance, we also investigated whether a pre-task focus on either form or meaning would be reflected in the learners' planning processes. In short, the current study investigates the effects of focusing on form or meaning in pre-task peer model video observations on both planning processes and three dimensions of task performance: accuracy, complexity and fluency.

1.1 Observational learning

In this section we will discuss how observational learning (i.e., learning from models) is related to language acquisition processes and which factors contribute to the effectiveness of observational learning. Furthermore, we will review few studies that examined the effects of focusing on form in pre-task modeling on subsequent task performance. To date, however, no studies have measured the effects of a focus on meaning in pre-task modeling observations on subsequent task performance.

Observational learning research draws on a social cognitive perspective on learning and claims that learners may 'acquire cognitive skills and new behavior by observing the performance of others' (Bandura, 1986, p. 49). According to Whitehurst and DeBaryshe (1989) 'language, like many other complex skills, is acquired in part through the process of observational learning' (p. 251). However, for learners to actually learn from observation, Bandura (1986) argues that they must pay attention to the model, understand the relevant aspects of what is modeled for retention, be able to reproduce (parts of) the modeled behavior, and be motivated to carry out the modeled behavior. This view on learning connects strongly with SLA theories on attention (Schmidt, 1980), output (Swain, 2000, 2005), practice (DeKeyser, 2007) and motivation (Dörnyei, 1998). In order to acquire language, learners should 1) pay conscious attention to written and/or oral language they understand; 2) produce language, so they can notice gaps in their linguistic knowledge and modify subsequent output; 3) practice to proceduralize language; and 4) feel motivated by teachers who give them interesting, relevant, feasible tasks to do and provide them with feedback.

Other factors that contribute to the effectiveness of observational learning are related to how observers perceive the model's competence (Schunk, 1987). Schunk and Zimmerman (1997) point out that observing successful peers may 'raise observers' efficacy and motivate them to try the task because they may believe that if others can succeed they can as well' (p.197). Furthermore, it has been shown that, in situations where learners learn new skills, they tend to emulate the example of com-

petent models because they want to learn the skills correctly (Schunk & Zimmerman, 1997).

Observational learning has been successfully applied to various language learning skills, such as written composition (Braaksma, Rijlaarsdam & Van den Bergh, 2002; Zimmerman & Kitsantas, 2002), speaking and listening with young children (Sonnenschein & Whitehurst, 1983, 1984), and reading and writing (Couzijn, 1999) (see Braaksma, Rijlaarsdam & Van den Bergh, 2002).

To date, relatively few studies have examined observational learning in terms of focusing on form and language development. Kim (2013), Kim & McDonough (2011), and LaPierre (1994, as cited in Swain, 1998) investigated the effects of different sorts of pre-task modeling videos. In a French immersion context, LaPierre (1994) investigated whether pre-task modeling affected the number of syntactical language related episodes (LREs) during collaborative task performance. During LREs language learners 'talk about the language they are producing, question their language use, or correct themselves or others' (Swain & Lapkin, 1998, p. 326). Participants were assigned to two different conditions, and asked to carry out one dictogloss task a week, for three weeks. Before performing the tasks, half of the students observed the researcher and their teacher talking about grammatical forms through the use of explicit grammar rules, whereas the other students observed the researcher and teacher talking about grammatical forms without metalinguistic rule explanations. Results of students' conversations, while carrying out the subsequent dictogloss in dyads, revealed that students who observed explicit modeling of grammar rules, produced more LREs than students who were provided with implicit talk on grammar rules.

Kim and McDonough (2011) investigated whether pre-task modeling promoted the occurrence of LREs during collaborative task performance. Forty-four female Korean middle school ESL learners were assigned to either a modeling or non-modeling group, which did not receive instructions on how to perform the task, or watch the video models. In three different modeling videos (dictogloss, information-gap task, decision-making task) the researcher and the students' English teacher carried out the tasks, which included LREs referring to both vocabulary and grammar. The findings indicated that learners who watched the pre-task models generated more LREs and found more accurate solutions for the LREs than learners who had not observed the models. In Kim's (2013) study the effects of pre-task model observations were not only measured by the occurrence of LREs, but also included measures of task performance. Forty-five female Korean middle school students, learning English as a second language, were assigned to a modeling group with guided planning time and a no-modeling group with unguided planning time. Modeling videos of three collaborative tasks (information gap, picture difference, decision making) provided the learners with examples of how to pay attention to linguistic forms and demonstrated how to work in pairs. Results showed that the pre-task modeling videos contributed to both learners' attention to form and question development, especially during planning time. The learners' focus on form during planning time consisted mainly of explicit questions about question formation in the L1.

These three studies demonstrate that the observation of modeling videos during the pre-task may stimulate learning, in terms of learners focusing more on form and language development in their subsequent task performance. However, in these studies the modeling videos were specifically set up as instructional material by providing learners with examples of how to pay attention to linguistic forms. No additional instructions were provided.

It should be noted that the strategy observational learning is not limited to the use of instructional models nor to focusing on linguistic structures. Observational learning can also be aimed at more functional aspects of the task and include learner activities that stimulate learners to create input for their own task performance. In addition to observation, regulation or metacognitive strategies, such as evaluation and reflection (Braaksma, Rijlaarsdam, Van den Berg & Van Hout-Wolters, 2004) may also stimulate processes of attention, retention, and production which are essential for learning.

1.2 Guided planning strategies

The observation of models in the pre-task may affect learners' orientation to the task and how learners subsequently plan and perform this task. For that reason, pre-task modeling observations can be considered a type of guided planning strategy (Kim, 2013). Several planning studies have used other strategies to focus learners' attention during the pre-task. We consider it relevant to the present study to review these studies because they can give us insights in how a focus on either or both meaning and form can affect learners' planning processes and subsequent task performance. We will first discuss planning studies that investigated the effects of focused attention during the pre-task on learners' planning processes, and then on task performance, in terms of accuracy, complexity, and fluency.

To date, few studies have investigated the effects of a pre-task focus on either or both meaning and form on learners' planning processes. Wendel (1997), Sangarun (2005), and Kim (2013) attempted to uncover learners' attentional processes and strategic choices during pre-task planning, with help from research methods such as think- or plan-aloud protocols and retrospective interviews. In Wendel's (1997) study students were asked to plan and retell a narrative story which they had seen on video. He compared a form/meaning planning (vocabulary and discourse structure) condition to a minimal strategic planning condition. Post-task interviews showed that most learners focused on completing the story and not on grammar. In Sangarun's (2005) study learners were assigned to three different foci of strategic planning: meaning-focus, form-focus, and meaning-form focus. The analysis of think-aloud protocols revealed that, regardless of the condition, most learners focused on meaning instead of form. Finally, Kim (2013) investigated whether pre-task modeling encouraged learners to attend to question structures during planning time. The results of the think-aloud protocols and video-taped learner-learner interaction showed that pre-task modeling indeed enhanced learners' attention to form (measured by LREs), especially during planning time.

Overall, only Kim's (2013) study showed signs of learners focusing on form during planning time, after having been directed to it. The two other studies (Sangarun, 2005; Wendel, 1997) reported that, irrespective of the focus, learners attended to meaning instead of form. This may indicate that because of learners' dominant focus on meaning 'planning may help achieve a focus on form during the task rather than during planning' (Park, 2010, p.12).

Other studies investigated whether the focus of planning in the pre-task affected subsequent task performance, in terms of accuracy, complexity and fluency. Foster and Skehan (1999) compared the effects of a focus on meaning and form during pre-task planning on subsequent task performance. Based on theories on limited capacity processing (Skehan, 1996a, 1998a) they hypothesized that a focus on meaning during planning time leads to increased complexity of the subsequent task performance because 'the task itself will be complexified as subjects set out for themselves the propositional structure of what they will want to say' (p.224). This could lead to more risk-taking and ambition in what learners try to say (Skehan, 1998a). Conversely, a focus on form would lead to increased accuracy. In Foster and Skehan's study, sixty-six adult students, learning English as a foreign language, were asked to carry out a decision-making task in which they discussed which person needed to be forced out of a hot air balloon, which is losing altitude, in order to avoid a crash. Besides comparing two different planning foci, they also compared teacher-fronted with group-based planning. This led to six different planning conditions: teacher-fronted with language focus, group-based with language focus, teacher-fronted with content focus, group-based with content focus, solitary planning and no planning. Learners were stimulated to focus on form through either teacher-led instructions on the use of modal verbs and conditionals or instructions on paper to control for correct English. Participants were stimulated to focus on meaning by the teacher who held a presentation on ideas that each character might use to defend his or her right to stay in the balloon. The instructions on paper required learners to think about reasons why a certain person would not be thrown out of the balloon. Results showed no effects for focus of planning, which led Skehan and Foster to the conclusion that 'effects of planning are not simply attributable to whether subjects concentrate on one of these areas rather than another' (p. 239). What they found was a main effect on accuracy for the source of planning. The teacher-fronted condition produced higher levels of accuracy than the group-based condition. Since the teacher-fronted condition received explicit instructions on the use of modal verbs and conditionals and the group-based condition was only asked to make sure that the English was correct, it remains unclear whether this effect on accuracy can only be attributed to the source. It might also be possible that the differences in accuracy scores can be explained by the different degrees of explicitness in the planning instructions.

In Sangarun's (2005) study, forty Thai, eleventh grade EFL learners were assigned to four strategic planning conditions: minimal planning (MinP), form planning (FP), meaning planning (MP), meaning/form planning (MFP). Effects were measured of these conditions on quality of speech (i.e. accuracy, complexity, and fluency). The results showed positive effects on complexity for the MP and MFP conditions. On accuracy, the MFP, MP, and FP conditions outperformed the MinP condition, but no significant effects between the other three conditions were found.

In addition, positive results were found on speech fluency for the MFP, MP, and FP conditions. Based on these findings, Sangarun argues that learners should focus on both meaning and form during strategic planning because this appears to be more effective than focusing on either one of them.

To conclude, the current body of research shows no convincing effects on either planning processes or task performance due to differences in focus of attention. It might be the case that the planning instructions, used to direct learners to either form or meaning, are insufficiently connected to the task performance itself. That is to say, it remains unclear to the learners what successful task performance, according to the different foci, looks like. For that reason, we argue that observing other learners carry out the same or a similar task may give learners 'a clear image' of how a task should be performed (Zimmerman & Kitsantas, 2002, p. 660). With such clear image in mind, directions to either form or meaning through evaluation and reflection, can be connected more easily to that image and stored in the memory as 'imaginal and verbal codings' (Bandura, 1986), ready to be retrieved during subsequent task performance.

1.3 *The present study*

The present study compares the effects of focusing on meaning and form during pre-task peer modeling observations on both planning processes and task performance. To this end, three research questions were formulated.

- RQ1 Does a focus on either meaning or form during pre-task peer modeling observations affect the planning processes prior to subsequent task performance?
- RQ2 Does a focus on either meaning or form during pre-task peer modeling observations affect subsequent task performance in terms of accuracy, complexity, and fluency?
- RQ3 Is there a trade-off between accuracy, complexity and fluency?

With this last question we investigated Skehan's (1996a, 1998a) *Trade-off Hypothesis* which claims that learners have limited attentional resources. As a consequence, learners' attention to either an accurate, complex or fluent task performance would entail less attention to the other dimensions and thus diminish the performance therein.

2. METHOD

2.1 *Participants*

Participants in the present study were forty-eight ninth-grade students learning German as a foreign language (A2 level of the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001)). Participants' mean age was 14,2 years and they all had a Dutch language background. Participants had 2 hours of German per week for about 17 months. Participants were recruited from the same

Dutch secondary school. Each participant was randomly assigned to the two conditions: Focus on Form ($n = 25$) and Focus on Meaning ($n = 23$ ¹). Participants were used to working with tasks because the curriculum in these two classes consists of tasks which are based on the CEFR. In accordance with the protocol of the University of Amsterdam's Faculty of Humanities' Ethics Committee, all parents were informed about the study and the possibility of opting out of participation in the study.

2.2 Modeling task

As a pre-task, all participants watched two videos of two different girls describing the school canteen to future first formers to inform them about their school and recruit them for it. The video scripts were written by the first author who was also the teacher of both groups. The girls in the videos were excellent former students who had already performed the modeling task the year before. Before the videos were recorded, the former students had practiced the oral task at home and at school with the teacher.

Before watching these videos, participants had read instructions that told them they were going to perform a similar task afterwards. Video 1 lasted 1.26 minutes and Video 2 lasted 1 minute. To increase the chance of identification, confidence and motivation, the girls in the videos were about the same age as the participants. In this way, participants could experience that if other peers could perform this task successfully they could as well (Schunk & Zimmerman, 1997). Based on theories on perceived competence that suggest that learners of new skills tend to follow the example of competent models (Schunk & Zimmerman, 1997) the girls in the modeling videos performed the task accurately and with confidence.

2.3 Target structure

In order to carry out the main task 'describe the school canteen to inform and recruit other pupils for your school', the use of locative prepositions is indispensable (see task-essentialness, Loschkey & Bley-Vroman, 1993). For that reason, the dative case of an article after a two-way preposition (*in, an, auf, hinter, neben, unter, über, vor, zwischen*²) was chosen as the target structure of the current study (see also Van de Guchte, Braaksma, Rijlaarsdam, & Bimmel, 2015a). In this example, *auf dem Bett liegt ein Kissen* (on the bed lies a pillow), the preposition *auf* demands the dative case and consequently the neuter definite article *das* changes into *dem*. This structure is considered a complex, rule-based structure with no equivalents in the L1 Dutch. Since the 19th century, Dutch nouns and articles are no longer inflected for case (Van den Toorn et al., 1997) which means that a transfer between L1 and L2 is

¹ We started the study with 49 participants, divided between a form ($n = 25$) and meaning ($n = 24$) condition. One week after the pre-tests were administered, one student in the FonM condition changed school level and left the class.

² Translation: *in, at, on, behind, next to, below, above, in front of, between.*

not very likely. Students' pre-test results showed indeed a low level of knowledge of the target structure.

2.4 Research design and procedures

In this experimental study we carried out a pre-post-delayed posttest research design which was spread out over a period of six weeks (see Figure 1). The modeling task was carried out under two conditions: a Focus on Form (FonF) and a Focus on Meaning (FonM) condition.

	Condition Focus on Form	Condition Focus on Meaning
Wk 1	Pretest: Communicative task performance (5 min)	
Wk 3	Pre-Task Model Observation Instructions directed to <i>form</i> (20 min)	Pre-Task Model Observation Instructions directed to <i>meaning</i> (20 min)
	Think-aloud instruction by research-assistant: Think-aloud task planning (10 min)	
	Posttest-1: Communicative task performance (5 min)	
Wk 6	Posttest-2: Communicative task performance (5 min)	

Figure 1. Research design FonF in the pre-task.

The operationalization of the two conditions is as follows. All participants watched Video 1 and 2. Observation sheets with written, reflective questions directed participants' attention to either form (FonF) or meaning (FonM). While observing, the FonF group was asked to write down a total number of 12 linguistic structures students in the videos used to describe the place of the furniture and the accessories in the school canteen (see Appendix A). Analysis of the FonF students' notes showed that they indeed provided the target structures on the observation sheets ($M = 11, 2$; $SD = 1,70$). The FonM group was asked to compare the rhetorical structure of both presentations and what the students had done to make the presentation attractive and persuasive (see Appendix B). Participants in both conditions were allowed to pause or scroll through the video and to make notes.

The intervention was conducted in both classes by the first author, who also works as a teacher in these groups. Two weeks prior to the experiment, students performed a pre-test in which they were asked to describe the school canteen to future first formers to inform and recruit them for their school. On the day of the experiment, all students were seated in the school's computer lab and watched two videos. Students of the FonF group received form planning note-sheets, whereas learners of the FonM group were given meaning planning note-sheets. The students were aware that the task was part of a research project but were not informed about the different conditions. The students observed peer models performing the task and answered the questions on the note-sheets. A logging program on the computer tracked the students' actions (total time, pausing, scrolling), while observing the two videos. The mean observation time was 22.33 minutes ($SD = 7.07$) for the FonF group and 22.30

minutes ($SD = 8.11$) for the FonM group. No statistical significant differences in observation time were found between the conditions ($F(1, 46) = .00, p = .984$). Nor were there significant differences between groups at pausing ($F(1, 46) = 1.46, p = .234$) or scrolling ($F(1, 46) = 3.89, p = .056$). When students finished the observation task, they were asked to hand in the note-sheets to the research assistant. Afterwards each student went to a separate room where he received written posttest-1 instructions (see Appendix C) from a research assistant. Ten research assistants administered the tests with help of a protocol to guarantee that all tests were performed in the same way. The assistants instructed the students on how to think aloud, and students were subsequently asked to read aloud the main task. Then, students were asked to look at a picture of a school canteen and were required to think-aloud on how they would plan the main task. To this extent, students were asked two questions: Q1) How will you approach this task? and Q2) What are you going to tell? Students were given ten minutes of planning time to answer both questions and plan the main task performance. In addition, for the second question students were allowed to write down a maximum of ten keywords. Students were not allowed to write down everything in detail because this might have led to learners reciting the plans and may have hampered natural speech (Sangarun, 2005). Students were free to plan the task in Dutch (L1) or German because we wanted to create a natural situation which resembled what they were used to in a 'normal' classroom situation (Kim, 2013).

Not all students used the full 10 minutes provided. The mean number of planning time for the FonF group was 7.56 minutes ($SD = 2.95$) and 7.35 ($SD = 2.60$) minutes for the FonM group. No statistical significant differences between conditions were found ($F(1,46) = .070, p = .793$). After planning students handed over the think-aloud protocols to the research assistant and performed the main task (post-test 1) which was a different version of the modeling task. Students performed another version of the main task (post-test 2) three weeks later.

2.5 Testing tasks

In all testing tasks (pre-, post-1, post-2) students were asked to give a tour through the school canteen to inform and recruit other pupils. To this end, they were provided with a picture of the school canteen. Three parallel test versions were designed in which the nationality of the audience (German, Swiss, Austrian) and the style of the school canteen differed. To avoid a repetition effect, furniture and accessories were located in different places. For the pre-test, each participant received version A, for post-test 1 version B (see Appendix C), and for post-test-2 version C, thus all conditions received the same test at each measurement occasion.

Learners' oral task performances were audio-taped and transcribed on all three occasions to measure accuracy, complexity and fluency. Transcripts were coded for the Analysis of Speech Unit (AS-Unit) defined as 'a single speaker's utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either' (Foster, Tonkyn, & Wigglesworth, 2000, p. 365).

2.5.1 Accuracy

Regarding accuracy, we measured the *correct* use of the dative case of the article after a two-way preposition per clause. Because the intervention was aimed at the (correct) use of this structure, we opted for this measure of accuracy instead of a more general measure such as error-free clauses per AS-Unit.

2.5.2 Complexity

Complexity was operationalized by means of four different measures, frequently used in SLA research (Norris & Ortega, 2009): 1) the use of the target structure (dative case of the article after a two-way preposition) per clause. As learners of English (Newton & Kennedy, 1996), learners of German find it difficult to use prepositions in a sentence. For example, learners prefer saying *'links steht ein Tisch'* (on the left is a table) than the more complex *'auf der linken Seite steht ein Tisch'* (on the left side is a table). Therefore the use of the target structure is considered a measure of complexity; 2) general complexity, by means of the total number of clauses per AS-unit; 3) complexity by subordination (number of subordinated clauses per AS-unit); and 4) complexity by coordination (number of coordinated clauses per AS-unit). Following Bardovi-Harlig (1992), we included this measure of complexity, because for learners at lower production levels, such as those included in the present study, increased complexity is not only shown through subordination but also through an increase in coordinated sentences.

2.5.3 Fluency

With respect to fluency, we ran a script on the sound files written by De Jong and Wempe (2009) in PRAAT (Boersma & Weenink, 2007) with which we were able to calculate students' *speech rate* (total number of syllables divided by total duration of the task performance). However, as De Jong et al. (2012) point out speech rate takes 'breakdown fluency and speed fluency (...) together into one measure that encompasses aspects of pausing as well as speed of delivery' (p. 2). Therefore we included two measures that did not confound pausing and speed of delivery: *articulation rate*, that is, mean duration of syllables (speaking time divided by total number of syllables), and the *number of silent pauses* (number of silent pauses divided by speaking time). Following De Jong et al. (2012), silences of 250 ms. or longer were considered to be hesitations or pauses, and thus silences shorter than 250 ms., so-called micropauses (see, e.g., Riggensbach, 1991), were discarded.

2.6 Planning processes

The think-aloud method was used to enable us to analyze what learners actually thought while planning the task. Although think-aloud protocols are sometimes criticized for being obtrusive and hampering the nature of planning (Ortega, 1999), they are considered a well-established method for capturing learners' planning processes (Gass & Mac Key, 2000; Kim, 2013). The audio-recordings of the think-aloud pro-

protocols were transcribed by research assistants. Afterwards they were segmented into communicative c-units (see Foster & Skehan, 1996, 1999; Sangarun, 2005). Learners' attention to form (see form and meaning planning categories in Appendix D) was then coded by the researcher in terms of the occurrence of the linguistic target structure (dative case of the article after a two-way preposition), plus syntactical language related episodes referring to the use of these prepositions ('I should use the prepositions in, an, auf here'). Since many students used the L1 while planning the task, we included the Dutch equivalents (*in het midden - in the middle*) in the counting. In addition, learners also used LREs referring to the use of lexicon. These lexical LREs included mostly talk about translations³.

For learners' attention to meaning we adapted Sangarun's (2005) categories in the coding scheme for meaning planning (see Appendix D). Sangarun's coding scheme relies on two main principles of the constant comparative method (Glaser & Strauss, 1967). First, categories were derived from the data, instead of applying pre-conceived categories to the data. Second, we refined 'each coding category' and identified 'its properties through comparing all instances coded in one category with other categories' (Sangarun, 2005, p. 123). For the current study, this led to the following five categories for both planning questions: 1) organizing discourse; 2) describing furniture and accessories in the school canteen; 3) describing activities in the school canteen; 4) ideas to make the presentation attractive; and 5) ideas to persuade pupils to come to their school. Protocols of all participants were coded for meaning planning categories by both the first author and a trained assistant with an interrater reliability of 96 % for question 1 (Approach) and 94 % for question 2 (Content).

2.7 Data analysis

A series of univariate GLMs were performed, to analyze learners' accuracy, complexity and fluency. Condition was put in the model as a fixed factor. The corresponding pre-test results were used as a covariate for the post-test results. To determine whether participants showed development in fluency over time, we performed a GLM repeated measures analysis, in which time and condition were included as fixed factors. To establish whether there were trade-offs between accuracy, complexity and fluency, we performed a Pearson's correlation analysis. The alpha for achieving statistical significance was set at .05.

For the analysis of the form and meaning planning categories in the think-aloud protocols, a series of one-way ANOVAs were performed. Before doing so, we first checked whether the five meaning categories were correlated (Pearson) and could be reduced by means of a factor analysis (PCA). Varimax with Kaiser normalization was selected as rotation method. For both planning questions (Q1: Approach and Q2: Content), three of the five categories correlated at least .36, suggesting reasonable factorability. Two components were identified as the results of PCA (for rotated

³ Because the focus of the present study was the use of LREs about the target structure, we decided not to report on the use of lexical LREs.

component matrices see Table E1 in Appendix E), which explained 64,70 % (Q1) and 57,70 % (Q2) of the variance respectively. Table 1 explains the two principle components for meaning planning.

Table 1. Two components for meaning planning

Component		Coding categories
1	Description ideas (64,70 % of variance explained)	Describing furniture and accessories Describing activities in the canteen
2	Communicative effectiveness (57,70 % of variance explained)	Organizational discourse Ideas to make presentation attractive Persuasive ideas

3. RESULTS

The results are described in two main parts: Effects on planning processes and effects on measures of task performance. Results of a series of one-way ANOVAs showed no initial differences on the pre-tests, for all measures the mean alpha level was $p = .55$ varying between $p = .356$ and $p = .932$.

3.1 Effects on planning processes

Tables 2.1 and 2.2 report the descriptive statistics for planning questions 1 (Approach) and 2 (Content). For both planning questions no statistical significant differences were observed for either form planning or meaning planning.

Table 2.1 Planning Question 1 (Approach)

Focus of Planning	Measure	Condition			
		FonF		FonM	
		Mean	SD	Mean	SD
Form Planning	Selecting Target Structure	.88	1.83	.57	.79
	LRE Target Structure	.20	.50	.04	.21
Meaning Planning	Organizing Discourse	1.80	1.41	2.30	1.61
	Describing furniture and accessories	2.40	2.45	1.61	2.02
	Describing activities in the canteen	.48	.77	.39	.89
	Ideas to make presentation attractive	.76	1.09	.70	1.02
	Ideas to make presentation persuasive	.12	.44	.17	.65

Note. FonF = Focus on Form condition ($n = 25$); FonM = Focus on Meaning condition ($n = 23$)

Table 2.2 Planning Question 2 (Content)

Focus of Planning	Measure	Condition			
		FonF		FonM	
		Mean	SD	Mean	SD
Form Planning	Selecting Target Structure	1.84	1.72	1.04	1.19
	LRE Target Structure	.04	.20	.04	.21
Meaning Planning	Organizing Discourse	1.12	1.39	1.74	1.76
	Describing furniture and accessories	3.68	2.67	2.57	1.99
	Describing activities in the canteen	.72	.79	1.30	1.33
	Ideas to make presentation attractive	.68	1.28	.78	1.09
	Ideas to make presentation persuasive	.20	.41	.35	.57

Note. FonF = Focus on Form condition ($n = 25$); FonM = Focus on Meaning condition ($n = 23$)

3.2 Effects on measures of task performance

Table 3 reports the descriptive statistics for three measures of task performance: accuracy, complexity, and fluency.

3.2.1 Accuracy

Correct use of target structure per clause

At Post-1 the FonF condition outperformed the FonM condition ($F(1,46) = 14.49$, $p < .001$) with a large effect size ($d = 1.08$). No differences between conditions were observed at Post-2 ($F(1,46) = 3.74$, $p = .059$).

3.2.2 Complexity

Use of target structure per clause

At Post-1 the FonF condition outperformed the FonM condition ($F(1,46) = 18.90$, $p < .001$) with a large effect size ($d = 1.3$). No statistical significant differences between conditions were found at Post-2 ($F(1,46) = .78$, $p = .382$).

Clauses per AS-unit

At Post-1 the FonM condition outperformed the FonF ($F(1,46) = 4.32$, $p = .044$) with a nearly medium effect size ($d = .45$). At Post-2 the FonM condition also outperformed the FonF condition ($F(1,46) = 6.71$, $p = .013$) with a medium effect size ($d = .70$).

Number of coordinated clauses (per AS-unit)

At Post-1 the FonM condition outperformed the FonF condition ($F(1,46) = 5.22$, $p = .027$) with a medium effect size ($d = .63$). But no differences between conditions were found at Post-2 ($F(1,46) = 2.40$, $p = .128$).

Number of subordinated clauses (per AS-unit)

At Post-1 no significant differences were found between conditions ($F(1,46) = .226$, $p = .636$). At Post-2 the FonM condition outperformed the FonF ($F(1,46) = 4.85$, $p = .033$) with a medium effect size ($d = .57$).

Table 3. Measures of Task Performance for two conditions

Measure	Condition	Pre-test		Post-1		Post-2		
		Mean	SD	Mean	SD	Mean	SD	
Accuracy	Correct Use of TS per clause	FonF	.03	.06	.14	.10	.10	.14
		FonM	.04	.09	.04	.09	.04	.07
Complexity	Use of TS per clause	FonF	.12	.14	.40	.18	.28	.26
		FonM	.15	.16	.20	.14	.23	.18
	Clauses per ASU	FonF	1.15	.23	1.11	.17	1.13	.11
		FonM	1.10	.11	1.19	.15	1.25	.22
Coordination per ASU	FonF	.06	.13	.05	.12	.07	.15	
	FonM	.06	.12	.15	.19	.15	.22	
Subordination per ASU	FonF	FonF	.12	.21	.11	.15	.10	.09
		FonM	.07	.11	.11	.15	.19	.18
	FonF	FonF	2.58	.60	2.70	.52	2.77	.58
		FonM	2.43	.64	2.41	.54	2.69	.56
Fluency	Speech rate	FonF	3.88	.60	3.89	.40	3.92	.43
		FonM	3.84	.41	3.77	.46	3.93	.57
	Articulation rate	FonF	.62	.23	.58	.22	.54	.21
		FonM	.69	.25	.69	.24	.60	.18

Note. FonF = Focus on Form condition ($n = 25$); FonM = Focus on Meaning condition ($n = 23$).

Transcripts from the oral task performances showed that some students ($n=5$) of the FonM condition who used the conjunction *und* (and) to combine two clauses at post-1, used the conjunctions *wo* (where) or *worauf* (on which) for the same sentences at post-2 (see Figure 2). Apparently, being familiar with both content and formulation, made it possible for these FonM students to make their sentences more complex by converting non-hierarchical coordinated sentences into hierarchical subordinated sentences.

Post-1: *Es gibt Automaten und Schüler kaufen da Süßigkeiten.*
 (In the corner are vending machines and students buy candies over there.)
 Post-2: *Es gibt Automaten, wo Schüler Süßigkeiten kaufen.*
 (In the corner are vending machines, where students buy candies.)
 Post-1: *Es gibt mehrere Stühle und darauf kann man sitzen.*
 (There are several chairs and you can sit on them.)
 Post-1: *Es gibt mehrere Stühle, worauf man sitzen kann.*
 (There are several chairs, on which you may sit.)

Figure 2. Examples of students converting coordinated sentences into subordinated sentences.

3.2.3 Fluency

No effect of conditions was observed for either: 1) speech rate (Post-1: $F(1,46) = 2.55, p = .117$; Post-2: $F(1,46) = .054, p = .817$) 2) articulation rate (Post-1: $F(1,46) = .96, p = .333$; Post-2: $F(1,46) = .018, p = .893$) or 3) number of silent pauses (Post-1: $F(1,46) = 2.50, p = .121$; Post-2: $F(1,46) = .053, p = .819$). Furthermore, we observed a main effect of time for both conditions on speech rate ($F(1,46) = 4.67, p = .014$) indicating that students of both the FonF and FonM condition gained in fluency over time.

3.3 Trade-off

A significant negative correlation between the correct use of the target structure (accuracy) and two complexity measures was observed at Post-2 for the FonF condition: the amount of clauses per ASU ($r = -.58, p = .002$) and the ratio of subordination ($r = -.54, p = .024$). In addition, we found significant negative interactions between two complexity measures. At Post-1, this interaction was found for the FonF condition between the use of the target structure and the ratio of subordination ($r = -.42, p = .038$) and for the FonM condition between the use of the target structure and the amount of clauses per ASU ($r = -.45, p = .031$). At Post-2, this correlation was observed for the FonF condition between the use of the target structure and the amount of clauses per ASU ($r = -.62, p = .001$) and the ratio of subordination ($r = -.60, p = .002$).

4. DISCUSSION

In Table 4 a summary of the results is presented.

Table 4. Summary of results

Measures	Posttest-1	Posttest-2
<i>Accuracy</i>		
Correct use of target structure	FonF > FonM	No effect
<i>Complexity</i>		
Use of target structure	FonF > FonM	No effect
Clauses per ASU	FonM > FonF	FonM > FonF
Ratio of coordinated clauses	FonM > FonF	No effect
Ratio of subordinated clauses	No effect	FonM > FonF
<i>Fluency</i>		
Speech rate; Articulation rate; Silent pauses	No effect	No effect
<i>Planning processes</i>		
Form planning, Meaning planning	No effect	No effect

The purpose of this study was to examine whether focusing on meaning or form in pre-task video model observations would affect the learners' planning processes and their subsequent task performance in a foreign TBLT context. As far as we know, this is the first study that uses observational learning with video models to compare meaning and form conditions in pre-task planning activities. With the results of the study we aim to contribute to research on strategies that may be effectively used to direct learners' attentional resources (meaning or form) and affect learners' subsequent task performance.

Regarding the first research question, whether a focus on either meaning or form would affect learners' planning processes prior to subsequent task performance, the findings revealed no statistical significant differences between conditions. Because we had observed significant differences between conditions on measures of task performance, accuracy and complexity, we assumed that these differences would be reflected in how learners had planned the main task. We expected students in the FonF condition to focus more on the use of the target structure and students in the FonM condition to focus more on different aspects of the content of the task. However, the results of the planning protocols indicated that, in both conditions, students mainly planned their content and focused to a much lesser extent on the use of the target structures. In line with Park (2010), Ellis (2005b), and Skehan (1996a, 1998a), we suggest that learners primarily conceptualize their message during planning time, with very little attention left for accurate formulation. In comparison, during task performance, when learners already have in mind what they want to say sufficient attentional resources may be available to focus on the language they want to use.

Regarding the second research question, whether a 'focus on either meaning or form in pre-task peer modeling observations may affect subsequent task performance' findings demonstrated that students in the FonF condition outperformed those in the FonM condition on accuracy and on one complexity measure at posttest-1. It appeared that students in the FonF condition used the target structure significantly more often and more accurately. We presume that the observations, combined with written instructions that focused on the use of the dative structure, succeeded in making the target structure salient. Analysis of student's notes on the preposition plus dative structures on the observation sheets confirmed this assumption. This level of awareness, which Bandura (1986) considers an important condition for learning through observation, may have activated general and even accurate use of the target structure during immediate, subsequent task performance. However, at posttest-2, three weeks later, these differences disappeared. The reason for the absence of significant long-term effects may be that the use of the structures was induced by a 'short-term activation from a memory representation' (see Bock & Griffin, 2000, p. 177). In other words, learners had memorized the structure. To achieve long-term effects, the implicit focus in the pre-task may need to be followed by other additional learning activities at a later stage in the TBLT cycle. This view is supported by Whitehurst and Vasta (1975) who argue that an implicit focus on form is only a first step to introducing syntactic structures into the productive mode. Schunk (2007) proposes that the observation of models could be combined with guided practice and corrective feedback. Zimmerman and Kitsantas (2002) define this as the emulation level of, in their case, writing skill acquisition in which after the level of observa-

tion, learners learn to enact (parts of) the modeled behavior. At this level of learning, providing learners with feedback is essential.

It may also be argued that the observations plus writing down the dative structures was a too implicit way of focusing on form. More explicit foci could have led to gains in explicit knowledge, resulting in more sustainable acquisition. Nevertheless, the choice for an implicit focus on form was a carefully considered decision. In order to minimize the interruption to learners' communication of meaning, we argue for a focus on form in the pre-task that attracts learners attention to form and avoids metalinguistic talk about grammar (Doughty & Williams, 1998).

Results also indicated that students in the FonM condition outperformed those in the FonF condition on three complexity measures: 1) clauses per AS-unit at both post-1 and post-2; 2) amount of coordinated sentences per AS-unit at post-1; and 3) amount of subordinated sentences per AS-unit at post-2. These findings suggest that observing others perform a similar task, in conjunction with instructions directed towards the content of the task, may lead to a more complex subsequent task performance. It can be argued that the evaluation of and reflection on the modeled performances (Braaksma et al., 2004) made students think actively about how to give an attractive and persuasive presentation. This may have resulted in more risk-taking and experimentation with the language they wanted to use. As a consequence, more clauses per ASU were produced, including sentences that were combined through either coordination or subordination.

While the effects of the form intervention were short term ones, the effects of the meaning intervention were still visible in the long-term. The question is which elements of the intervention contributed to this long-term effect. Presumably, the observation of the modeled task, combined with learning activities that asked the students to evaluate and compare the two modeled presentations, enabled students to: 1) pay conscious attention to those elements of the task that were crucial for their own successful task performance; 2) store imaginal and verbal conceptions of the tasks in memory and connect them to previous knowledge about presentations; and 3) translate the task conceptions into language production (Schunk, 2007).

The third research question refers to the assumed trade-off between accuracy, complexity and fluency. A positive answer to this question would require a negative correlation between either one of them. No such effect was found between fluency and accuracy. This is not surprising considering the implicit nature of the focus on form. The syntactic structure was provided as a chunk without any references to grammar rules. As a result, learners could imitate the structure, without imposing a heavy burden on the attentional resources which could lead to lower fluency scores.

In contrast, we found a moderate negative relationship between accuracy and two complexity measures for the FonF condition at post-2. This implies that learners who used the target structures more accurately, generated less clauses per AS-unit and obtained lower scores on subordination. Another interesting trade-off was found between two complexity measures. At post-1 students of the FonF condition who used the target structure more frequently, scored lower on subordination, whereas students of the FonM who used the target structure more frequently, produced less clauses per AS-unit. The correlation analyses of post-2, as described in the results

section, showed that students of the FonF condition who used the target structure more frequently, generated less clauses per AS-unit and a lower ratio of subordination. In short, all these trade-offs suggest that the (accurate) use of syntactic dative structure, may have negative consequences for the complexity of the task performance, in terms of clauses per AS-unit and the ratio of coordination and subordination.

4.1 Limitations and suggestions for future research

Some limitations of the current study need to be acknowledged. First, the use of observational learning as a pre-task activity may conflict with the principle of task-based language teaching, that allows learners to choose their own linguistic and non-linguistic resources to fulfill the task (Ellis, 2003; Willis, 1996). Providing learners with a model, may promote imitation instead of learners' use of their own creativity. On the other hand, imitation is a basic learning principle in foreign language acquisition.

Second, one may argue that the focus on linguistic structures was too implicit to call it a focus form, in the sense that it did not make learners aware of the rules that are connected with the used prepositions plus a dative structure. As a result, the use of the structure may then be limited to the situation presented in the video and could never result in general use of the rule in different circumstances. Although we do not regard the pre-task as the right place for explicit grammar teaching, it could be interesting to investigate the effects of a more explicit focus on form.

Third, the discussed studies on observational learning in the literature review included studies that measured the impact of observational learning on general education including L1 learning. It must be acknowledged that the effects of those studies may not be fully applicable to L2 learning, since these are considered two different areas.

Fourth, we measured the effects of two types of observation strategies on the learners' performance in (1) the pre-task planning and (2) the main task. However, we cannot be sure that the conducted think-aloud planning method right after the first phase, may have influenced the learners' performance as well.

Finally, since the aim of the current study was to compare the effects of two types of pre-task modeling observations, the results do not provide insights into whether learners who are provided with models benefit more than learners who are not provided with them.

A challenging task for further research would be to investigate the effects of different focus on form and meaning instructions in relation to other grammatical structures and task types.

5. CONCLUSIONS

The current study aimed to investigate whether the direction of learners' attention to either form or meaning, through observational learning, would affect their planning processes and subsequent task performance. The results indicate that a focus on

form may lead to gains in accuracy and complexity (use of the target structure), whereas a focus on meaning may promote a more complex task performance in terms of clauses per AS-unit, and the amount of coordination and subordination used. In some cases, these gains were achieved at the expense of other aspects of performance, supporting Skehan's (1996a, 1998a) Trade-off Theory, which claims that learners' have limited attentional resources that compete with each other. We found trade-offs between accuracy and complexity and between different measures of complexity. Although we found a clear effect on the task performance between the conditions, this effect was not observed in the think aloud planning protocols.

Considering that the current study has been carried out under real classroom conditions, we argue that the findings of our study may have an important implication for language pedagogy in a foreign task-based language learning context. Observational learning can easily be carried out in the classroom, given that electronic devices are often already present. A teacher can video-tape a learner carrying out a task, and then use this video in the forthcoming years. Also 'live modeling' with either the teacher or a student, acting as a model, is possible

Depending on the purposes of the lesson, instructions may guide learners' attention to either meaning or form. In doing so, observational learning may provide a useful and naturalistic strategy to achieve balanced language development in the modern TBLT classroom.

APPENDIX A

FOCUS ON FORM INSTRUCTIONS VIDEO MODEL OBSERVATIONS

Video fragment I: Tara

- 1) Observe the video fragment in which *Tara* presents the school canteen.
You are allowed to scroll through the fragment, pause or review the fragment.
- 2) Write down six different examples of how *Tara* describes where all the furniture and accessories are placed in the school canteen.

For example: An der Wand hängt ein Poster.

Video fragment II: Laura

- 1) Observe the video fragment in which *Laura* presents the school canteen.
You are allowed to scroll through the fragment, pause or review the fragment.
- 2) Write down six different examples of how *Laura* describes where all the furniture and accessories are placed in the school canteen.

For example: In der Mitte stehen Tische.

APPENDIX B

FOCUS ON MEANING INSTRUCTIONS VIDEO MODEL OBSERVATIONS

Video fragment I: Tara

1) Observe the video fragment in which Tara presents the school canteen.
You are allowed to scroll through the fragment, pause or review the fragment.

2) Which steps does Tara use to set up her presentation?

3) Besides structuring the presentation, what does Tara do to make the presentation attractive for the pupils?

4) What does Tara do to persuade the pupils to come to her school?

Video fragment II: Laura

1) Observe the video fragment in which Laura presents the school canteen.
You are allowed to scroll through the fragment, pause or review the fragment.
Which steps does Laura use to set up her presentation?

2) Besides structuring the presentation, what does Laura do to make the presentation attractive for the pupils?

3) What does Laura do to persuade the pupils to come to her school?

4) Now compare the two presentations and decide which student gives the best presentation and explain why. Tara/Laura gives the best presentation, because

APPENDIX C

POST-TEST 1 - THINK ALOUD PROTOCOL AND TASK - VERSION B

Name: Time:until.....

1. Read aloud the following task.

On the website Etwinning, schools may get in contact with foreign schools to organize a school exchange. Our school, the ARHC, is looking for a German exchange partner for the year 2014 and for this purpose wishes to upload an attractive and inviting video clip on the Etwinning website. You are going to record the text for this video clip in which you lead a guided tour through the school canteen. The goal is to present the school canteen to Austrian ninth-graders and persuade them to come to our school this summer.

2. Look at the picture of the school canteen.
3. How will you approach this task? Think aloud.
4. What are you going to tell? Think aloud.
(You are allowed to make short notes)
 - 1.
 - 2.
 - 3.
 - 4.
 - 5.
 - 6.
 - 7.
 - 8.
 - 9.
 - 10.
5. Hand this note-sheet over to the research assistant and then carry out the task.

APPENDIX D

CODING SCHEME: CATEGORIES OF THE THINK-ALOUD PLANNING
PROTOCOLS * ADAPTED FROM SANGARUN (2005)

Coding Categories	Definitions and examples from the protocols
1. <i>Form Planning</i> 1.1 Selecting the target structure	Protocol statements in this category show that participants select the target structure for their message Example: ' <u>In der Mitte</u> stehen die Stühle und <u>an der Wand</u> hängt ein Poster.'
1.2 LREs referring to target structure	Protocol statements in this category show that participants refer to the use of the target structure. Example: 'I should use a preposition like auf or an here.'
2. <i>Meaning Planning</i> 2.1 Organizing Discourse	Protocol statements in this category show that participants are organizing their speech beyond sentence level. Example: 'I should start by introducing myself and then I should welcome the students.'
2.2 Describing furniture and accessories	Protocol statements in this category show that participants formulate what the canteen looks like by describing the furniture and accessories. Example: 'An der Decke hängen verschiedene Lampen und in der Ecke stehen Mülleimer.'
2.3 Describing activities in the canteen	Protocol statements in this category show that participants generate ideas about the activities that can be carried out in the canteen. Example: 'You can sit down with your friends here and play cards or listen to music.'
2.4 Ideas to make presentation attractive	Protocol statements in this category show that participants formulate what makes the school canteen attractive. Example: 'Everybody is very happy, so there is a great atmosphere in the canteen.'
2.5 Ideas to make presentation persuasive	Protocol statements in this category show that participants generate ideas to persuade the pupils to sign up for their school. Example: 'I hope that you will come to our school next year.'

APPENDIX E

TABLE E1

*Rotated Component Matrices Factor Analysis Meaning Planning**E1.1. Planning Question 1 (Approach)*

Meaning Planning Categories	Component	
	1	2
Organizing discourse	-.494	.405
Describing furniture and accessories	.738	
Describing activities in the canteen	.817	
Ideas to make presentation attractive	.423	.789
Ideas to make presentation persuasive		.852

E1.2. Planning Question 2 (Content)

Meaning Planning Categories	Component	
	1	2
Organizing discourse	-.632	.473
Describing furniture and accessories	.692	
Describing activities in the canteen	.618	.367
Ideas to make presentation attractive		.689
Ideas to make presentation persuasive		.853

CHAPTER 4

LEARNING NEW GRAMMATICAL STRUCTURES IN TASK-BASED LANGUAGE LEARNING: THE EFFECTS OF RECASTS AND PROMPTS¹

In the present study we examined the effects of prompts and recasts on the accurate use of two new and different grammar structures in a task-based learning environment. Students were randomly assigned to two experimental conditions: one receiving prompts, the other recasts. These experimental conditions were compared to a control condition, which was an intact class. The study involved two subsequent interventions: The first targeted a complex structure, dative case after a preposition; the second a more simple structure, comparatives. Pre-tests, immediate, and delayed post-tests included written and oral accuracy as well as oral fluency. Statistical comparisons on both written and oral post-tests showed that prompts and recasts were effective, when compared to the control group, with prompts being superior to recasts. Furthermore, the effect of recasts depended on the structure: Recasts were more effective for the comparative than for the dative on written accuracy, as compared to prompts.

1. INTRODUCTION

Over the last three decades, researchers as well as teachers have shown much interest in task-based language teaching (TBLT). There is steadily growing empirical evidence that task-based language learning promotes language acquisition through different kinds of processes such as negotiation of meaning, uptake of corrective feedback, noticing the gap between incorrect performance and correct use of the target structure, metalinguistic reflection, and automatization (see Robinson's review article on task-based language learning, 2011). In an increasing number of classrooms, the classic PPP-model (presentation–practice–production) makes room for learning languages by means of tasks. Such tasks primarily focus on meaning. In this process, rich and authentic input (Krashen, 1981) asks students to produce output (Swain, 2000, 2005) and promotes interaction (Long, 1996; Long & Robinson, 1998), thus stimulating language acquisition.

A central issue in task-based language teaching, however, is whether input, output, and interaction are sufficient to acquire language (Skehan, Xiaoyue, Qian, & Wang, 2012). Several researchers have expressed serious doubts (Doughty & Varela, 1998;

¹ This chapter is an adapted version of an article that is published as: Van de Guchte, M., Braaksmā, M., Rijlaarsdam, G., & Bimmel, P. (2015a). Learning new grammatical structures in task-based language learning: The effects of recasts and prompts. *The Modern Language Journal*, 99(2), pp. 246-262.

Long, 1991; Skehan, 1996a). They agree that tasks should initially focus on meaning, but suggest that in addition there should also be attention to form.

One way to focus on form is to provide the student with corrective feedback (CF).

In recent years, the effects of implicit as well as explicit CF have been studied extensively (for a review, see Lyster, Saito & Sato, 2013; Schoorman & Schlak, 2012 and meta-analyses by Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007; Russel & Spada, 2006). Although these meta-analyses show that, in general, CF is beneficial to the acquisition of the target language, the issue is still, which form of CF is most effective under which conditions and for which linguistic structures.

Most studies comparing the effects of recasts and prompts (e.g., Ammar & Spada, 2006; Ellis, Loewen & Erlam, 2006; Loewen & Nabei, 2007; Lyster & Izquierdo, 2009; Yang & Lyster, 2010) have reported on the acquisition of linguistic features with which the students were already (partially) familiar. To our knowledge, no classroom studies on new linguistic features, which also differ in complexity and relatedness to the L1, exist. It is for this reason that we designed the present experiment. We investigated the effects of recasts and prompts on two *new* and *different* grammar structures in a task-based language learning environment.

1.1 Corrective feedback

One way to stimulate attention to formal structures during meaningful communication is the provision of corrective feedback (CF), defined by Lightbown and Spada (1999) as ‘any indication to the student that their use of the target structure is incorrect’ (p.171).

How CF contributes to language acquisition can be explained by both skill acquisition theories and theories on implicit learning. In addition, we will discuss Skehan’s (1996a, 1998a) dual mode system on information processing. Anderson (1993, 2000) considers the acquisition of a language akin to learning any other skill, with declarative knowledge turning gradually into procedural knowledge. In the case of language acquisition, this could mean that students memorize, for example, a grammar rule and through practice (proceduralization of knowledge) they are able to use the rule without thinking about it. Both Lyster (2004) and DeKeyser (2010) support practice of declarative knowledge through CF because of its ability to promote restructuring of the interlanguage. CF as a way to practice, however, can only be effective if embedded in communicative interaction and when students’ individual differences and the teaching context are taken into account (DeKeyser, 2010). Ellis (1993), supporter of theories of implicit learning, sees a more limited role for explicit knowledge. He points out that explicit knowledge as provided by CF may facilitate implicit learning in only two ways. First CF may contribute to the process of noticing: ‘[I]f students are armed with explicit knowledge of a linguistic feature, they are more likely to notice its occurrence in the communicative input they receive and thus to learn it implicitly’ (p. 149). Second, CF may assist noticing the gap because a student who possesses explicit knowledge of certain language structures will, probably, be able to notice the gap between their own incorrect and the correct target structure given by a teacher or peer.

An additional view to language acquisition is Skehan's (1996a, 1998a) dual mode system, which describes two ways of processing information. The *exemplar-based* system includes discrete lexical items as well as ready-made formulaic chunks of language while the *rule-based system* consists of abstract representations of the underlying patterns of the language. These systems cooperate and 'combine in a synergistic manner to yield results, and degrees of learning, that are more than simply the sum of the parts' (p. 43). Drawing on this dual model system, Skehan (1998a) argues that exemplar-based CF 'may not be so effective' because 'there is not the (...) connection with a rule which can produce general change' (p. 89). The rule-based system, however, 'is more likely to be more sensitive to feedback since the precision and system which accounts for rule-organization will make the feedback more informative' (p. 88). Nonetheless, achieving greater accuracy is not without consequences for other dimensions of the task performance. According to Skehan's (1996a, 1998a) Trade-off Hypothesis, learners possess limited attentional resources and a focus on, for instance, accuracy, may have a negative impact on fluency and/or complexity. Robinson (2001, 2005), on the other hand, disagreeing with this hypothesis, claims that people have multiple attentional resources, which are not in competition with each other but can be used at the same time.

Having reviewed three different theories on language processing, several key factors on how CF may promote language learning emerge: The feedback should be noticeable, create opportunities for practice (output), and trigger access to the rule-based system.

1.2 Recasts and prompts

A key concern is the extent to which the two CF types recasts and prompts promote noticing (the gap), create opportunities for practice, and trigger access to the rule-based knowledge system. As defined by Lyster and Ranta (1997), recasting is 'the teacher's reformulation of all or part of a student's utterance minus the error' (p. 46). Recasts are considered to be an implicit form of CF and are by far the most frequent form of negative feedback in classrooms of all kinds (Long, 2007). The second feedback type is prompting, which, according to Lyster and Ranta (1997), comes in various shapes and types: clarification requests, repetitions, metalinguistic feedback, and elicitations.

Based on L1 acquisition research that shows that children notice and use linguistic information (e.g., repeating the utterance), Doughty and Varela (2008) regard recasts as an ideal way to focus on form in the communicative classroom. In addition, it is also the ability of correcting a learner's mistake without breaking down the communicative flow (Long, 2007) that make recasts very suitable for meaning-focused TBLT. The effectiveness of recasts, however, is disputed. Nicholas, Lightbown, and Spada (2001), for instance, comment that the student may not always notice the corrective element of a recast. Similarly, Lyster (1998) found that students perceived recasts as confirmation of meaning rather than feedback on form. Another point of criticism is that students may not be aware of the exact location of the error and as a consequence may not notice it. Finally, De Bot (1996) comments

that no trace in memory is left by recasts, because the student is not actively involved but only listens to the input. He reasons that it is this low level of attention that generates only weak or no connections in memory, hence constraining restructuring of the interlanguage. In regard to these negative claims about recasts, Goo and MacKey (2013) recently pointed out some methodological and interpretive problems in the small number of studies on which the claims are based, including issues like form-focused instruction. They state that in some recast-versus-prompts studies form-focused instruction has been included as a part of the experimental treatment. They doubt whether the differential effects of prompts over recasts can be fully attributed to the feedback type because of the 'moderating role of the form-focused instruction' (p.152).

The ambiguous findings on the effectiveness of recasts have led to a comparison of other forms of CF, like prompts. In the majority of classroom studies following this line of investigation, prompts appeared to be more effective than recasts (Ammar, 2008; Ammar & Spada, 2006; Ellis et al., 2006; Loewen & Philp, 2006; Yang & Lyster, 2010). Factors such as noticeability (Schmidt, 1990), the ability to generate modified output (Swain, 2000, 2005), also considered as a way of practicing, and the metalinguistic feedback being rule-driven, may explain the superiority of prompts over recasts.

1.3 Previous research on corrective feedback

An ever-growing number of classroom and laboratory studies have compared the effects of recasts and prompts on L2 acquisition. In this review, we focus on studies that compared prompts and recasts by measuring student outcomes in terms of acquisition of linguistic structures. Thus, we exclude studies that measured uptake and repair and studies that focused on the acquisition of vocabulary (Dilans, 2010).

Before comparing recasts with prompts, we briefly summarize the findings of studies that investigated the effectiveness of recasts. Several studies have reviewed their effects (Ellis & Sheen, 2006; Long, 2007; Lyster & Saito, 2010; Miller & Pan, 2012; Nicholas et al., 2001) and conclude that recasts have proven their effectiveness in laboratory settings (Carroll & Swain, 1993; Egi, 2007, 2010; Han, 2002; Leeman, 2003; Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998). However, recasts were not effective, or only to a limited extent, in classroom settings (Dougherty & Varela, 1998; Ellis et al., 2006; Lyster, 2004).

In the majority of laboratory studies that compared recasts with prompts, no significant differences were found. Carroll and Swain (1993) examined the effects on the acquisition of English dative alternation of explicit correction, recasts and two types of prompts, and compared them with data obtained in a control group with no feedback. All of the treatment groups performed better than the control group; the group receiving explicit correction outperformed the other groups, but no significant differences were found between prompts and recasts. The same results are reported by McDonough (2007) and Lyster and Izquierdo (2009). They investigated the effects of recasts and prompts on the acquisition of, respectively, the simple past and progressive activity verbs in a Thai EFL context (Lyster & Izquierdo, 2009) and

grammatical gender by adult second language students of French (McDonough, 2007). Lyster and Izquierdo concluded that 'students receiving recasts benefited from the repeated exposure to positive exemplars as well as from the opportunities to infer negative evidence, whereas students receiving prompts benefited from the repeated exposure to negative evidence as well as from opportunities to produce modified output' (pp. 453–454). Nassaji (2009) compared recasts with elicitations in dyadic interactions between native-speaker English teachers and adult ESL students. His study favored recasts by demonstrating that these led to a higher percentage of immediate post-interaction correction than the elicitations did. In addition, explicitness seemed to be a key factor: For both recasts and elicitations, the more explicit forms led to more immediate and delayed post-interaction correction than the implicit forms.

These classroom studies can be categorized into studies with either young or adult students. Lyster (2004) investigated the effects of form-focused instruction (FFI) in combination with corrective feedback (recasts versus prompts) on 10–11-year-old immersion students' ability to accurately assign grammatical gender in French. Results of the written tasks in particular, and to a lesser degree the oral tasks, revealed that FFI is more effective when combined with prompts than with recasts or no feedback. Ammar and Spada (2006) also focused on younger students with their study including 12-year-old francophones as participants in an intensive ESL course. They examined the impact of recasts in comparison to prompts and no corrective feedback on students' acquisition of English third person determiners. On the oral task, both the recast and the prompt group outperformed the control group on the posttests. The delayed posttest results also revealed that the prompt group was even better than the recast group. On the written task, the prompt group outperformed the recast group on the immediate and delayed posttest. However, the investigators do not conclude that prompts are the ideal CF technique because analyses by proficiency level showed that low-performance students benefited more from prompts than recasts, whereas high-proficiency students benefited equally from both prompts and recasts. This was affirmed by Ammar's 2008 study, which was a secondary analysis of Ammar and Spada's (2006) earlier data. Comparisons of prompts and recasts showed that prompts may be more effective than recasts in leading to L2 morphosyntactic development, especially in the case of low-proficiency students. Results from the computerized task showed no differences between the groups in terms of accuracy. In addition, Mackey and Philp (1998) found that developmentally more advanced students benefit more from recasts than developmentally unready students in facilitating an increase in production of targeted higher-level morphosyntactic forms.

The participants in all other classroom studies were adult students of a second or foreign language. Ellis et al. (2006) reported on the effects on the acquisition of the past tense. The participants, 77% of whom were of East Asian origin, were low-intermediate students of L2 English. Implicit knowledge was measured by an oral imitation test, whereas explicit knowledge was measured by an untimed grammaticality judgment and a metalinguistic knowledge test. Although no significant differences were found on the immediate posttests, prompts outperformed recasts on the delayed imitation and grammaticality judgment posttests. In a similar study, Ellis

(2007) not only measured acquisition of the English past tense but also the acquisition of comparatives. It appeared that prompts were overall more effective than recasts, but more so for comparatives than for past tense. As a result, Ellis concluded that the effectiveness of the feedback depends on the grammatical structure. We shall address this point later in the article.

In studies by Loewen and colleagues (Loewen & Philp, 2006; Loewen & Nabei, 2007), no significant differences were found between prompts and recasts on accuracy. Nevertheless, Loewen and Philp found on the immediate posttests that prompts led to an accuracy rate of 75%, whereas recasts led to an accuracy rate of only 53%. Although prompts led to a higher success rate, Loewen and Philp found the achievements of the recast group encouraging for classroom teachers by suggesting that recasts are likely to be productive for students and have pedagogical benefits: They are considered 'less threatening to students' confidence and less intrusive to the flow of communication' (p. 551). The most recent study to date, conducted by Yang and Lyster (2010), concerns a quasi-experimental investigation held in China in EFL classrooms at university level. The study compared the effectiveness of recasts, prompts, and no feedback on the use of regular and irregular English past tense. The effects of prompts were larger than those of recasts in terms of increasing accuracy in the use of regular past tense, whereas prompts and recasts had similar effects on improving accuracy in the use of the irregular past tense.

The differences in findings among studies comparing recast and prompts in laboratory and classroom settings might lie in the fact that recasts are more salient in laboratory interactions than in the classroom. According to Nicholas et al. (2001) the laboratory setting makes the student more aware of the feedback being corrective and the fact that the feedback was restricted to only one or two features made it easier for the students to recognize what was intended by the feedback.

1.4 CF and linguistic structure

CF research shows that the effectiveness of CF depends heavily on the grammatical structure being investigated (Lyster et al., 2013). What works for one linguistic structure may not be effective for another one (Sheen, 2011). Although there is no agreement on what is meant by *type of structure*, the differences in target structures are often explained by factors as simple/complex, easy/difficult to learn, rule-based/exemplar-based, and L1–L2 relatedness.

The current body of research into what type of instruction is beneficial to which target structures shows no agreement. Hulstijn & De Graaff (1994), for example, claim that complex rules can be taught best in an explicit way and simple rules in an implicit way. They argue that learners are not able to notice complex structures in the input and that therefore explicit learning of the rule is required. In comparison, simple rules may be noticed more easily by the learner, which may lead to acquisition. Krashen (1981) takes the opposite point of view by claiming that only simple rules can be taught and that difficult rules are best learned in an implicit way. Regarding L1–L2 transfer, Andringa, De Glopper and Hacquebord (2011) found that

explicit instruction was more effective for simple structures if these structures were more or less similarly realized as in a student's L1.

To date, only two classroom studies have been designed to compare the effects of oral prompts and recasts on different target structures. Ellis (2007) compared the effects of recasts and prompts on the acquisition of the English past tense and comparatives. He considered the comparative to be more difficult than the past tense, because the rule for the comparative requires both morphologic and syntactic analysis, while the past tense only asks for a morphologic analysis. Ellis found that prompts were overall more effective than recasts, but more so for the 'difficult' comparative than for the more 'simple' past tense. Yang and Lyster (2010) also investigated the effects of recasts and prompts on two different structures: the regular and irregular English past tense. The prompt group performed significantly better than the control group on irregular past tense forms at the delayed oral posttest. In addition, the effects of prompts were larger than those of recasts on the oral production test of regular past tense forms. However, prompts and recasts had similar effects on the oral production test of irregular past tense forms. Drawing on Skehan's (1998a) dual knowledge system, they argued that prompts had more effect on regular past tense because this target structure is considered a rule-based feature (add *-ed* to the base form of a regular verb). They reasoned that prompts, more than recasts, trigger access to the rule-based system. Since there are no clear rules for irregular past tense forms, these are considered to be exemplar-based features.

So far, German language structures have received very little attention in empirical investigations of CF (Lochtman, 2002). To enlarge the scope of linguistic structures, in relation to the effectiveness of CF, we selected two different German linguistic structures for the current study.

1.5 The present study

The current study focuses on whether previous findings on the effects of prompts and recasts on (partially) known target structures also applies to two new German grammar structures that differ in complexity and relatedness to the L1. By investigating Skehan's (1996a, 1998a) Trade-off Hypothesis, the study also contributes to the literature on attentional resources in task-based language learning. Five research questions guided the study:

- RQ1 Do recasts have a positive effect on the accuracy of new grammar structures?
- RQ2 Do prompts, operationalized as metalinguistic feedback followed by elicitation, have a positive effect on the accuracy of new grammar structures?
- RQ3 Which type of CF is more effective?
- RQ4 Does the effectiveness depend on the targeted structure?
- RQ5 Does a student's focus on accuracy have a negative effect on oral fluency?

2. METHOD

2.1 *Participants*

Sixty-four ninth-grade students learning German as a foreign language participated in this study. They were recruited from three school groups at a Dutch secondary school. The majority of the participants (96%) were native speakers of Dutch. Apart from German, participants also learn English and French as foreign languages at school. The mean age was 14.3 years. The students had been engaged in learning German for 19 months at the A2 level of the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001), for 2 hours per week. In accordance with the protocol of the University of Amsterdam's Faculty of Humanities' Ethics Committee, all parents were informed about the studies and the possibility of non-participation.

2.2 *Target structures*

In line with our purpose to examine whether corrective feedback facilitates the learning of a completely *new* structure, we chose two German target structures with which the students at A2-level of the CEFR were not already familiar. Moreover, we wanted to examine the interaction of prompts and recasts with a complex and a simple German structure that also differ in their relatedness to the L1.

2.2.1 *Dative case after a two-way preposition (Dative)*

The target structure in Task 1 and Testing Task 1 concerns the dative case of an article after a two-way preposition in German (*in, an, auf, hinter, neben, unter, über, vor, zwischen*²). A two-way preposition may be used with a dative to indicate the current location; combined with an accusative it signals direction towards something or someone. In the present study, only the locative dative was used. In this example, *auf dem Bett liegt ein Kissen*³, the preposition *auf* demands the dative case and consequently the neuter definite article *das* changes into *dem*. We define this structure as complex because (a) learners have to undertake several analytical steps to arrive at the correct form (see Hulstijn & De Graaf, 1994, for complexity and difficulty of target structures). Specifically, learners are required to apply the syntactic rule: The article *der* of a masculine word changes into *dem*, the article *die* of a feminine word changes into *der*, the article *das* of a neutral word changes into *dem*, and the article *die* of a plural word changes into *den*. (b) Learners have no existing knowledge regarding the dative structure because it has no equivalents in Dutch. Transfer between L1 and L2 is therefore unlikely.

² Translation: *in, at, on, behind, next to, below, above, in front of, between.*

³ Translation: *on the bed (the) lies a pillow.*

2.2.2 Comparative

For Task 2 and Testing Task 2, comparatives were selected as the target structure. We consider this structure to be more simple than the dative structure because (a) most of the comparatives are formed by simply adding *-er* to the adjective or adverb; and (b) the forming of German comparatives is very similar to the forming of comparatives in the L1 (Dutch). Learners may apply the same rule, add *-er* to the adjective or adverb. In addition, German suppletive comparative forms (*besser* [better] – *mehr* [more]) show similarities to the L1 forms. However, adding an umlaut mark to a comparative form (*größer* [bigger] – *höher* [higher]) is not found in Dutch.

2.3 Design

The study comprises two interventions, where each consisted of a task with a focus on a particular language structure. The design entailed three testing sessions for each task. One week prior to the start of Task 1, the students performed two pretests: the oral test first and then the written test. One week after the intervention period (which lasted three weeks) students were engaged in the posttests, and three weeks later the students performed the delayed posttests. Task 2 (which lasted three weeks) started two weeks after Task 1 and followed the same pattern of measurements. Two intact classes were randomly assigned to the recast ($n = 20$) and prompt ($n = 21$) conditions. The control group ($n = 23$), an intact class, followed the regular textbook-based curriculum. These students did not receive experimental treatment. During the intervention periods, the students of the experimental groups worked 1 hour a week on the treatment tasks and during the remaining hour they read a book and practiced listening skills without any grammar instruction or feedback on grammar structures. Because the teachers were required to provide the students with approximately 1 or 2 feedback moments every week, they needed the whole hour to do so for 20 or 21 students. Sometimes, students did not need a full hour to work on the task and continued with reading activities. Then, the task materials remained on the table, which enabled the teacher to look at the material, ask questions, and provide feedback. The control group worked one hour in their text books, which included written exercises and structured oral dialogues that targeted both two-way prepositions plus the dative case and comparatives. During the other 60 minutes they performed the reading and listening activities.

2.3.1 Prompt group

As shown in Figure 1, the prompt was given by the teacher in two phases: (a) provision of metalinguistic feedback on the student's false utterance and (b) elicitation of the correct answer.

In order to ensure a valid comparison, none of the learner groups received instruction and practice on the new target structures (see Goo & Mackey, 2013, on instruction and prompts) prior to task performance. The students of the prompt group received information on the grammar structures during the first part of the

prompts (metalinguistic feedback), enabling them to respond to the teacher's request to moderate their output (elicitation).

St: In das Mitte steht ein Stuhl.[Error: Grammatical] [In the middle is a chair]
 T: Almost, it's not, in das Mitte. After the preposition 'in' follows the dative case. [FB: metalinguistic]
 T: Okay, try again. In d----- Mitte? [FB: Elicitation]
 St: In der Mitte

Note. St = student; T = teacher; FB = feedback

Figure 1. Example prompt on dative case after two-way preposition.

2.3.2 Recast group

The teacher of the recast group was asked to reformulate the student's false utterance, minus the error (see Figure 2).

St: In das Mitte steht ein Stuhl. [Error: grammatical] [In the middle is a chair]
 T: Ach so, in der Mitte steht ein Stuhl. [FB: recast]

Note. St = student; T = teacher; FB = feedback

Figure 2. Example recast on dative case after two-way preposition.

Like the students in the prompt group, the students in the recast group were not instructed in the new structures but received information on the correctness of the target structure through the positive example of the recast.

2.3.3 Control group

Because of school regulations, the control group could not perform the two oral tasks but were required to follow the standard curriculum of the course book, which was strongly focused on forms (for the distinction between focus-on-form and focus on forms, see Long, 1991). The two target structures, dative and comparative, were part of their curriculum, during the period of the interventions. They did not receive personal oral feedback on these grammar structures. According to the regular practice in those language classes, students were instructed on the target structures, performed both written exercises and structured oral dialogues in their workbooks and reviewed them with help from the correction model. At the end of the lesson, during a joint class moment, the teacher asked the students whether they had any questions about the exercises.

2.4 *Instruction materials and procedures*

2.4.1 *Task 1*

The students worked in pairs and designed the room of their dreams with a virtual budget of 10,000 euros. First, the students performed pre-task activities that equipped them with useful vocabulary for the task performance. The goal of the during-task phase for each pupil was to describe the room in triads. The two other students were not able to see the room and were asked to draw the described room on paper. Afterwards, they compared their drawings and posed questions on any furniture that was placed incorrectly. After finishing the task, the students presented their dream bedroom individually to a research assistant, who recorded their performance on video camera. This oral presentation was subsequently assessed by the teacher using a scale ranging from 1 to 10 (low to high, with 6 judged sufficient).

2.4.2 *Task 2*

The students were asked to participate as researchers in a German television program for consumers entitled *Deutschland testet* ('Germany tests'). They were given the assignment to compare two products. From the input in the pre-task, the students selected vocabulary that could be useful for task performance. The during-task phase was oriented toward an oral presentation of the comparison test that had been conducted in triads. Each student presented the comparison of his two products, and the two other group members took notes and explained, which product they wanted to buy. After finishing the task, the students presented their product comparisons individually to a research assistant, who recorded their performance on video camera. This oral presentation was assessed by the teacher using the same scale as in Task 1.

2.5 *Procedures*

The two experimental groups were taught by two different teachers because their classes were taught at the same time. The groups were randomly assigned to the teachers, who both hold a master's degree in teaching German language and literature and were trained in the feedback type they were to provide⁴. During the training session, both teachers observed examples of possible student failures and practiced with providing either recasts or prompts.

During the pre- and main task phase, the teachers of the intervention groups walked around the classroom to provide all students with one or two feedback moments per lesson (one hour). They looked at the students' design of their dream room (Task 1) or their product-comparison (Task 2), followed by questions, such as '*Wo steht das Bett in deinem Zimmer?*' (Where is the bed in your room?), '*Wo hängt das Poster in deinem Zimmer?*' (Where is the poster in your room?) regarding Task

⁴ *The teacher of the prompt group is the first author of this article.*

1 and 'Warum möchtest du ein I-phone kaufen? Gib mir drei Gründe' (Why do you want to buy an I-phone? Give me three reasons) regarding Task 2.

The control group was taught by a third teacher who also holds a master's degree in teaching German language and literature. The researcher received a copy of the lesson plans of the control group and visited the classes on a regular basis. As for the intervention groups, a colleague in the department regularly observed the lessons with the help of a feedback logbook in order to ascertain whether both teachers adhered to expectations regarding the provision of recasts and prompts and to take note of the number of feedback moments per student. The observations revealed no differences in treatment except in terms of feedback type. In each lesson, both teachers wrote down the number of feedback moments for each student. Each student of the recast group received a mean sum of 6.3 ($SD = 4.37$) recasts on the dative structure and 4.5 ($SD = 1.79$) on the comparative structure, spread out over three weeks. In short, approximately 2.1 recasts on the dative and 1.5 recasts on the comparative per hour. Students of the prompt group received a mean sum of 5.33 ($SD = 2.11$) prompts on the dative structure and 5.30 ($SD = .85$) on the comparative structure. In other words, 1.78 prompts on the dative and 1.77 on the comparative, per hour. Because students made less failures in the comparative than the dative structure, they received more feedback on the dative. No statistical significant differences in the number of feedback moments were found between conditions on either the dative ($F(1,39) = .83, p = .37$) nor the comparative structure ($F(1,39) = 3.71, p = .06$).

2.6 Testing materials and procedures

For the written tests a fill-in-the-gap test was chosen in order to measure monitored language of the German grammar structures; by comparison, the communicative oral testing tasks were intended to measure the students' implicit knowledge (Ellis, 2005a) and oral fluency. In order to avoid the retrieval of explicit knowledge on the structures, students performed the oral task first, followed by the written task. For the oral test on comparatives, no delayed posttests were administered in the control group because the students were not motivated to perform them. We chose not to force them because of possible negative influence on the reliability of the results.

2.6.1 Written accuracy tests

Following DeKeyser (1993) a fill-in-the-gap test was chosen to assess accuracy. Not only was this format most familiar to the students; it also forced them to provide a correct alternative rather than simply indicating the occurrence of an incorrect structure, a limitation of many grammaticality judgment tests. The written tests were administered by the teacher during the lessons. In both settings, students were given sufficient time, which turned out to be 15 minutes maximum. The test was administered in three versions: For the pretest, each participant received version A, for the posttest version B, and for the delayed posttest version C, thus all conditions received the same test at each measurement occasion. All written tests for the dative and comparative structure were piloted, prior to the intervention, in two 10th-grade

classes ($n = 48$) which did not take part in the intervention. Cronbach's alpha scores varied between .67 and .79, suggesting reasonable reliability. Finally, the different versions of the tests were randomly assigned to a measurement occasion.

The written test on the dative included 14 fill-in-the-gap sentences where students were to supply the correct case of an article after a two-way preposition that was used with its meaning of static location. In order to disentangle application of the dative rule from knowledge of noun gender, grammatical gender was noted in brackets after the noun (*Tisch (der)*) [table (the)]. Reliability (Cronbach's alpha) was .85 for the pretest, .95 for the immediate posttest, and .93 for the delayed posttest.

The written test on the comparative included eight fill-in-the-gap sentences where students were asked to provide the correct comparative form. Reliability (Cronbach's alpha) was .65 for the pretest, .81 for the posttest and .70 for the delayed posttest.

2.6.2 *Oral accuracy tests*

Oral accuracy in both tasks was assessed during a classroom period by means of an oral communication test held with one of the three interlocutors: two native speakers of German and the first author. The tests consisted of communicative tasks similar to those used in the treatment sessions. Once again, three parallel test versions were designed. The interlocutors, located in different rooms, had received instruction regarding test administration and followed a fixed protocol to ensure identical test administration. Each student was tested individually and audio-recorded. We avoided a nested structure where students were nested in interlocutors. Therefore at each measurement occasion a student was tested by another interlocutor. The recordings were rated for accuracy by a native speaker of German and the first author. The first author evaluated all recordings; the second rater evaluated 70%. For both tasks, inter-rater reliability was high: 95% for the dative and 90% for the comparative.

Knowledge of the dative was tested with a picture description test (Appendix A). The interlocutor used a 3-dimensional picture of a bedroom; the German noun plus its gender of the furniture and accessories in the picture were provided. Not only would this approach focus on students' grammatical over their lexical knowledge; it would also facilitate the occurrence in their production of the target structures. The interlocutors were asked to elicit a minimum of 15 target structures. Accuracy scores were expressed as percentages correct (correct use of dative case after a two-way preposition / total number of articles used after a two-way preposition). Use of prepositions other than the target structures was disregarded. Assessment focused on the repairs, not on the utterance before the repair.

For the comparison test, the interlocutor used a picture of two mobile phones that included a table indicating price, dimensions, weight, memory, internet speed, touchscreen resolution, and battery size for each mobile phone (Appendix B). Once again, the German noun forms were provided. Students were asked to compare nine features of the mobile phones. Accuracy scores were expressed as percentages of correct forms (correct use of comparatives / total number of comparatives used). Incorrect use was indicated when either the comparative form itself was incorrect or

when it was used incorrectly in context. When a student repaired, the repair was assessed and not the utterance before the repair. For the oral test on comparatives, no delayed posttests were administered in the control group because the students were not motivated to perform them. We did not force the issue in order to avoid a potentially negative influence on the reliability of the results.

2.6.3 *Oral fluency tests*

Fluency in the oral tests was evaluated by two raters, a native speaker of German and the first author. In order to achieve rating consistency, the raters listened to 15 preselected performances for each task at different levels in the scale. The fluency rating scale ranged from 1, very low fluency, to 5, very high fluency. Discrepancies were discussed until agreement was reached. The order of recordings was randomized and each sample was rated independently by each rater. The second rater evaluated 70 % of the recordings. Inter-rater reliability was high in both studies: 91% in the first experiment and 90% in the second.

2.7 *Statistical analysis*

The effects of the two variables (i.e., target structure and feedback type) were investigated through an immediate and a delayed posttest. Because of the correlational nature of the data (i.e., two types of feedback, two target structures, two posttests) a general linear mixed model (GLMM) statistical analysis was used that included fixed and random factors for each dependent measure. Test time (pre-test = M0, post-test 1 = M1, post-test 2 = M2), type of feedback, and target structure were entered in the model as fixed factors; the random factor was subjects. In addition, the model included three 2-way interactions and one 3-way interaction. An alpha level of .05 was used for all statistical tests.

3 RESULTS

3.1 *Written accuracy*

Table 1 presents the means and standard deviations resulting from the written accuracy tests. The results for written accuracy showed a main effect, for (a) feedback type ($F(2, 61) = 61.77, p < .001$), (b) time ($F(2, 304) = 64.81, p < .001$), and (c) structure ($F(1, 304) = 353.16, p < .001$). Pairwise comparisons showed that the recast group and prompt group outperformed the control group (recasts vs. control, $MD = .068$, 95 % confidence interval [CI] = 0.14, $p = .05$; prompts vs. control, $MD = .35$, 95% CI = .29, .42, $p < .001$). The prompt group outperformed the recast group ($MD = .29$, 95% CI = .22, .35, $p < .001$). The main effect for feedback type is moderated by target structure and time. The analyses revealed a significant feedback type/structure interaction ($F(2, 304) = 16.21, p < .001$), feedback type/time interaction ($F(4, 304) = 44.83, p < .001$), and feedback type/structure/time interaction ($F(4, 304) = 2.61, p = .036$.) The last interaction indicates that the effect of recasts and prompts differed according to structure over time (see Figure 3).

Table 1. Group means (proportions correct) and standard deviations for written accuracy tests for three measurement occasions

	Groups	M0		M1		M2	
		M	SD	M	SD	M	SD
Dative	Recasts	.03	.07	.09	.12	.07	.11
	Prompts	.07	.22	.72	.32	.50	.35
	Control	.09	.09	.09	.10	.09	.10
Comparative	Recasts	.41	.18	.52	.14	.47	.14
	Prompts	.42	.16	.85	.14	.73	.04
	Control	.31	.15	.30	.15	.29	.19

Note. Recast group $n = 20$; Prompt group $n = 21$; Control group $n = 23$.

With regard to the effect of prompts on the dative, we see a very large increase between M0 and M1 ($d = 2.42$) and a decrease after M1 ($d = -.67$). For the effect of prompts on the comparative we also see a very large increase between M0 and M1 ($d = 2.89$) and a decrease after M1 ($d = -1.16$). With regard to the effects of recasts on the dative, there is an increase between M0 to M1 ($d = .58$) and a small decrease after M1 ($d = -.16$). For the comparative, as well, there is an increase between M0 and M1 ($d = .68$) and a decrease after M1 ($d = -.34$). The interaction effect shows that, compared to the prompts, recasts had a significantly larger effect on the comparative than on the dative ($t(304) = -3,13, p = .002$).

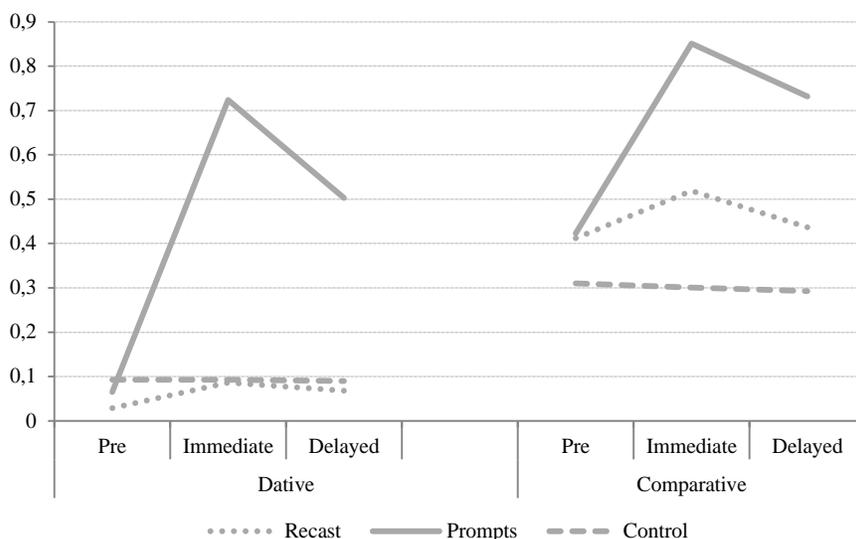


Figure 3. Interaction effects target structure /feedback type/ time for written accuracy.

3.2 Oral accuracy

Table 2 displays the means and standard deviations of the oral accuracy tests. The results for oral accuracy showed main effects for (a) feedback type ($F(2,66) = 27.51$, $p < .001$), (b) time ($F(2,284) = 31.05$, $p < .001$), and (c) structure ($F(1,284) = 210.99$, $p < .001$). Pairwise comparisons showed that the recast group and prompt group outperformed the control group (recasts vs. control, $MD = .097$, 95 % confidence interval [CI] = .023, .17, $p = .011$; prompts vs. control, $MD = .302$, 95% CI = .23, .38, $p < .001$). The prompt group also outperformed the recast group ($MD = .205$, 95% CI = .13, .28, $p < .001$). The main effect for feedback type is moderated by time. A significant feedback type/time interaction ($F(4,284) = 15.4$, $p < .001$) indicates that the effect of prompts and recasts differed over time. Reviewing the effect of prompts on the dative, we see an increase between M0 and M1 ($d = 1.77$) and a decrease after M1 ($d = -.11$). For the effect of prompts on the comparative we see the same pattern: an increase between M0 and M1 ($d = 1.66$) and a decrease after M1 ($d = -.56$).

Observing the effects of recasts on the dative, we see an increase between M0 to M1 ($d = .81$) and a decrease after M1 ($d = -.30$). Regarding the effects of recasts on the comparative we observe an increase between M0 and M1 ($d = .22$) and an increase after M1 ($d = -.22$). The intervention effect for feedback type was not moderated by structure. In other words, no structure/feedback type interaction ($F(2, 284) = 1.57$, $p = .209$) was found.

Table 2. Group means (proportions correct) and standard deviations for oral accuracy tests for three measurement occasions

	Groups	M0		M1		M2	
		M	SD	M	SD	M	SD
Dative	Recasts	.07	.16	.21	.19	.16	.17
	Prompts	.06	.13	.52	.34	.49	.34
	Control	.03	.05	.06	.12	.04	.06
Comparative	Recasts	.38	.21	.44	.30	.49	.21
	Prompts	.40	.29	.81	.18	.69	.24
	Control	.43	.22	.43	.19	—	—*

* For the oral test on comparatives, no delayed posttests were administered
 Note. Recast group $n = 20$; Prompt group $n = 21$; Control group $n = 23$.

3.3 Oral fluency

The results for oral fluency showed a main effect for (a) feedback type ($F(2,62) = 5.3$, $p = .007$), (b) time ($F(2,259) = 54.86$, $p < .001$), and (c) structure ($F(1,259) = 4.46$, $p = .036$). The analyses also revealed a significant structure/feedback type interaction ($F(2,259) = 3.09$, $p = .047$). This suggests that the effect of recasts and prompts on fluency differed according to structure: The recast group performed the dative task more fluently than the prompt group did, while for the comparative task

no differences were observed. The analyses revealed no significant differences for either feedback type/time interaction $F(3,259) = .216, p = .885$ or structure/feedback type/time interaction $F(3,259) = .540, p = .655$.

Table 3. Group means (1–5) and standard deviations for oral fluency tests for three measurement occasions

	Groups	M0		M1		M2	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dative	Recasts	2.70	.87	3.30	.66	3.70	.73
	Prompts	2.24	.89	2.76	.70	3.14	.73
	Control	2.13	.76	2.74	.81	2.75	.72
Comparative	Recasts	2.35	.88	3.05	.76	3.20	.69
	Prompts	2.10	.70	2.91	.77	3.33	.91
	Control	1.96	.83	2.57	.59	—	—*

* For the oral test on comparatives, no delayed posttests were administered
Note. Recast group $n = 20$; Prompt group $n = 21$; Control group $n = 23$.

3.4 Accuracy versus fluency

The means and standard deviations for oral fluency appear in Table 3. To examine Skehan's (1996a, 1998a) Trade-off Hypothesis, correlation coefficients (Pearson's r) were calculated between oral accuracy and oral fluency. For the dative after preposition task, a significant negative correlation between accuracy and fluency could be ascertained only for the prompt group in the delayed posttest ($r = -.518, p = .016$), not for the recast group ($r = -.436, p = .055$). For the comparative task no significant correlations were found between oral accuracy and oral fluency.

4 DISCUSSION

The present study investigated the effect of (a) recasts and prompts on secondary school students' accuracy of two new and different grammar structures, (b) prompts and recasts with two kinds of target structures that differed in difficulty and their relatedness to the L1, and (c) on oral fluency of a focus on accuracy.

The results demonstrate that recasts can affect the accurate use of new grammar structures. Results of both the oral and written tests show that the recast group performed significantly better than the control group. Lyster and Izquierdo (2009) concluded in their study that the recast group either benefited from repeated exposure to positive exemplars of the grammar structure and/or was able to compare the recast and their own incorrect utterance, in which case the recasts served as negative evidence. Since the recast group was not provided with instruction on the two target structures, it would appear that the students mostly benefited from positive exemplars. The combination of hearing positive exemplars of the grammar structures via recasts and the students' use of the structures during the oral presentations in the

classroom may have resulted in oral proficiency gains. This might also explain the difference in effect sizes for the oral and written tests: From pre- to posttest 2, the mean effect size for the written tests for both structures was small ($d = .41$) and for the oral tests medium ($d = .54$).

Besides the exposure to positive and negative exemplars, the effectiveness of the recasts might also derive from the fact that the recasts provided were short and consisted of only one corrective change (Goo & Mackey, 2013). This may have added to their saliency while making few demands on cognitive capacity. Even without providing overt language-focused instruction, then, recasts proved to be effective in enabling learners to acquire new grammatical structures.

RQ 2 addressed whether prompts had an effect on the accuracy of newly acquired grammatical structures. This question, too, can be answered in the affirmative. On both oral and written tests, the prompt group outperformed the control and the recast group. Disaggregating that effect was the focus of RQ3. Specifically, prompts were more effective than recasts in promoting the accurate use of new grammar structures. This finding is in line with earlier studies (Ammar, 2008; Ammar & Spada, 2006; Ellis et al., 2006; Loewen & Philp, 2006; Lyster, 2004; Yang & Lyster, 2010). Prompts may be superior because they are more salient to the students and are therefore noticed more easily. Another possible explanation for their superiority is that their first part, the metalinguistic feedback, triggers access to the rule-based knowledge system. The student makes a connection with a rule and is therefore able to apply the rule in a different setting. The second part of the prompt, the elicitation, might also trigger learning because students are asked to engage in self-repair (Swain, 1995). Lyster (2004) characterizes this as an opportunity for students to practice, which could lead to restructuring of the interlanguage. In line with De Bot (1996), we observed that the elicitation led to students' active involvement in the process, with a high level of attention. This active role can generate strong connections in memory and therefore may lead to subsequent learning.

Learner age and performance level may also be a contributing factor. The 14-year-old participants in the study performed on a low intermediate level and might not have been able to benefit from recasts as much as developmentally more advanced students (Mackey & Philp, 1998). At the same time, students were eager and willing to respond to prompts, reporting that it was motivating to get the teacher's attention and that they 'learnt a lot' during this moment of attention. In line with Skehan (1998a), we conclude that conscious awareness of the feedback is crucial, because appreciation for receiving corrections and the required transformation into new output 'predispose the student towards a rule-based perspective which is more likely to lead to longer-term change' (p. 57).

The development of a rule-based perspective might also explain the differences in mean effect sizes between the oral and written tests for the prompt group from pre- to posttest 2. The mean effect size for both structures of the written tests ($d = 2.13$) was more than one and a half times larger than the mean effect size of the oral tests ($d = 1.35$). In contrast to recasts, prompts might have stimulated the acquisition of explicit knowledge, which the written test, as compared to the oral tests, tapped into particularly well.

RQ 4 investigated whether the effectiveness of recasts and prompts depends on the target structure. An affirmative answer would require significant interaction between feedback type and structure. Indeed, we found such an interaction effect on written accuracy, which indicates that the difference in effect on the dative and comparative structure was larger for the recast than for the prompt group. Compared to prompts, recasts had a larger effect on the comparative structure than on the dative structure, a difference that may be attributable to the nature of the target structures. The dative structure is considered to be a complex, syntactic, and rule-based feature, with no similarities to the students' L1. The recast group was not informed on the rules of the dative structure during the feedback moment and, as a consequence, lacked the kind of explicit knowledge that could have facilitated performance on the written test. It would have been difficult for learners to deduce rules from the input because complex syntactic structures appear to be difficult to notice (Hulstijn & De Graaff, 1994).

That recasts were more successful in promoting the comparatives is plausible because the students already possessed explicit knowledge on the forming of the comparative from their L1. This facilitated applying the rule as well as recognizing the correct form in the recast and comparing it with their own incorrect utterance. These results support Long's view (2007) that recasts serve the acquisition of new linguistic structures that bear meaning and are noticeable. It appears that the structure's relatedness to the L1 also enhances its salience and its connection to meaning.

Another interaction between group and structure was found for oral fluency. The recast group performed the dative task more fluently than the prompt group did, while for the comparative task no differences were observed. Presumably, the students of the recast group did not think about a rule while carrying out the dative task and because of this lack of focus on accuracy, they freed up attentional resources that could be devoted to fluency. The prompt group, on the other hand, may have attempted to retrieve the rule from long-term memory in order to produce the dative form correctly.

Moreover, within the prompt group we found a trade-off between accuracy and fluency, the focus of RQ5. When students' accuracy and fluency scores were correlated, results indicated a negative correlation for the prompt group in the delayed oral posttest of the dative structure. In other words, students who performed the task more accurately showed less fluency, a result that seems to confirm Skehan's (1996a, 1998a) Trade-off Hypothesis, which claims that increased attention to accuracy will lead to decreased fluency and vice versa. By contrast, for the morphologically simpler comparative that is also similar to the L1, no negative correlation between accuracy and fluency was found, suggesting that, under specific conditions, attentional resources are not in competition with each other.

4.1 Limitations and future research

Despite the positive findings of the present study, certain limitations need to be considered. First, the students of the control group followed the regular curriculum in which the target structures of the present study were embedded through both written

exercises and structured oral dialogues. Although the students of the control group did practice the oral use of the target structures through structured dialogues, they did not participate in the same communicative tasks as the experimental groups did. This might have disadvantaged them on the oral tests. It may also explain why students in the control group were not motivated to participate in the delayed oral test on the comparatives. They stated that they were tired of participating in the tests, while their curriculum showed no differences with what they were used to. In future studies on CF it might be advisable to let the control group participate in the treatment tasks, though without providing CF (see also Lyster et al., 2013).

The oral accuracy test for the comparative (i.e., ‘compare two mobile phones’) might also have limitations in that it did not always elicit the intended target structures: For example, some students used the same target structure for different elements, resulting in incorrect as well as correct utterances. We counted and judged these as different utterances.

In the current study we made use of human raters to measure students’ fluency. Although research (see De Jong et al., 2013) demonstrates that several aspects of human-rated fluency (e.g., speech rate and pausing) correlate with more objective measures of fluency, we cannot rule out the possibility that the human raters’ judgment for all conditions was influenced by students’ accuracy or accent.

Although the number of feedback moments did not differ between the two experimental conditions, the variation within the condition for this measure was sometimes quite large, indicating that some students received more feedback than others. We did not find a relation between the number of feedback moments and learning gain, but nevertheless, in a subsequent study we would control for the amount of feedback moments per hour to ensure equivalence between participants. Obviously, because the study measured the effects of CF on only two different German target structures, additional research is needed to specify more precisely which linguistic structures benefit from which type of CF and in what way the L1 may contribute to this process.

5 CONCLUSIONS

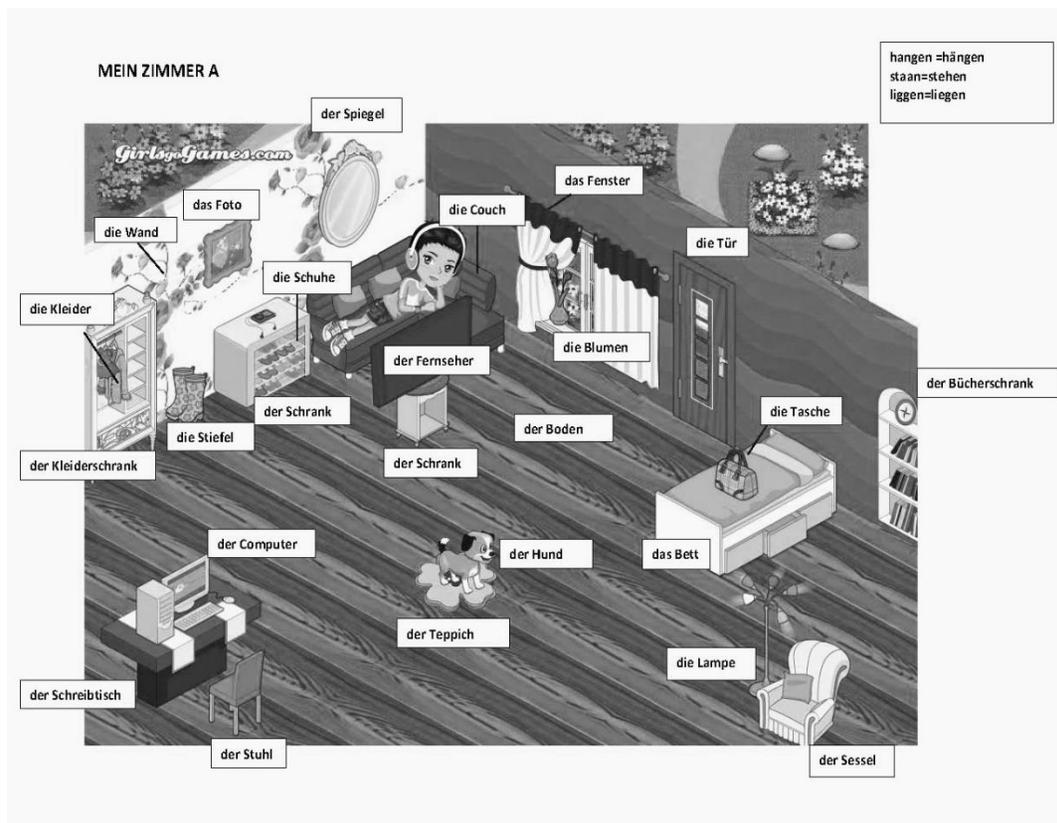
To summarize, the results of the two experiments show that both recasts and prompts contribute to the accurate use of *new* structures, with prompts being more effective than recasts. They confirm earlier research on the effects of recasts versus prompts and contribute to a more nuanced understanding of the effect of feedback in relation to the difficulty and L1-relatedness of the linguistic structure: There exists an interaction effect between structure and feedback type for written accuracy and oral fluency. For written accuracy recasts, as compared to prompts, had a larger effect on the comparative than on the dative structure. For oral fluency, results showed that the recast group performed the dative task more fluently than the prompt group.

It seems appropriate, then, to conclude that prompts may be used in the classroom effectively for both complex as well as simple rules. But because prompts may interrupt the flow of communication quite obtrusively, correction of simple rules, which are related to the L1, may also be done effectively through recasts. The struc-

ture's relatedness to the L1 makes it easier for the students to notice the recasts, which may enable them to compare the target-like structure with their own false utterance and promote acquisition. Finally, with a complex syntactic structure it appears that attention for accuracy comes at the expense of fluency.

APPENDIX A

EXAMPLE ORAL ACCURACY AND FLUENCY TEST ON THE DATIVE



APPENDIX B

EXAMPLE ORAL ACCURACY AND FLUENCY TEST
ON THE COMPARATIVE

1. Vergleiche Preis, Größe, Speicher und so weiter von diesen zwei Handys miteinander.
2. Welches Handy würdest du kaufen und warum?



	Blackberry Bold 9900	iPhone 4S
Preis (price)	€549	€695
Länge (length)	115 mm lang	116 mm lang
Breite (width)	66 mm breit	59 mm breit
Dicke (thickness)	12 mm dick	10 mm dick
Gewicht (weight)	131 Gramm schwer	145 Gramm schwer
Touchscreen	Ja, -960×640 Pixel Auflösung	Ja, 640 x 480 Pixel Auflösung
Batteriedauer battery length)	5 Stunden	6 Stunden
Speicher (memory)	12 GB groß	14 GB groß
Downloadgeschwindigkeit (Download speed)	16.5 Megabit pro Sekunde	15.4 Megabit pro Sekunde

CHAPTER 5

FOCUS ON FORM THROUGH TASK REPETITION¹

Because there has been little research on focus on form during the post-task phase in task-based language teaching, this experimental study investigates the effects of task repetition after having directed learners' attention to form during the main task. The study comprises two interventions, where each consisted of a task with a focus on a particular language structure. Forty-eight ninth-grade students learning German as a foreign language were randomly assigned to two conditions: one group repeated a similar task (R); the other group did not (NR). The first intervention targeted the German dative case after a preposition; the second German comparatives. Pre-tests, immediate and delayed post-tests included metalinguistic knowledge, written and oral accuracy as well as oral fluency. Results showed that on written accuracy and metalinguistic knowledge, the R condition outperformed the NR condition on both structures. No statistical significant differences between conditions were found on oral accuracy.

1. INTRODUCTION

Over the last thirty years, task-based language teaching (TBLT) has become a field of interest to SLA researchers, curriculum designers, language teachers, and teacher educators (Ellis, 2003; Shehadeh & Coombe, 2012; Skehan, 1998a; 2014; Van den Branden, Bygate & Norris, 2009; Willis, 2006). Several researchers, however, have expressed their concerns whether TBLT is able to teach specific grammar forms to learners of second and foreign languages (Sato, 2011; Swan, 2005). Ellis (2009a) responded to this criticism by stating that, although TBLT may not have an explicit grammar syllabus, all supporters of TBLT see a role for grammar in terms of a focus on form during meaningful communication.

Several studies have made learners focus on form at all three stages of the TBLT framework. According to this framework learners start the TBLT sequence by carrying out a pre-task, subsequently perform the main task and then optionally perform some post-task activities (Skehan, 1996a; Willis, 1996). Strategies such as guided planning (Foster & Skehan, 1999), different forms of input enhancement (see Doughty & Williams, 1998), and modeling (Kim, 2013) have been used to make learners attend to form in the pre-task stage. A focus on form in the during-task stage is mostly achieved through the provision of corrective feedback (see Lyster, Saito & Sato, 2013). In the post-task researchers have experimented with the anticipation of a public performance (Foster & Skehan, 2013) and post-task transcription of the oral performance (Qian, 2014). In this paper we will investigate another post-

¹ This chapter is an adapted version of an article that will be published as: Van de Guchte, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (in press). Focus on form through task repetition in TBLT. *Language Teaching Research*.

task strategy that may predispose learners' attention to accuracy, namely the repetition of a task.

Task repetition has proven to positively change learners' task performance in terms of complexity, accuracy and fluency (Ahmadian & Tavakoli, 2011; Bygate, 1996, 2001; Gass, Mackey, Alvarez-Torres, & Fernandez-Garcia, 1999; Lynch & McLean, 2000, 2001). Since this paper investigates the effects of task repetition on accuracy measures, we are especially interested in accuracy effects. Following Bygate (1996), we argue that giving learners the possibility to perform a task again, may lead to gains in oral accuracy. That is to say, because learners are familiar with the content through the first performance, the second time they may pay more attention to its correct formulation.

In most studies task repetition is operationalized as repetition of the same task (Bygate, 2001). In a pilot-study on same task repetition we observed, however, that learners found it 'boring' to repeat the task and were not motivated to perform them. Therefore, in the present study we opted for repetition of a similar task, which was another version of the initial task. Since several studies have demonstrated that the beneficial effects of task repetition do not carry over to a new task (Bygate, 2001; Gass et al., 1999), Ahmadian (2012) suggested that learners may need feedback on their initial performance (Sheppard, cited in Ellis 2009b) for effects to carry over to new contexts. Therefore, in the present study learners were provided with feedback at the during-task phase.

Both Swan (2005) and Sato (2010) question whether TBLT is successful in teaching learners predetermined grammar structures. For that reason, the present study did not use overall measures of accuracy (occurrence of errors per t-unit, Bygate 2001) but investigated the effects of similar task repetition on the accurate use of two different German grammar structures. In addition, we examined whether learners' focus on accuracy would come at cost of fluency (see Trade-off Hypothesis, Skehan, 1996a, 1998a)

1.1 Theoretical and empirical background

Why would task repetition be an effective strategy to make learners focus on form and promote the correct use of grammar structures? Based on theories on learners having limited attentional resources, Skehan (1996a, 1998a) argues that the first time learners carry out a task they are so preoccupied with completing the task that they focus almost completely on the content and not necessarily on linguistic accuracy. Levelt (1989) defines this process as the conceptualization stage in which learners perform different kinds of mental activities, such as selecting and ordering information and thinking about 'what will I say?'. It is not until the formulation stage that learners try to find the right words, grammar rules, and 'phonetic plan' (p. 12) to bring their message across. According to Bygate (2001), it is exactly during task repetition that learners can redistribute their focus from conceptualization to formulation (words and grammar) and articulation (pronunciation) because they are already familiar with the content and language knowledge.

One of the earliest studies on task repetition was carried out by Bygate (1996). In his small-scale laboratory study, he showed one learner of English a Tom and Jerry cartoon and asked to re-tell the story. Three days later, the student was asked to perform the same task again. Analysis of the oral output showed changes in complexity and grammatical variety and small gains in accuracy. In Bygate's (2001) second study he investigated the effect of learners repeating a task versus learners practicing a type of task. Forty-eight ESL students carried out a narrative and an interview, twice with ten weeks in between. Results showed a strong effect for task repetition in increasing fluency and complexity but not on accuracy. This effect did not carry over to a new task. Bygate commented, however, that some results showed a trend in that direction and he concludes that the 'notion of 'discourse competence' - the capacity to process certain types of discourse more easily than others - does appear to have some empirically identifiable psychological reality...' (p. 43).

Gass et al. (1999) investigated whether task repetition led to more sophisticated language use and whether this would transfer to a new context. One hundred and three students of Spanish as a foreign language were assigned to a control, same content, or different content group. They watched video fragments of Mr. Bean and recorded their own version of what was happening, at the same time. Gass et al. observed that task repetition led to greater overall proficiency, greater accuracy in morphosyntax (trend towards more target-like production of the verb *estar*), and lexical sophistication. Gass et al. reasoned that '... freeing up attention to meaning allows learners to gain greater control over their linguistic knowledge' (p. 573). These results, though, did not transfer to a new context.

A variation on task repetition, task recycling was investigated by Lynch and McLean (2001). They asked 14 oncologist students, learning English for specific purposes, to present their research six times in a poster presentation carousel. In contrast to other repetition studies with larger intervals, repetition in this study was immediate with an interval of 3 minutes. Findings revealed that participants' speech was more accurate and fluent.

Sheppard's (2006) study is the only study that demonstrated that effects of repetition can transfer to new contexts. In addition to task repetition, Japanese students were provided with feedback on their initial performance. After repetition students advanced in fluency, complexity, and clearly in accuracy. Transfer effects were observed for grammatical complexity.

Hawkes' (2012) study is the only one that investigated the effects of task repetition after having directed learners to form in the main task. After having carried out the main task, 13-14 year old Japanese students studying English as a foreign language focused on form by highlighting and practicing different target structures and vocabulary. During the final phase participants repeated the exact same task as the main task. Results showed an increase in the number of form and pronunciation-focused corrections. From these findings Hawkes concluded that students paid more attention to form during the repeated task performance.

1.2 *The present study*

Drawing on theories on limited attentional resources and Levelt's (1991) Model of Speech we designed an experimental study which investigates the effects of similar task repetition, preceded by rule deduction and form-focused feedback during the main task, on the accurate use of two German grammar structures. The design incorporated two subsequent interventions with each a different German grammar structure because we attempted to examine whether the effects of task repetition would depend on the target structure. In addition, this study enabled us to investigate Skehan's (1996a, 1998a) Trade-off Hypothesis.

To this end, three research questions were formulated:

- RQ1 Does task repetition as a post-task activity promote the accurate use of two new grammar structures?
- RQ2 Does the effect depend on the grammar structure involved?
- RQ3 Does students' focus on oral accuracy have a negative effect on oral fluency?

2. METHOD

2.1 *Participants*

The study involved forty-eight ninth-grade students learning German as a foreign language (A2 level of the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001)). Participants were recruited from two classes of a Dutch secondary school. The majority of the participants (95%) were native speakers of Dutch. Apart from German, participants also learn English and French as foreign languages at school. Their mean age was 14,2 years. The students had been engaged in learning German for 18 months, for 2 hours per week. In accordance with the protocol of the University's Faculty of Humanities' Ethics Committee, all parents were informed about the study and the possibility of non-participation. No parent objected.

2.2 *Design*

In this study we carried out an experimental pre-post-delayed posttest research design for two subsequent interventions (see Figure 1). Both interventions examined the effects of task repetition but each on a task with a different target structure. Each intervention was spread out over a period of three weeks. We performed a randomized experiment in two intact classes, each taught by a different teacher. Within each class students were randomly assigned to the Repetition (R) ($n = 24$) and No-Repetition (NR) ($n = 24$) condition. Thus, in each class were students of the R and the NR condition. The operationalization of the two conditions was equal for both interventions. All students received form-focused feedback on the particular target structure during main task preparations. Two weeks later, the R condition repeated a similar task, the NR condition did not and instead performed a filler task, which was

not related to the target structure. The performance of both the repetition and filler task, including preparations, took students approximately 25 minutes. To avoid that students of both conditions would cooperate, the Repetition group was placed at the right part of the classroom and the No- Repetition group at the left part. In addition, lesson materials were designed for independent working.

Because of earlier negative experiences in a pilot study with students immediately repeating the same task afterwards (students were bored and not engaged) we altered the timing and slightly the content of the post-task. The Repetition Task did not immediately follow the performed Main Task, but was performed 2 weeks later as a post-task of the subsequent Main Task (see Figure 1). As a result, the content of the repetition tasks needed to be adjusted to the topic of the subsequent main tasks. For example, instead of asking students to describe their dream bedroom at home (Main Task 1) they were required to describe their dream bedroom on a winter holiday (Repetition-Task 1). Although some adjustments to the content of the repetition tasks were made, the linguistic demands and grammar structures were the same as in the main tasks. For that reason, the main and repetition tasks can be regarded as ‘task families’, where a group of tasks resemble one another and may well have similar language or cognitive demands’ (Skehan, 1996a, p. 56).

The tasks, designed for this study, were embedded in an ongoing school program where the first author is active as a teacher of German. As part of the school curriculum, students’ performances of the tasks, needed to be assessed with the school grade scale ranging from 1 to 10 (where 1 is low, 10 is high and 6 is sufficient).

	Condition Repetition (R)	Condition No-Repetition (NR)
Wk 1	Pretest: Dative	
Wk 2	Main Task 1 (dative), plus feedback on target structure	
Wk 3	Posttest-1 Dative (measuring the effects of Main Task 1) Pretest Comparative	
Wk 4	Main Task 2 (Comparative) plus feedback on the target structure Repetition Task 1 (Dative)	Filler Task 1
Wk 5	Posttest-1 Comparative (measuring the effects of Main Task 2) Posttest-2 Dative (measuring the effect of Repetition 1)	
Wk 6	Main Task 3 Repetition Task 2 (Comparative)	Filler Task 2
Wk 7	Posttest-2 Comparative (measuring the effects of Repetition Task 2) Posttest-3 Dative (measuring the delayed effects of Repetition Task 1)	
Wk 9	Posttest-3 Comparative (measuring the delayed effects of Repetition Task 2)	

Figure 1. Research design repetition task.

2.3 Target structures

We selected two tasks with each a different target structure. We opted for a task with a simple target structure with similarities to the L1, the comparative, and for a task with a more complex target structure with no similarities to the L1, the dative structure.

2.3.1 Dative case after a two-way preposition

The target structure in Main Task 1 and Repetition Task 1 concerned the dative case of an article after a two-way preposition (*in, an, auf, hinter, neben, unter, über, vor, zwischen*²). We consider this structure complex because (a) learners have to undertake several analytical steps to arrive at the correct form (see Hulstijn & De Graaf, 1994); and (b) learners have no existing knowledge regarding the dative structure because it has no equivalents in Dutch. Transfer between L1 and L2 is therefore unlikely.

2.3.2 Comparative

In contrast to the dative case, we consider German comparatives to be a more simple structure because (a) most comparative forms in German are composed by a simple rule: Attach *-er* to the adjective or adverb. In addition, some German irregular comparative forms get an umlaut mark (*größer* [bigger] – *höher* [higher]) or have suppletive forms (*besser* [better] – *mehr* [more]); and (b) the forming of German comparatives is realized similarly as comparatives in the L1 (Dutch).

2.4 Treatment tasks

2.4.1 Main task 1- room of my dreams

This task is a description task for which each student was asked to design and describe the bedroom of his or her dreams with a virtual budget of 10,000 euros (adapted from Van de Guchte, Braaksma, Rijlaarsdam & Bimmel, 2015a). To increase the chance that the target structure, a two-way preposition plus dative case, would be used students were required to describe the place of the furniture and accessories (minimal amount of 15 pieces) in the room to two other students. These other students were not able to see the room and were asked to draw the described room on paper. Afterwards, they compared their drawings and posed questions on any wrongly positioned furniture. After finishing the task, each student presented his or her dream bedroom individually to a teacher trainee who recorded their performance on video camera. This oral presentation was assessed afterwards by the teacher with the help of a grade scale ranging from 1 to 10.

² Translation: *in, at, on, behind, next to, below, above, in front of, between.*

2.4.2 *Main task 2 – compare two ski holidays*

Main task 2 is a comparison task for which students compared two ski holidays in Austria. The use of the target structure comparatives was triggered by the additional instruction that students needed to convince their parents in an oral discussion with at least 7 arguments to go to either the quiet St. Johann or the more vibrant St. Anton. The information that was needed to compare the two holiday destinations was put in a written structured chart. The elements that students compared in the chart were e.g. the number of ski lifts, the price of a ski pass, the altitude of the area, measurements of the apartment, the temperature, quality of the apartment, beauty of the area and kindness of the staff. After finishing this task, students presented their seven arguments individually to a research assistant, who recorded their performance on video camera. This oral presentation was assessed afterwards.

2.4.3 *Repetition task 1 – holiday room of my dreams*

All students in the R condition were told to design the bedroom of their dreams for a ski holiday at the destination of their choice and make a phone call to a reservations center to find such a room. During this phone call (with a teacher or teacher assistant) students were required to describe the place of the furniture and accessories in the room.

2.4.4 *Main task 3 - summer memories*

Main Task 3 was designed to embed Repetition Task 2 as a post-task. This task did not take part in the measurements. For this task all students were asked to bring a holiday picture into the class and tell about the picture and the holiday during a speed date session with their fellow students.

2.4.5 *Repetition task 2 - compare summer holidays*

Students were asked to compare aloud two summer holidays in Italy and convince their friend in an oral discussion with at least 7 arguments to go to either one of them. The information that was needed to compare the two holiday destinations, was put in a written structured chart. The elements that students compared in the chart were, in a few cases, adjusted to the topic ‘summer holiday’ but elicited the exact same comparative forms as those in Main Task 2. Students compared, for example, the number of swimming pools, the price of the flight, the altitude of the area, measurements of the apartment, the temperature, quality of the apartment, beauty of the area, and kindness of the staff.

2.5 *Procedures*

The two groups were taught by two different teachers. Because the lessons for the two groups were scheduled in parallel it was impossible to have both groups taught

by the same teacher. The groups were randomly assigned to the teachers. Both teachers held a master's degree in teaching German language and literature and were trained in providing the feedback type prompts by means of a protocol and a training session.

2.5.1 *Focus on form during the main task*

All students' attention was directed to form during main task preparations by means of rule deduction and corrective feedback. After the pre-task, in which students were shown both oral and written examples of the upcoming tasks and gathered useful vocabulary, students were asked to discover the grammar rule, which could help them carrying out the main task. We opted for instruction at the beginning of the main task, because after having read and heard the language within a communicative context, students would be able to understand the relationship between communicative meaning and linguistic form (Nunan, 2004).

Prompts were selected as the type of feedback students received, because several studies have demonstrated their effectiveness in fostering the acquisition of grammar structures (Ammar, 2008; Ellis, 2007; Van de Guchte et al., 2015a; Yang & Lyster, 2010). When students made an error in one of the target structures, they first received metalinguistic information on the target structure, followed by an elicitation part in which they were asked to answer again (see example in Figure 2). Both teachers wrote down the number of feedback moments per student and were observed by a research assistant who also noted the amount of feedback for each student. The mean sum of feedback moments for the R condition on the dative structure of Task 1 was 2.88 ($SD = .68$) and on the comparative structure of Task 2, 2.08 ($SD = .41$). For the NR condition the mean sum of feedback moments on the dative structure of Task 1 was 3.04 ($SD = .95$) and on the comparative structure of Task 2, 2 ($SD = 0$). No significant differences in the number of feedback moments were found between conditions on either the dative ($F(1,46) = .49, p = .49$) nor the comparative structure ($F(1,46) = 1, p = .32$).

St : Das Wetter ist guter in Sankt Anton. [Error-Grammatical]
 T : Remember that gut has an irregular comparative form. [FB- metalinguistic]
 T : Please try it again. [FB-Elicitation]
 St : Das Wetter ist besser in Sankt Anton.
 T : Yes, well done.

Note: St = Student, T = Teacher, FB = Feedback

Figure 2. Example feedback episode at during-task phase of repetition task 2 (Comparative).

2.6 *Dependent measures*

Drawing on Macrory and Stone's (2000) study, investigating the relationship between knowing and using the perfect tense in French, we attempted to examine both

students' implicit and explicit knowledge (Ellis, 2005a) on the target structures, by means of the following three ways of examining acquisition.

- 1) Learners verbalizing the rules on a) a dative case after a two-way preposition and b) the comparative in German in a metalinguistic knowledge test (MKT).
- 2) Their actual knowledge on using these rules in a fill-in-the gap exercise (FITG).
- 3) Their ability to use both structures in two meaning-based oral tasks.

We did not measure the complexity of the performance because the testing tasks were structured and focused on the use of the two target structures. The task elicited only short sentences and there was not much variety in the way the tasks could be carried out.

To avoid the retrieval of explicit knowledge on the structures students performed the oral task first followed by the written task (FITG) and the metalinguistic knowledge test MKT. The same measurement design was used for Task 1 and 2 (see Figure 1). One week prior to the start of each task students performed pretests on written accuracy and oral accuracy and fluency. One week after completing the task students took part in posttests-1 on MK, written accuracy and oral accuracy and fluency. Posttests-1 only measured the effects of instruction and task performance which enabled us to examine whether there were any significant differences between conditions. Results of the linear mixed model analysis showed no significant differences at Post-1 for the two structures (MKT: $t(230) = 1.73, p = .085$; FITG: $t(322) = 1.087, p = 2.78$; Oral Accuracy: $t(322) = .136, p = .892$; Oral Fluency: $t(322) = -.562, p = .575$). Two weeks after each main task performance students of the R condition repeated a similar task. The NR condition performed a filler task which was not related to the use of the target structures. One week later students participated in posttests-2 which measured the effects of task repetition. Two weeks later students performed posttests -3 which measured the delayed effects of task repetition.

2.6.1 Metalinguistic knowledge tests (MKT)

In the MKT on the dative structure students were asked to complete the following rule: When you describe a room, using the following prepositions: *in, an, auf, hinter, neben, unter, über, vor, zwischen*, a *der* -word changes into ...; a *die*-word changes into..., a *das*-word changes into...; a *die*-plural changes into.... Scores were expressed in percentage correct (total amount of correct dative cases divided by four). For the MKT on comparatives, students were invited to write down the standard rule to form the comparative in German and two additional irregular 'rules' to form the comparative. Scores were expressed in percentage correct (total amount of correct rules divided by three). Students did not perform the MKT at the pretest because we did not want to predispose them towards particular target structures, before the upcoming meaning-focused task performance

2.6.2 *Written accuracy tests*

As in our earlier study (Van de Guchte et al., 2015a) in which we compared the effect of prompts and recasts on the German dative structure and the comparative, we used a written gap-fill test to measure the students' actual use of both rules³. Written tests were administered by the teacher during the lessons, and took about 15 minutes. There were four similar versions of the written test: For the pretest, each participant received version A, for posttest-1 version B, for posttest-2 version C, for posttest-3 version D; thus all conditions received the same test at each measurement occasion.

Dative

The written tests on the dative included 14 gap-fill sentences in which students were asked to write down the correct case of an article after a two-way preposition (Appendix C). In order to disentangle application of the dative rule from knowledge of noun gender, grammatical gender was noted in brackets after the noun (Schrank (der)) [cabinet (the)]. Reliability (Cronbach's alpha) showed .70 for the pretest, .90 for Post-1, .92 for Post-2, and .93 for Post-3.

Comparative

The written test on the comparative included 13 gap-fill sentences in which students were asked to write down the correct comparative. Reliability (Cronbach's alpha) was .66 for the pretest, .66 for Post-1, .72 for Post-2, and .57 for Post-3.

2.6.3 *Oral accuracy tests*

Oral accuracy in both tasks was assessed by means of an oral communication test, held with one of the three interlocutors: two native speakers of German and the first author. The tests consisted of communicative tasks similar to those used in the treatment sessions. For both tasks, the oral test was administered in four versions: For the pretest, each participant received version A, for posttest-1 version B, for posttest-2 version C, for posttest-3 version D. The interlocutors, all seated in different rooms, administered the tests with help of a step-by-step plan to guarantee that all tests were performed similarly. Each student was tested individually and audio-recorded, out of class. The recordings were presented to a native speaker of German and the first author who rated them blind for condition for accuracy. The first author evaluated all the recordings; the second rater evaluated 70%. For both tasks inter-rater reliability was very high: 99% for the dative and 86% for the comparative.

³ With the fill-in-the gap exercise we attempted to tap explicit knowledge, but as Ellis (2005a) points out we cannot guarantee that the students did not also draw on their implicit knowledge.

Dative

For the oral picture-description test (Appendix A) the interlocutor used a 3-D picture of a room, including the German translations (plus gender) of the furniture and accessories in the picture. Students were asked to look at the picture and to describe the room. For the same reason as in the written tests, grammatical gender was noted in brackets after the noun. The interlocutors were asked to elicit a minimum of 15 target structures. Accuracy scores were expressed as percentages correct (correct use of dative case after a two-way preposition / total number of articles used after a two-way preposition). When a student used other prepositions than the target structures to describe the room, these were not counted. Furthermore, repairs were assessed and not the utterance before the repair.

Comparative

During the comparison test the interlocutor showed the student a chart with information on two ski-holidays regarding twelve elements: the number of ski lifts, length of slopes, the cost of a ski pass, snow reliability, altitude of the ski area, après ski, beauty of area, weather, temperature, surface, quality of the apartment, and kindness of the staff (Appendix B). Students were asked to compare twelve elements of the two holidays concerned. Accuracy scores were expressed as percentages correct (correct use of comparatives / total number of comparatives used). Incorrect use of the comparative meant that either the comparative form itself was incorrect or that the comparative was not used properly in the context.

2.6.4 Oral fluency tests

With respect to fluency, we ran a script on the sound files written by De Jong and Wempe (2009) in PRAAT (Boersma & Weenink, 2007), with which we were able to calculate students' *speech rate* (total number of syllables divided by total duration of the task performance). However, as De Jong et al. (2012) point out speech rate takes 'breakdown fluency and speed fluency (...) together into one measure that encompasses aspects of pausing as well as speed of delivery' (p. 2). Therefore we included two measures that did not confound pausing and speed of delivery: *articulation rate*, that is, mean duration of syllables (speaking time divided by total number of syllables), and the *number of silent pauses* (number of silent pauses divided by speaking time). Following De Jong et al. (2012), silences of 250 ms. or longer were considered to be hesitations or pauses, and thus silences shorter than 250 ms., so-called micropauses (see, e.g., Riggenbach, 1991), were discarded.

2.7 Statistical analysis

To examine whether task repetition was effective in promoting the accurate use of the two new grammar structures we carried out a linear mixed model analysis (LLM) for each dependent measure (metalinguistic knowledge, written accuracy, oral accuracy and oral fluency). We investigated the effects of two independent variables (i.e., grammar structure and condition) on four measurement occasions: pre-

test, post-1, measuring the effect of the instruction which was the same for both groups, post-2, measuring the effect of the repetition intervention, and post-3 measuring the delayed effect of the repetition intervention. The fact that students were exposed to a Repetition and No-Repetition condition, two target structures and four measurement occasions made the resulting observations correlated (i.e., not independent). For that reason, we opted for a statistical approach that could manage correlated errors accurately and therefore used a general linear mixed model (GLMM). The model included fixed and random factors for each dependent measure. The three fixed factors were time of measurement, grammar structure and condition. We analyzed the model, including three 2-way interactions and one 3-way interaction. The random factor was participants. The alpha level was set at .05.

3. RESULTS

3.1 Metalinguistic knowledge tests

Table 1 presents the means and standard deviations resulting from the metalinguistic knowledge tests. On the MKT's we found a main effect for time ($F(2, 230) = 3.47, p = .033$), for condition ($F(1, 46) = 12.66, p = .001$), and for structure ($F(1, 230) = 11.43, p = .001$). Pairwise comparisons showed that the Repetition condition outperformed the No-Repetition condition ($MD = .206, 95\% \text{ confidence interval [CI]} = .089, .323, p = .001$). The main effect for condition was moderated by time ($F(2, 230) = 4.4, p = .013$). At Post-2 (1 week after task repetition) the scores of the Repetition condition increased on both structures with large effect sizes, compared to the No-Repetition condition (Dative, $d = .94$; Comparative, $d = 1.22$). At Post-3 (3 weeks after task-repetition) the Repetition condition showed a decrease on the dative structure ($d = .60$) and showed no change for the comparative structure.

Table 1. Group means (proportions correct) and standard deviations for metalinguistic knowledge tests for three measurement occasions

Structure	Condition	Post-1		Post-2		Post-3	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dative	R	.69	.28	.92	.19	.75	.34
	NR	.58	.32	.65	.36	.54	.37
Comparative	R	.78	.31	.96	.11	.96	.15
	NR	.71	.38	.63	.37	.71	.36

Note. R = Repetition group (n = 24); NR = No-Repetition group (n = 24)

3.2 Written accuracy tests

Table 2 demonstrates the means and standard deviations resulting from the written accuracy tests.

Table 2. Condition means (proportions correct) and standard deviations for written accuracy tests for four measurement occasions

Structure	Condition	Pre		Post-1		Post-2		Post-3	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dative	R	.10	.10	.74	.27	.87	.19	.82	.29
	NR	.07	.08	.63	.31	.61	.35	.55	.33
Comparative	R	.28	.11	.90	.11	.88	.10	.82	.14
	NR	.29	.16	.80	.21	.73	.25	.74	.19

Note. R = Repetition group ($n = 24$); NR = No-Repetition group ($n = 24$).

Results of the fill-in-the-gap tests showed a main effect for time ($F(3, 322) = 207.42, p < .001$) and for condition ($F(1, 46.24) = 9.74, p = .003$). Pairwise comparisons revealed that the Repetition condition outperformed the No-Repetition condition ($MD = .126, 95\%$ confidence interval [CI] = .045, .207, $p = .003$). In addition, the main effect for condition was moderated by time $F(3, 322) = 4.22, p = .006$. At Post-2, for the Repetition condition, we observed an increase for both structures with large effect sizes, compared to the No-Repetition condition (Dative, $d = .92$; Comparative, $d = .77$). At Post-3, we noted a decrease for both structures (Dative, $d = .88$; Comparative, $d = .45$).

3.3 Oral accuracy tests

Table 3 presents the means and standard deviations resulting from the oral accuracy tests. On the oral accuracy tests we found a main effect for time ($F(3, 322) = 90.11, p = .001$) and for structure ($F(1, 322) = 281.38, p = .001$). No significant effect was found for condition ($F(1, 46) = 1.44, p = .237$). Pairwise comparisons showed that the mean scores on the comparative structure were significantly higher than the scores on the dative structure ($MD = .367, 95\%$ confidence interval [CI] = .324, .401, $p < .001$). Furthermore, the main effect for structure was moderated by time ($F(3, 322) = 5.24, p = .002$).

Table 3. Group means (proportions correct) and standard deviations for oral accuracy tests for four measurement occasions

Structure	Condition	Pre		Post-1		Post-2		Post-3	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dative	R	.02	.03	.40	.40	.59	.38	.63	.39
	NR	.02	.05	.44	.39	.47	.38	.46	.38
Comparative	R	.51	.19	.87	.13	.84	.15	.88	.14
	NR	.46	.15	.85	.16	.76	.18	.81	.17

Note. R = Repetition group ($n = 24$); NR = No-Repetition group ($n = 24$).

3.4 Oral fluency tests

In Tables 4 a, 4b, and 4c the means and standard deviations resulting from the fluency tests are presented. On speech rate we found a main effect for time ($F(3, 322) = 66.19, p < .001$), but not for condition ($F(1, 46) = .079, p = .780$) or structure ($F(1, 334) = .376, p = .540$). On articulation rate we found a main effect for time ($F(3, 322) = 4, p = .008$), but not for condition ($F(1, 46) = .025, p = .875$) or structure ($F(1, 336) = 1.49, p = .223$). On silent pauses we found a main effect for time ($F(3, 322) = 11.6, p < .001$), and structure ($F(1, 322) = 139.77, p < .001$), but not for condition ($F(1, 46) = .510, p = .479$). It appeared that students produced more silent pauses in the preposition plus dative task than in the comparative task.

From Pre- to Post-3 students of the R condition, on the dative task, increased on speech rate by .18, on articulation rate by .04, and decreased on the number of silent pauses by .04. On the comparative task students of the R condition increased on speech rate by .53, on articulation rate by .13, and decreased on the number of silent pauses by .34. Students of the NR condition, on the dative task, increased on speech rate by .23, on articulation rate by .10, and on the number of silent pauses by .01. On the comparative task students of the NR condition increased on speech rate by .60, on articulation rate by .22, and decreased on the number of silent pauses by .28.

Table 4a. Group means and standard deviations for speech rate for four measurement occasions

Structure	Condition	Pre		Post-1		Post-2		Post-3	
		M	SD	M	SD	M	SD	M	SD
Dative	R	1.71	.39	2.03	.49	2.03	.46	1.89	.52
	NR	1.76	.43	1.78	.27	2.09	.39	1.99	.44
Comparative	R	1.98	.53	2.39	.47	2.22	.50	2.51	.55
	NR	1.93	.32	2.31	.39	2.21	.33	2.53	.46

Note. R = Repetition group ($n = 24$); NR = No-Repetition group ($n = 24$)

Table 4b. Group means and standard deviations for articulation rate for four measurement occasions

Structure	Condition	Pre		Post-1		Post-2		Post-3	
		M	SD	M	SD	M	SD	M	SD
Dative	R	3.69	.31	3.83	.33	3.92	.40	3.73	.43
	NR	3.75	.57	3.69	.33	4	.51	3.85	.39
Comparative	R	3.61	.38	3.86	.53	3.69	.39	3.74	.45
	NR	3.52	.32	3.78	.64	3.64	.41	3.74	.37

Table 4c. Group means and standard deviations for silent pauses for four measurement occasions

Structure	Condition	Pre		Post-1		Post-2		Post-3	
		M	SD	M	SD	M	SD	M	SD
Dative	R	1.17	.18	1.13	.34	1	.22	1.13	.30
	NR	1.09	.25	1.13	.17	1	.25	1.10	.29
Comparative	R	1.03	.35	.84	.25	.91	.28	.69	.24
	NR	.97	.20	.85	.22	.89	.28	.69	.19

3.5 Accuracy versus fluency

The trade-off effect between accuracy and fluency measures was tested through correlations. For the Repetition condition on the dative task we found a significant moderate negative correlation on Post-2 ($r(24) = -.57, p = .004$) between oral accuracy and speech rate. Further, no significant correlations were found between accuracy and other fluency measures.

4 DISCUSSION

In this experimental study we aimed to investigate whether task repetition could be an effective FonF strategy to promote the accurate use of German grammar structures. In addition, we investigated whether the effects of task repetition were different for the two tasks with the different grammar structures. Furthermore, the present study made it possible to examine Skehan's (1996a, 1998a) Trade-off Hypothesis.

As for the first research question, 'Does task repetition as a post-task activity promote the accurate use of two new target structures?' results showed positive effects on both metalinguistic knowledge on the rule (MK) and on written application of the rule (FITG). Since we have no online task-performance data that show the students really focused on the grammar features during task repetition, we can only assume that the better post-test performance of the R group can be attributed to students' attention to these structures during task repetition. Because the students in the R group had more opportunity to practice using the target forms, this may probably have raised or re-awakened their metalinguistic awareness. Nevertheless, this can only be claimed for declarative knowledge but not for procedural knowledge on the target structures. That is to say, the oral tests, which attempted to tap the procedural knowledge on the target structures, revealed no significant results in favor of the R condition. In addition, results showed significant correlations between rule-knowledge and oral accuracy for the dative structure but not for the comparative structure. These findings might be explained by the fact that the forming of the dative case after a preposition used with its meaning of static location, is considered a syntactic rule of which Hulstijn and de Graaff (1994) argue that the acquisition strongly depends on rule-based learning. However, it might have been the case that some frequent dative forms such as *in der Mitte* (in the middle) have been acquired

through exemplars. The forming of the comparative on the other hand, is only partly based on a rule (*add -er* to the adjective or adverb) but also knows many irregular forms which are acquired through exemplars.

Several explanations why students in the R condition did not outperform the NR condition on oral task performance can be given. From a skill acquisition (Anderson, 2000) point of view, we might argue that repeating a task once does not lead to the amount of communicative practice (DeKeyser, 1998) that one needs to turn explicit knowledge into implicit knowledge. In other words, students need more practice to fully automatize the target structures. This point has also been made by Gass et al. (1999), who comment that ‘learners may have knowledge of certain features, but they may not have acquired control over that knowledge’ (p. 554).

Not finding an effect in the oral measurement might be caused by how we operationalized the task repetition in terms of a) a similar not an identical task and b) the interval of 2 weeks between the main task and the repeated task. Possibly, learners did not regard the main and repetition as ‘family’ (Skehan, 1996a) which did not give them the advantage of being familiar with the content and therefore able to focus on form. In addition, practicing twice in three weeks, may not have been sufficient to proceduralize the dative and comparative structures. In line with Bygate (2001) and Sato (2010), we suggest that for learners to be more successful on similar tasks more practice or more task exposure might be required.

Another perspective on why no significant results were found for the Repetition condition on the oral accuracy test may be gained by analyzing why the No-Repetition condition performed so well on this test. Probably, the assessed public performance (motivated by school report requirements) at the end of the main-task phase accounts for the considerable scores of this condition. Skehan and Foster (1997) investigated the effects of the ‘threat’ of a public performance, as a post-task activity, and indeed found improvement in accuracy. They concluded that because learners knew that a private task performance in pairs would lead into a public performance, their attention was, to some extent, more focused on form and learners ‘were more attentive to error during actual task performance’ (Skehan et al., 2012, p. 174).

A positive answer to the second research question, ‘Does the effect depend on the grammatical structure involved?’, requires an interaction effect between condition and structure. The results did not reveal any of these interactions, so, we therefore conclude that the effect of task repetition does not depend on the structure.

We now turn to the third research question, ‘Does a student’s focus on oral accuracy have a negative effect on oral fluency?’. With this question we aimed to contribute to the literature on attentional resources in task-based language learning. For the Repetition condition on the dative task we found a moderate *negative* linear relationship between accuracy and fluency on Post-2 (one week after task repetition) which indicates that students who performed the task more accurately showed less fluency. This finding lends support to Skehan’s (1996a, 1998a) Trade-off Theory of limited attentional resources. We presume that retrieval and application of the dative rule, during oral performance, confronted students with a heavy cognitive load, resulting in less attention to fluency. Since students already had exemplars of compar-

ative structures in their long-term memory, retrieval was much easier, and did not interfere with a fluent performance.

4.1 Limitations and directions for future research

Despite the positive findings of the present study, some limitations need to be acknowledged. First, we only have post-test data and no online data to measure the effects of task repetition. Online-recordings of the repeated performance would have provided us with more insights in whether students actually paid attention to the grammatical structures and how this could have led to the significant findings on the MK and FITG tests.

Second, our choice to use ‘accuracy as a reflection to focus on form’ (Foster & Skehan, 1996, p. 280) might have been too conservative to get a full insight into the effects of task repetition. We agree with Foster and Skehan that, with taking complexity into account, one can also capture the learner’s risk-taking and experimenting, which connects with gains in interlanguage development. A challenging task for further research on task repetition is to investigate how the use of similar tasks in task repetition could promote accurate use of linguistic structures in oral task performance and which interval of repetition would be effective.

5 CONCLUSIONS

Altogether, the results of the present study have led us to conclude that repetition of a similar task may be an effective strategy to promote the accurate use of grammatical structures. It appeared that students who received corrective feedback at the during-task stage and in addition repeated a similar task were better in verbalizing and applying the rules on a simple and more complex German grammar structure. We find this encouraging because these findings suggest that previous knowledge on the target structure is available to the learners to build on in subsequent performance. Possibly, more practice is needed to automatize these structures in oral performance.

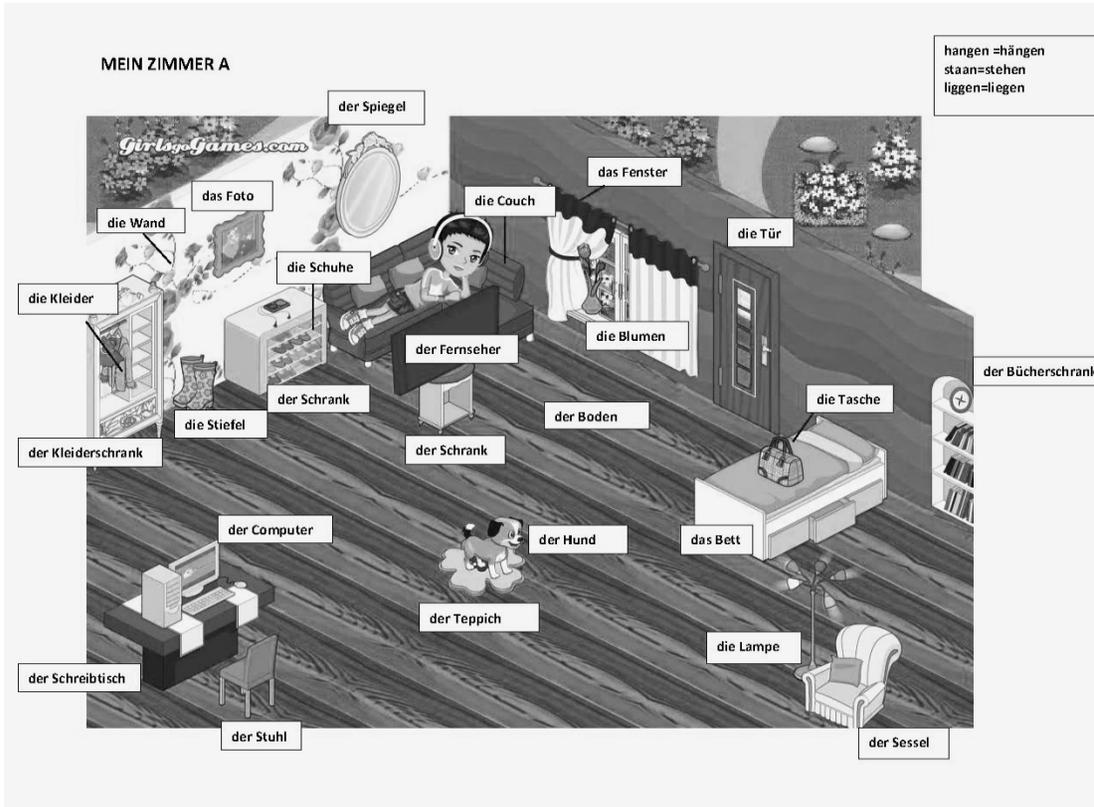
In addition, no interactions were found between the effect of condition and task, which indicates that the effect of task repetition did not depend on the structure. Furthermore, this study provided additional evidence for Skehan’s (1996a, 1998a) Trade-off Theory which claims that people have limited attentional resources. The negative interaction we found within the Repetition condition on the dative structure provides additional evidence that accuracy may come at the expense of fluency. However, this was not the case for the comparative structure, which has many similarities with the L1, suggesting that under specific conditions accuracy and fluency are not in competition with each other.

A strength of the present study is that it has been carried out in real classrooms and that task repetition has been embedded in an ongoing program. For that reason, we argue that the findings of our study may have an important implication for language pedagogy in a foreign task-based language learning context. The Common European Framework of Reference for Languages asks for a more communicative approach in language teaching but, at the same time, does not provide sufficient

guidance on how grammar education can be a part of meaning-focused language education. The results of the present study underline that similar task repetition can be an effective strategy to focus on form in task-based language teaching, when students have already been directed to form in the during-task stage. Selecting and re-using similar tasks systematically in a task-based curriculum may promote learners getting better at a task in terms of accuracy, complexity and/or fluency (Bygate, 2001) and in addition keep students interested and motivated to perform the tasks.

APPENDIX A

EXAMPLE ORAL TEST DATIVE



APPENDIX B

EXAMPLE ORAL TEST COMPARATIVE

Lust auf Skiurlaub? Davos oder St. Moritz?

You are looking forward to go on a ski-trip in Switzerland. You have not yet decided on where you want to go. In front of you lies a chart with information on two ski destinations from the website www.skigebiete-vergleichen.ch

- 1) Please compare the ski areas Davos and St. Moritz on all the 12 elements with each other.
- 2) Decide where you want to go to and explain why.



	Davos	St. Moritz
Lifte	34	60
Pisten	80 km lang	100 km lang
Ski-Pass (pro Tag)	€38	€42
Schneesicherheit	90% sicher	75% sicher
Höhe Skigebiet	2590 Meter hoch	1856 Meter hoch
Après-Ski	Langweilig☹	Toll!
Schönheit der Umgebung	Nicht so schön	Sehr schön
Wetter		
Temperatur	8°C warm	1°C warm
Größe des Skigebietes	320 km groß	350 km groß
Qualität des Appartments	Sehr gut	Gut
Freundlichkeit des Personals	freundlich	Nicht so freundlich

APPENDIX C

EXAMPLE FILL- IN- THE-GAP EXERCISE DATIVE

- 1) Über mein_____ Couch (die) hängt ein Gemälde.
- 2) An d _____ Wand (die) hängt ein Poster.
- 3) Auf dein____ Schrank (der) stehen viele Bücher und CDs.
- 4) Hinter d _____ Computer (der) liegen fünf Bleistifte.
- 5) In sein_____ Zimmer (das) ist es gemütlich.
- 6) Neben d _____Stuhl (der) liegt ein Tennisschläger.
- 7) Gegenüber mein_____ Bad (das) liegt ein Teppich.
- 8) Vor d _____ Lampe (die) liegt ihr Hund.
- 9) In d _____ Fenster (das) stehen schöne Pflanzen.
- 10) Neben mein_____ Spiegel (der) hängt ein Foto.
- 11) Hinter d _____ Sofa (die) liegen meine Schuhe.
- 12) Über sein_____ Bücherschrank (der) hängt eine Uhr.
- 13) Ich sitze oft vor d _____ Fernseher (der).
- 14) Auf d _____ Bild (das) siehst du meine Freundin.

CHAPTER 6

DISCUSSION AND CONCLUSIONS

In three experimental studies, we investigated the effects of different FonF strategies in all three stages of the task-based language teaching framework (pre-task, during-task, and post-task). In addition, we examined in Study 2 and 3 whether these effects depended on the grammatical structure involved. Furthermore, we investigated Skehan's (1996a, 1998a) trade-off hypothesis, which claims that learners possess limited attentional resources and a focus on, for instance, accuracy, may have a negative impact on their fluency or linguistic complexity. In this concluding chapter, we will present an overview of the main results and will attempt to interpret these results. Next, we will raise several methodological issues that need to be considered. Finally, we will make some suggestions for future research and discuss potential implications for classroom practices.

1. AN OVERVIEW OF THE MAIN RESULTS

From the findings of the three empirical studies, the overall picture seems to be that a focus on form (FonF) is possible and effective in the pre-task, during-task, and post-task three stage of the TBLT framework. In this section we will review and interpret the effects of the FonF strategies based on the research questions of the three studies.

- RQ1 How did the different FonF strategies affect learners' declarative knowledge of the new grammatical structures and their oral task performance in terms of accuracy (Study 1, 2, 3) and complexity (Study1)?
- RQ2 Did these effects depend on the grammatical structure involved? (Study 2, 3)
- RQ3 Did learners' focus on oral accuracy come at the expense of oral fluency (Study 1, 2, 3) or oral complexity (Study 1)?

1.1 Effects on declarative knowledge and oral task performance

Table 1 summarizes the results of all three studies in this dissertation. In Study 1 we observed that a pre-task focus on form during video model observations is beneficial to learners' subsequent task performance but does not affect learners' planning processes prior to the task performance. Students who focused on form during the video model observations used the target structure significantly more often and more accurately during subsequent task performance. Another finding was that students who focused on the content (meaning) of the modeled task, during pre-task observations, conducted a more complex performance afterwards. In sum, directing learners' at-

tention in the pre-task to either meaning or form may result in a more accurate or complex task performance. Depending on the purpose of the lesson, teachers can decide what they want learners to focus on (content or form).

Table 1. Summary of results of the three experimental studies

<i>Study</i>	<i>Measures</i>	<i>Results</i>	
1. Pre-task		<i>Post-1</i>	<i>Post-2</i>
	<i>Oral Accuracy</i>		
	Correct use TS	FonF > FonM	No effect
Conditions	<i>Oral Complexity</i>		
a) Form	Use of TS	FonF > FonM	No effect
b) Meaning	Clauses per ASU	FonM > FonF	FonM > FonF
	Coordination	FonM > FonF	No effect
	Subordination	No effect	FonM > FonF
	<i>Planning</i>		
	Form	No effect	No effect
	Meaning	No effect	No effect
2. During-task			
	<i>Written Accuracy</i>		
Conditions		Recasts > Control	
a) Prompts		Prompts > Control	
b) Recasts		Prompts > Recasts	
c) Control		Larger effect of Recasts on Comparative than Dative structure	
	<i>Oral Accuracy</i>		
		Recasts > Control	
		Prompts > Control	
		Prompts > Recasts	
3. Post-task			
	<i>Metalinguistic Knowledge</i>		
Conditions		Repetition > No-Repetition	
a) Repetition	<i>Written Accuracy</i>		
b) No-Repetition		Repetition > No-Repetition	
	<i>Oral Accuracy</i>		
		No effect	

In Study 2 we observed that both recasts and prompts contribute to the accurate use of German grammatical structures, with prompts being more effective than recasts. On both written and oral accuracy, the recast and prompt group outperformed the control group. In addition, the prompt group performed significantly better than the recast group. However, for oral accuracy the main results for feedback type were

moderated by time (see Chapter 4). For written accuracy, the effect of recasts and prompts differed according to structure over time (see Figure 1 in Chapter 4). Compared to prompts, recasts had a larger effect on the comparative structure than on the dative structure.

In Study 3 we observed that the repetition of a similar task during the post-task, after rule deduction and having provided learners with feedback in the form of prompts during the main task, resulted in gains in learners' metalinguistic knowledge and written accuracy of the two German grammatical structures. We assume that the use of these forms during task repetition may have raised or re-awakened learners' metalinguistic awareness of the structures. However, repetition as a focus on form strategy did not result in effects on learners' oral accuracy of these structures. Perhaps, gains in oral accuracy may require more practice of the same task type (Bygate, 2001).

1.2 Did the effectiveness depend on the grammatical structure?

In both Study 2 and 3 we investigated whether the effectiveness of the FonF treatment depended on the grammatical structure involved. This was only the case in Study 2. It appeared that recasts, as compared to prompts had a larger effect on the comparative structure than on the dative structure. From this finding we are led to the conclusion that grammatical structures with simple underlying rules that are similarly realized in the L1 may also be corrected effectively through recasts. Apparently, recasts are easier to notice when the L2 target structure is related to the L1. This saliency may enable learners to compare the correct structure with their own incorrect utterance and thus promote its acquisition. Furthermore, the recast group performed the dative task more fluently than the prompt group. Conceivably, the students of the recast group did not think about a rule while accomplishing the dative task and because of this lack of focus on accuracy, they freed up attentional resources that could be devoted to fluency.

Prompts appeared equally effective for both structures. However since prompts can break the communicative flow quite obtrusively, it may be preferable to correct simple L1-related structures through less intrusive recasts.

In Study 3 the effects of task repetition were not significantly different for the two target structures: the dative and the comparative structures. We therefore draw the conclusion that task repetition as a FonF strategy may be equally effective in enhancing metalinguistic knowledge and written accuracy on the dative and comparative structures.

1.3 Investigating Skehan's Trade-off Hypothesis

Skehan's (1996a, 1998a) Trade-off Hypothesis was investigated and confirmed in all three studies. Skehan's hypothesis claims that learners have limited attentional resources. As a consequence, learners' attention to either an accurate, complex or fluent task performance would entail less attention to the other dimensions and thus diminish the performance therein.

From the findings of both Study 2 and 3 it appears that paying attention to accuracy comes at the expense of oral fluency regarding the complex, syntactic Dative rule. Since this was only the case for the Prompt (Study 2) and Repetition (Study 3) group, we presume that learners who explicitly focus on complex rules, without a connection to the L1, have less attention left for a fluent performance. In comparison, no such trade-off was found for the comparative structure. This may suggest that the retrieval of an L2 rule that is similarly realized as the L1 rule does not interfere with a fluent performance. Table 2 demonstrates a summary of the trade-off effects in all three studies.

Table 2. Summary of trade-off effects in all three studies

Study	Trade-off
1. Observation of models in pre-task	<p><i>Oral accuracy and complexity</i> For the Dative task, for the Form Observation group in the delayed post-test</p> <p><i>Between different complexity measures</i> Use of target structure came at the expense of clauses per AS-Unit and subordination.</p>
2. Feedback during-task	<p><i>Oral accuracy and fluency</i> For the Dative task, for the prompt group in the delayed post-test</p>
3. Task repetition in post-task	<p><i>Oral accuracy and fluency (speech rate)</i> For the Dative task, for the Repetition group in the immediate post-test</p>

In Study 1, the FonF through observation study, no trade-off between accuracy and fluency was found for the complex dative structure. This may imply that, in contrast to an explicit focus on accuracy, an *implicit* focus that does not include rule-based knowledge, does not interfere with fluency. An interesting finding, however, was that this implicit focus on accuracy did come at the expense of oral complexity, in terms of clauses per AS-unit and the ratio of subordination. Another remarkable trade-off was found between two complexity measures. At post-1, students of the FonF condition who used the target structure more frequently, scored lower on subordination, and students of the FonM who used the target structure more frequently, produced less clauses per AS-unit. At post-2, students of the FonF condition who used the target structure more frequently, generated less clauses per AS-unit and a lower ratio of subordination.

In sum, the more explicit foci such as prompts and task repetition on a complex grammatical structure (preceded by rule deduction and prompts during the main task) came at the expense of fluency, whereas the more implicit foci on this same complex structure, such as recasts and the observation of models did not. Furthermore, an implicit focus on a complex structure during the observation of models did

lead to lower scores on complexity measures and resulted in trade-offs between different complexity measures.

On the basis of these results we argue that an explicit focus on a complex grammatical structure at the during and post-stage of the TBLT framework may lead to lower scores on learners' oral fluency. However, an implicit focus on that same structure through modeling also comes at a price. Such a focus may negatively affect learners' performance in terms of complexity.

2. METHODOLOGICAL ISSUES

In this section we will consider the methodological decisions and other choices we made regarding issues of both internal and external validity. According to Cook and Campbell (1979) internal validity 'refers to the approximate validity with which we infer that a relationship between two variables is causal or that the absence of a relationship implies the absence of cause' (p. 37). In other words, can we ascribe the results we found to the FonF treatment or do other variables play a role? External validity refers to the generalizability of the treatment outcomes across different settings.

2.1 *Interventions*

2.1.1 *Operationalization of the FonF strategies*

The most important variable in all three studies is the focus on form treatment variable. Several critical remarks can be made about how we operationalized the FonF strategies in our three studies.

In Study 1, while observing the videos, the FonF group focused on the use of two-way prepositions that require the dative case, by means of written questions. One may put forward that providing learners with models in the pre-task may promote imitation instead of learners' use of their own creativity. Notwithstanding this argument, we argue that in the foreign language context, the classroom is the only place where learners are exposed to input. From our point of view, models can be used as a rich source of input that provides learners with new language and may prime online-noticing (see Swan, 2005 for a discussion on input in TBLT).

Another point of criticism may be that, in comparison to the other two studies, the focus on the target structures in Study 1 was so implicit that it did not make learners aware of the rules behind the use of these structures. We agree that regarding cross study comparisons it may not be fair to compare the results of such an implicit strategy with more explicit ones such as prompts. However, because we attempted to investigate FonF strategies in all three stages, we opted for a strategy which we thought was appropriate for a pre-task. From our point of view, the primary focus in the pre-task should be on meaning and a FonF at this stage should be aimed at making the structures salient, not on rule explanation.

One could say that the control group in Study 2 was not a good comparison group to inform us about the effects of feedback because students did not participate in the

treatment tasks. One could therefore argue that these students had no previous exposure to the content of the task nor any experience with performing such a task, which may have disadvantaged them. This is partially true, these students did indeed not have experience with the exact same communicative tasks as the experimental groups did. However, they did have the opportunity to practice the oral use of the target structures through structured dialogues. These structured dialogues in their textbooks included the description of a living room and comparisons of different objects. The study was designed this way because there was no opportunity to embed the communicative tasks into the existing curriculum for German in the control group. The control group was part of the regular department of the school in which teachers do not design their own tasks, but are required to finish the textbook before the end of the year. In future studies on corrective feedback we would advise to let the control group participate in the treatment tasks, without providing the corrective feedback (see also Lyster, Saito, & Sato, 2013).

Another issue that could have been a threat to the internal validity of Study 2, is the number of feedback moments in the two experimental conditions. Although this number did not differ statistically significantly between conditions, the variation within the condition for the number of feedback moment was sometimes quite large. This may indicate that some students received more feedback than others. Although we did not find a relation between the number of feedback moments and learning gain, in a subsequent study we would control for the amount of feedback moments per hour to ensure equivalence between participants. For that reason, in Study 3 we controlled for the number of feedback moments per student per hour.

In Study 3 the attention of the students was directed to form in the during-task phase through deduction of the grammar rules behind the target structures and feedback in the form of prompts (see the remark above on number of feedback moments). The Repetition group was asked to repeat a similar task and we expected them to focus on the accuracy of the grammatical forms that were elicited by the task. A potential point for discussion may be the way in which we operationalized task repetition. Although several studies have demonstrated that effects of task repetition do not carry over to new contexts, task repetition in our study involved similar tasks instead of the same tasks. The most important reason why we opted for the use of similar tasks is that in an earlier pilot study we had negative experiences with students repeating the exact same task, immediately after the main task. Students commented that it was boring and demotivating to perform the exact same task again. We observed that students treated the repetition task as an exercise, made no real effort and tried to finish the task as quickly as they could. To keep learners motivated, we altered the content and the timing of the repetition tasks. The content of the repetition tasks was slightly different from the initial main tasks in the sense that they were adapted to the topic of the subsequent main task in the curriculum. Because the main and repetition task in Study 3 elicited the same structures, used more or less the same language and imposed the same cognitive demands, we considered them as tasks belonging to the same family (Skehan, 1996). In line with Bygate (2001), we expected that familiarity with a certain task would give learners the opportunity to focus on the formulation and articulation of the message, which could have a positive effect on the accuracy of their subsequent task performance.

Following Sheppard's (2006) study that demonstrated that effects of repetition can transfer to new contexts, when providing students with feedback, the students in our study received corrective feedback through prompts on the target structures during their first task performance.

Regarding the timing of the repetition task, we did not ask students to perform the repetition task immediately, but two weeks later as a post-task of the next main task. Perhaps we did not find an effect in the oral measurement because we operationalized the task repetition in terms of a) a similar but not the same task and b) the interval of 2 weeks between the main task and the repeated task. Possibly, learners did not regard the two tasks as 'family' (Skehan, 1996b), which did not give them the advantage of being familiar with the content and therefore enable them to focus on form. In addition, practicing twice in three weeks, may not have been sufficient to proceduralize the dative and comparative structures. In line with Bygate (2001), we suggest that learners may get better at a task, when tasks would systematically be selected and reused in language teaching.

Altogether, we argue that the implemented strategies in our studies were valid representations of a Focus on Form. We therefore conclude that the effects found in the experimental studies were due to the use of the FonF strategies and not by other variables in the interventions.

2.1.2 *Design of the treatment tasks*

To ascertain that the treatment tasks adhered to TBLT principles, all tasks were reviewed by the first author's colleagues at her school who were all experienced task-based curriculum designers and by a linguistics professor at the University of Amsterdam. The most debated issue was the tension between the task eliciting the use of the target structure and the primary focus on meaning. Moreover, we did not want learners to see the tasks as exercises to only practice the use of the target structures. To avoid this, two important decisions in task design were taken. First, the pre-task provided learners with useful language for the main task performance but never mentioned the use of a certain grammatical structure (see Ellis, 2003). Second, the task outcome was aimed more at communicating meaning than the learning of a grammatical form.

In Study 2 and 3 we opted for the use of focused tasks which increased the chance that learners actually used the target structure. This enabled us to provide the learners with feedback on these structures. However, learners' primary focus needed to be on the meaning of the task. We found that the focus on meaning was more easily achieved for the comparative than for the dative tasks. Using comparatives in a comparison task (compare two mobile phones in Study 2; compare two holidays in Study 3) was more natural than the use of a two-way preposition plus dative in a room description task (Study 2 and 3). That is to say, to elicit a preposition-plus-dative-case in the latter task, it was necessary to ask the students to describe the place of the furniture or accessories in their dream room. Without these specific instructions, the description of the dream room could also have been performed in other ways, with different structures. We acknowledge that the description tasks in

Study 2 and 3 may have focused less on meaning than we intended them to. For that reason, one may criticize that these tasks fit more in a task-supported rather than a task-based curriculum. We agree with Loschky and Bley-Vroman (1993) that it is very challenging to design production tasks for which the use of a certain grammatical structure is essential, while the primary focus is still on meaning.

For Study 1 we chose to compare a pre-task focus on form to a focus on meaning condition. For that reason, we opted for the more open task ‘give a tour through the school canteen for upcoming pupils and persuade them to join our school’. This task was designed in this way for two reasons: a) it elicited the use of prepositions plus a dative case; but b) it also allowed for more complex interpretations of the task.

In sum, considering the fact that the treatment tasks in our studies were designed to be used in a research situation which requires a high level of control and precision (see dimensions underlying the study of tasks, Bygate, Skehan & Swain, 2001, p. 5), we argue that we did as much as we could to make them valid representations of tasks in TBLT. First, learners were never told about the grammatical structures in the pre-task. Second, the task outcome was always focused on meaning instead of language and finally tasks were connected to real-life situations.

2.2 *Experiments in natural settings*

2.2.1 *Teacher effects*

Several measures were taken to control for threats of internal validity such as teacher differences and differences between classes.

In Study 1, the observational learning study, the first author, a teacher at the school, conducted the experiment in both experimental groups. The modeling task was observed individually in the computer lab, in conjunction with written instructions for each student, to ensure that students could not work together,

For Study 2 the threat to internal validity was the largest because the oral feedback in the experimental groups was provided by two different teachers. To control for this threat, a colleague in the language department regularly observed the lessons with the help of a feedback logbook, to make sure that both teachers adhered to the protocol regarding the provision of recasts and prompts. In addition, the colleague took note of the number of feedback moments per student. The observations revealed no differences in treatment except in terms of feedback type. To make sure that the control group followed the regular textbook the researcher received a copy of the lesson plans and visited the classes on a regular basis.

In Study 3 students were randomly assigned to the two conditions, within classes, each taught by another teacher. Thus, both conditions were present in each classroom. To avoid students from both conditions working together, the Repetition group was placed on the right side of the classroom and the No- Repetition group on the left side. In addition, lesson materials were designed for students to work independently .

2.2.2 *Participants*

Since the three experiments found place in three successive years, participants were not the same for each study. However, there are no reasons to believe that the participants of the experimental groups differed much from each other. First, all participants were 14-year old ninth-grade students in secondary education and were all recruited from the task-based Quest-department of the school. Second, they had all been learning German for 17 to 19 months at the A2 level of the Common European Framework of Reference for Languages (CEFR, Council of Europe, 2001), for 2 hours per week.

Furthermore, all participants of the experimental groups were used to working with communicative tasks in all second and foreign language classes (English, French and German). Therefore, the lesson series were not new for the students which could have influenced learners' motivation and consequently their performance. In other words, no Hawthorne effect will have occurred (Izawa, French & Hedge, 2011). In addition, the three experimental studies were conducted in real classrooms during students' regular classes.

In Study 2 we incorporated a control group consisting of students from the regular department of the school. Students in this control group followed their own, more form-focused, textbook-based curriculum, which included oral practice of the target structures through structured dialogues but not through the use of communicative production tasks. This may have disadvantaged the control group on the oral tests in comparison with the experimental groups.

2.3 *Measurements*

In this section we will discuss the measurements in the order in which the studies were conducted over time. This is in contrast to the other sections in which we discussed the studies in the natural order (pre, during, post) of the TBLT cycle. The reason for this order is that the measurement choices were revised due to new insights which arose during this PhD project.

We started the PhD project with Study 2, the feedback study. Because the oral feedback focused on two grammatical structures, we initially chose to measure learners' procedural knowledge by means of oral accuracy of the two structures. In addition, we used written accuracy measures to tap learners' declarative knowledge on the structures. For this measure we used a fill-in-the-gap test. One could ask why we did not use a grammaticality judgment test instead, because these tests have proven to be a valid measure to tap learners' declarative knowledge (see Ellis, 2005). Based on the limitation of many grammaticality judgment tests, that ask students to simply indicate the occurrence of an incorrect structure, we decided to use a test which forced students to provide the correct answer (DeKeyser, 1993). In addition, this format was most familiar to the students.

Another methodological issue is how we measured students' oral fluency. In our first study (Study 2) we based this on judgements by human raters. They assessed learners' fluency on a scale from 1, very low fluency, to 5, very high fluency. We based our choice for human ratings on the findings of De Jong, Steinel, Florijn,

Schoonen, & Hulstijn (2012) who concluded from the results of several studies that some objective measures of pausing and some measures of speech rate correlate positively with measures of perceived fluency by human raters. In addition, De Jong et al. (2012) noted that the aspects of fluency human raters judged, depended strongly on how these raters were instructed. Therefore, the raters in Study 2 were instructed to judge two aspects that correlate most with objective measures: speech rate and silent pauses. In addition, the raters listened to 15 preselected performances for each task at different levels in the scale to achieve rating consistency. All these measures were taken to ascertain that the judgments by human raters would be a valid measure of fluency. It appeared that the two human raters achieved very high interrater reliability, 91% for Task 1 and 90 % for Task 2.

A reviewer of the article that we submitted to a research journal argued, nevertheless, that the fluency analysis could be strengthened by using a measure that is more common in the literature on complexity, accuracy and fluency (e.g., syllables per second). Although we consider the judgment of human raters a valid measure of assessing fluency, this point of criticism was taken into account in the subsequent two studies by using more objective measures of fluency with help from a script written by De Jong and Wempe (2009) in the software program PRAAT (Boersma & Weenink, 2007). With help from this script we measured learners' speech rate (number of syllables / total time), articulation rate (number of syllables / speaking time), and the number of silent pauses.

To establish whether judgments of human raters and objective fluency measures of PRAAT correlated, we performed a Pearson's correlation analysis for Study 3. The results are presented in Table 3. For the measure *speech rate* we found moderate correlations for both the dative and comparative task. For the number of *silent pauses* we found a weak negative correlation for the dative and a moderate negative correlation for the comparative task. For *articulation rate* we only found a weak correlation on the dative task. These findings confirm earlier research (see De Jong & Wempe, 2009; De Jong et al., 2012) that demonstrated that objective measures of speech rate and pauses are predictors of subjective fluency. However, it must be recognized that these correlations were moderate and not high. It may be the case that, in addition to learners' speech rate and pausing, the raters may have taken into account learners' accent, the correctness of grammatical structures and the use of vocabulary as well (Freed, 1995). For example, some students achieved high levels of speech rate when measured by PRAAT, but when listening to the performance raters observed that the students spoke very quickly in a language that was more Dutch (L1) than the target language German. For that reason, the human raters had not judged that performance as highly fluent in German.

Table 3. Correlations between human- and automatically-measured fluency in Study 3

<i>Task</i>	<i>Speech Rate</i>	<i>Articulation Rate</i>	<i>Silent Pauses</i>
Dative	$r(190) = .56, p < .001$	$r(190) = .21, p = .003$	$r(190) = -.38, p < .001$
Comparative	$r(190) = .58, p < .001$	ns	$r(190) = -.63, p < .001$

When comparing the results for both objective and human rated fluency measures for trade-off effects for Study 3, we observed that the objective measures revealed a significant moderate negative correlation on Post-2 ($r(24) = -.57, p = .004$) between oral accuracy and speech rate for the Repetition condition on the dative task. The human-rated measures revealed a similar significant trade-off on Post-2 ($r(24) = -.49, p = .016$) and addition also one on Post-3 ($r(24) = -.56, p = .004$).

Another issue is that the oral accuracy test for the comparative task (i.e., ‘compare two mobile phones’) did not always elicit the intended target structures. To avoid this situation in the next study, Study 3, we enhanced the reliability of the outcomes of the oral comparative test by training the research assistants to instruct the students to use the adverbs that were included in the test instructions.

Finally, in Study 2 we encountered the problem that the control group, which did not receive treatment, were not motivated to perform the delayed oral test on the Comparative structure. To control for a valid outcome on this measure, we chose not to force the students to perform the test.

Study 3 was the second study we conducted in the research project and investigated the effects of task repetition as a strategy to focus on form. Because all the students in the study had discovered grammar rules for both structures in the during-task phase, we did not only want to measure learners’ application of this rule through a fill-in-the-gap test but also learners’ metalinguistic knowledge of those rules. Because learners used the target structures in an oral task and the Repetition group repeated a similar oral task we also measured oral accuracy. As we only have post-test data and no online task-performance data that could show that students really focused on the grammatical features, we do not have insights into what students actually did during task repetition and how this could have led to the significant findings on the MK and FITG tests. Therefore we can only assume that the repetition task provided more opportunity to practice using the target form and that the use of these forms may have raised or re-awakened metalinguistic awareness of the structures.

Finally, opting for ‘accuracy as a reflection of focus on form’ (Foster & Skehan, 1996, p. 280) might not have given us complete insight into the effects of task repetition. We agree with Foster and Skehan that, by taking complexity into account, one can also capture the learner’s risk-taking and experimenting, which is linked to gains in interlanguage development. The reason we only measured accuracy and not complexity was that we used focused testing tasks that elicited the use of grammatical structures and only short sentences. As a consequence, there were very few options to make the task more complex by experimenting with other language forms. There-

fore, in Study 1, the last study conducted, we used a less focused oral testing task which enabled us to not only measure learners' gains in accuracy but also in complexity in terms of 1) the use of the grammatical structure per clause; 2) total number of clauses per Analysis of Speech-Unit (ASU); 3) number of subordinated clauses per ASU; and 4) number of subordinated clauses per ASU.

Regarding the last study, Study 1, we measured the effects of two types of observation strategies on the learners' performance (1) in the pre-task planning and (2) in the main task. It must be recognized that we cannot be sure that the conducted think-aloud planning method right after the (1) phase, may have influenced the learners' performance in (2) as well. However, since the study was also aimed at identifying learners' thoughts on how they planned their performance, the use of well-established think-aloud methods (Gass & Mac Key, 2000; Kim, 2013) seemed the most appropriate option.

In summary, we note that it is challenging to design oral tests that elicit the use of different forms of a predetermined target structure, without compromising a focus on meaning. We argue that this is a spin-off of the use of tasks in a testing situation, which 'elicit data which can be used for purposes of measurement' (Bygate, Skehan, & Swain, 2001, p. 12). According to Bygate, Skehan, and Swain, it is clear that in the case of testing tasks it is highly relevant that one achieves 'control over what happens with the task (...) and 'some degree of precision and control of variables in establishing causality in the effects which may be found' (pp. 11-12). For that reason, we find it defensible to use oral tests that direct learners to the use of predetermined grammatical structures.

Besides the use of post-test data to measure the effects of learners focusing on form, we suggest that the use of online measures is also highly valuable. For example, the occurrence of language related episodes (LREs) recently have proven to be a valid measure of learners' focusing on form during collaborative task performance (see Kim, 2013; Kim & McDonough, 2011). Therefore, we used LREs in Study 1 to measure learners' talk about the use of the target structure, during planning time.

As for measuring fluency in a research setting, we would recommend the use of objective measures such as speech rate, articulation rate, and the number of silent pauses since such measures exclude judgements of accuracy, accent or lexicon and may be used for cross-study comparisons. Correlations with human-rated measures were statistical significant for speech rate and number of silent pauses. In addition, both human and automatic measures demonstrated a similar trade-off result. When using human raters we would suggest to instruct them to assess for speech rate and pausing.

2.4 Generalisation

It is important to consider that the findings of the three studies are based on the effects on the accurate use of only two German grammatical structures in a foreign language context. The effects of the form foci may be different for grammatical structures of other second or foreign languages. Moreover, the complexity of the

structures in our studies is highly determined by the relatedness of the L2 to the learners' L1, which is Dutch. Since Dutch and German are related Germanic languages that share similarities, we cannot be sure whether the findings of our studies also apply to educational settings in which learners have a different language background.

A threat to the external validity could also be the fact that all individual participants came from just one Dutch secondary school. However, we have no specific reasons to think that these students differ much from other ninth-grade students in secondary education in the Netherlands. However, it should be recognized that the results of the current set of studies were collected from students who were used to working with communicative tasks.

3. DIRECTIONS FOR FUTURE RESEARCH

We have investigated the effects of a focus on form in all three stages of the TBLT framework (see experiments in chapter 3, 4, 5). What do the findings mean for further research on FonF in task-based language teaching? We will address this question for all three stages of the TBLT framework.

3.1 FonF in the pre-task: Different focus on form and meaning instructions

From the findings of Study 1 we can only draw conclusions on the effectiveness of a form versus a meaning condition in pre-task modeling. Study 1 does not provide insights into whether learners who are provided with models benefit more than learners who are not provided with them. Further research could investigate this issue with, for example, a model/no-model design. Moreover, it would also be interesting to examine the differences in effect between the observation of models as a pre-task, and another type of pre-task, the performance of a similar task (Prabhu, 1987). Furthermore, it would be interesting to experiment with different focus on form and meaning instructions in relation to other grammatical structures and task types..

3.2 FonF during-task: Which intervention for which structure?

Study 2 is the only study in which the effectiveness of the treatment depended on the target structure that was taught. Since we only investigated two different German target structures, and German target structures are scarce in other empirical research on corrective feedback (Lochtman, 2002), additional research could specify more precisely which German grammatical structures benefit from which type of corrective feedback and in what way the L1 could contribute to this process.

3.3 FonF in the post-task: Consolidating effects

In Study 3 the interval between the initial and repeated task performance was two weeks. This did not lead to more accurate use of the dative and comparative forms in

oral task performance. Further research could explore which interval between the initial and repeated task would promote more accurate use of the target structures. In addition to the interval, it would also be interesting to examine whether the number of repetitions could affect the accurate use of the target structures. For example, studies by Lynch and Maclean (2000, 2001) showed that when students repeat their task performance several times in a ‘poster presentation carousel’, their oral accuracy improved.

4. IMPLICATIONS FOR EDUCATIONAL PRACTICE

We argue that the findings of our studies may have important implications for language pedagogy in a foreign task-based language learning context, as all three studies have been carried out in real classroom settings, and included lesson materials and tests that can be used in school practice. In addition, this PhD project has shown that TBLT research can be conducted in regular classroom situations by teachers who are in search of empirical evidence for what they are doing.

The fact that a focus on grammar in meaning-focused TBLT is effective and feasible in real classrooms may motivate other second and foreign language teachers to experiment with such tasks. We would like to underline however that this study is not a plea to convince teachers to exchange their textbooks for a task-based syllabus. For teachers who find that the ‘strong’ meaning focused approach to TBLT is not applicable in their second or foreign language lessons, it is also possible to use an approach that allows for the use of the textbook alongside the more communicative TBLT tasks (see Ellis 2003; 2009). It could, for instance, prove effective to introduce a task at the beginning of a textbook chapter and not just at the end (see also ‘whole-task-first-model’, Janssen, 2012). By presenting the task at the end, continuation of the traditional PPP (present-practice-produce) practice could be very likely (see Ellis, 2014). Presenting the task at the beginning of each chapter, could be a first step towards a more meaning-based second and foreign language education in textbook based school contexts.

Regarding a focus on form in TBLT, we conclude from our findings that this can be effectively achieved in all three stages of the TBLT framework. The choice for the place of a FonF and the FonF strategy itself may depend on: 1) how explicit the teacher wants to focus on form, and 2) in the case of feedback on complexity difficulty of the target structure. What are the choices to be made in foreign language classes based on the results of this study?

During the *pre-task*, we would not recommend explicit rule teaching because it is very likely that learners will only focus on the correct use of the grammatical structure and not on the content (meaning) of the task. At this stage, we argue that grammatical structures should be made salient to enable learners to notice them. Such an implicit focus could be achieved through the observation of models (as in Study 1), or through different forms of visual input enhancement, such as highlighting, color-coding, bolding, italics, and underlining (Doughty & Williams, 1998). These pre-task techniques, however, can only introduce linguistic structures into the productive system (Whitehurst & Vasta, 1975). Undoubtedly, further assistance in using the

input correctly is necessary. An effective way to give learners the chance to practice the use of the grammatical structures is by means of communicative focused tasks that elicit the use of such structures. In this way, learners are pushed to produce the structures and may notice a gap between what they want to say and what they are able to say (Swain, 1995). Noticing this gap may result in a search for the correct answer and lead to modified output, hence restructure learners' interlanguage. In addition, more complex grammatical rules may require rule explanations when learners notice a gap in their knowledge, while preparing their own task performance.

In the *during-task* stage learners prepare, practice, and perform their oral task performance. In our classes, learners often prepare their performance with help from written notes and then practice with another student, while the teacher wanders through the classroom and provides learners with corrective feedback when they make mistakes. As the results of Study 2 showed, the implicit feedback form recasts, may be effectively used for more simple grammatical structures which underlying rules are related to the L1 and therefore easily discovered by the students themselves. In addition, these simple rules do not require much rule processing. Results also showed that the feedback type prompts, which provides learners with metalinguistic information and subsequently elicits the use of the structure, can be effectively used for both simple and complex grammar rules. Considering the fact that prompts may interrupt the flow of communication quite obtrusively, we would suggest the use of prompts for complex grammatical structures that are not related to the L1, which require much rule processing, have large scope and are reliable. A German two-way preposition that requires the dative case is a reliable rule with a large scope and learners have to take several steps to arrive at the correct dative form. In other words, the effort of learning a rule is justified here. By contrast, we observed that in the case of the more simple comparative structure, learners deduced the rule 'add -er to the adjective/adverb' from the input. We only made them aware of the similarities with the rule in their L1, Dutch. The additional rule 'add an Umlaut mark over a, o, or u in most one-syllable adjectives/adverbs' has too many exceptions to be learnt as a rule. We therefore argue that these comparative forms are more easily acquired as lexical chunks. This also applies to the irregular comparative forms such as *besser* [better] and *mehr* [more] which are very similar to the forms in Dutch (L1).

For the *post-task*, several possibilities emerge. Learners may repeat the same task or when they find this too monotonous, the teacher may present a different but similar version of this task. Being familiar with the content of the task may free up learners' attentional resources for a focus on accuracy (Bygate, 2001). Another option is that a similar version of this task is not immediately repeated after the main task, but later in the curriculum. Post-task results in Study 3 showed that this strategy was effective in promoting learners' explicit knowledge of the grammatical structures but not in their oral performance. This may indicate that more practice is needed (Bygate, 2001). An interesting option for repetition would be to let learners present their oral task several times to other students. One could think of poster presentations (as suggested by Lynch and Maclean, 2000, 2001) in different groups but also of speed dates in which learners successively present themselves to other stu-

dents. Another suggestion could be to let learners present their bike-skateboard-tennis racket to different interested buyers. The other students would not only have to listen to these presentations but also ask questions. The presenting student could make lexical, grammatical or phonetic adjustments after each presentation. In this way, task repetition could not only lead to a more accurate but also to a more complex or fluent task performance (Bygate, 1996, 2001)

Other researchers have presented more explicit FonF strategies that can be used effectively in the post-task. These are the anticipation of a public performance or a transcription of the main oral performance. Both activities draw on the notion that learners will attend to accuracy and attempt to avoid errors when they know they will be confronted with a post-task that asks them to present their task to the whole class or that will ask them to transcribe their own words, including their own mistakes (Foster & Skehan, 2013). Skehan and Foster (1997) investigated the effects of the ‘threat’ of a public performance, as a post-task activity, and indeed found improvement in accuracy. Both Skehan and Foster (2013) and Qian (2014) examined the effects of post-task transcribing on learners’ oral task performance and found effects on both accuracy and complexity measures. In the case of post-task transcribing, learners are required to transcribe their oral performance and, for example, note their mistakes (Skehan & Foster, 2013). Transcribing can be done individually or in pairs and with or without rewriting options (Qian, 2014).

In line with Doughty and Williams (1998) we suggest that it is perfectly possible to combine implicit and explicit FonF strategies in different stages of the framework. One could, for instance, start with an implicit focus in the pre-task to make learners aware of a grammatical structure and subsequently provide them with feedback in the during-task stage or ask for a transcription of the oral performance in the during-task stage.

To conclude, the Common European Framework of Languages (CEFR, 1997) has been implemented in the Netherlands and used to help set the core objectives for foreign language education in basic secondary education as well as in the development of the examination programs for the modern foreign languages (Feskens, Keuning, Van Til, & Verheyen, 2014). The CEFR requires a communicative approach in language teaching but, at the same time, does not provide guidance on how grammar education can be a part of meaning-focused language education. Secondary teacher training programs and training courses for teachers should provide students and teachers with evidence-based options to design TBLT lessons and in addition should make them aware of the options on how learners attention can be directed to form, during communicative task performance. We hope that this thesis might help teacher educators, curriculum designers and teachers by making more informed decisions on task design and the selection of focus on form strategies.

REFERENCES

- Ahmadian, M. J. (2012). Task repetition in ELT. *ELT Journal*, 66, 380-382.
- Ahmadian, M. J., & Tavakoli, M. (2011). 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), 23-49.
- Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax. *Language Teaching Research*, 12, 183-210.
- Ammar, A., & Spada, N. (2006). One size fits all? Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition*, 28, 543-74.
- Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.
- Anderson, J. R. (2000). *Learning and memory. An integrated approach*. New York: John Wiley.
- Andringa, S., De Glopper, K., & Hacquebord, H. (2011). Effect of explicit and implicit instruction on free written response task performance. *Language Learning*, 61, 868-903.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.
- Bardovi-Harlig, K. (1992). The relationship of form and meaning: A cross sectional study of tense and aspect in the interlanguage of learners of English as a second language. *Applied Psycholinguistics*, 13, 253-278.
- Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General*, 129, 177-192.
- Boersma, P., & Weenink, D. (2007). *Praat: Doing phonetics by computer*. Retrieved from <http://www.praat.org>. Accessed 15.02.2015
- Braaksma, M.A.H., Rijlaarsdam, G., & Van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology*, 94, 405-415.
- Braaksma, M.A.H., Rijlaarsdam, G., Van den Bergh, H., & Van Hout-Wolters, B. H. A. M. (2004). Observational learning and its effects on the orchestration of writing processes. *Cognition and Instruction*, 22(1), 1-36.
- Breen, M. (1987). Learner contributions to task design. In C. Candlin & D. Murphy (Eds.), *Language Learning Tasks* (pp. 23-46). Englewood Cliffs, N.J.: Prentice Hall.
- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp.21-46). Amsterdam: John Benjamins Publishing.
- Bygate, M. (1996). Effects of Task Repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp.136-146). Oxford: Macmillan Heinemann.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23-48). Harlow, UK: Longman.
- Bygate, M., Skehan, P., & Swain, M. (2001). Introduction. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 1-20). New York: Pearson Education Limited.
- Carless, D. (2012). TBLT in EFL settings: Looking back and moving forward. In A. Shehadeh & C. A. Coombe (Eds.), *Task-based language teaching in foreign language contexts: Research and implementation* (pp. 345-358). Amsterdam: John Benjamins.
- Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition*, 15, 357-386.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Company.

- Council of Europe (2001). *Common European framework of reference for languages*. Cambridge: Cambridge University Press.
- Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction*, 2, 109–142.
- Criado, R. (2013). A critical review of the presentation-practice-production model (PPP) in foreign language teaching. In R. Monroy (Ed.), *Homenaje a Francisco Gutiérrez Díez* (pp. 97-115). Murcia: Ed-it.um.
- Crookes, G. (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11, 367-383.
- De Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning*, 46, 529–555.
- De Jong N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2012). Facets of speaking proficiency. *Studies in Second Language Acquisition*, 34, 5-34.
- De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34, 893 - 916.
- De Jong, N. H., & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, 41(2), 385-390.
- De Graaff, R. (1997). The "Experanto" experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, 19(2), 249-276.
- DeKeyser, R. (1993). The effect of error correction on L2 grammar knowledge and oral proficiency. *Modern Language Journal*, 77, 501–514.
- DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition*, 17, 379-410.
- 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 second language acquisition* (pp. 42–63). Cambridge, UK: Cambridge University Press.
- DeKeyser, R. (2010a). Monitoring processes in Spanish as a second language during a study abroad program. *Foreign Language Annals*, 43, 80-92.
- DeKeyser, R. (2010b). Practice for second language learning: Don't throw out the baby with the bathwater. *International Journal of English Studies*, 10, 155–165.
- Dilans, G. (2010). Corrective feedback and L2 vocabulary development: Prompts and recasts in the adult ESL classroom. *The Canadian Modern Language Review*, 66, 787–815.
- Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114–138). Cambridge: Cambridge University Press.
- Doughty, C., & Williams, J. (1998). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197–261). New York: Cambridge University Press.
- Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching*, 31, 117-135.
- Egi, T. (2007). Interpreting recasts as linguistic evidence: The roles of linguistic target, length and degree of change. *Studies in Second Language Acquisition*, 29, 511-537.
- Egi, T. (2010). Uptake, modified output, and learner perceptions of recasts: Learner responses as language awareness. *The Modern Language Journal*, 94, 1-21.
- Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition*, 9, 1-20.
- Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly*, 27, 91–113.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005a). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141-172 .
- Ellis, R. (2005b). Planning and task-based performance: theory and research. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 4–34). Philadelphia, PA: John Benjamins.
- Ellis, R. (2007). The differential effects of oral corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 339–360). Oxford: Oxford University Press.
- Ellis, R. (2009a). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221-246.

- Ellis, R. (2009b). The differential effects of three types of task planning on the fluency, complexity and accuracy in L2 oral production. *Applied Linguistics*, 30, 447–509.
- Ellis, R. (2014). Taking the critics to task: The case for task-based teaching. In M.K. Aishah, S.K. Bhatt, W.M. Chan, S.W. Chi, K.W. Chin, S. Klayklung, M. Nagami, J.W. Sew, T. Suthiwan, I. Walker (Eds.), *Knowledge, Skills and Competencies in Foreign Language Education* (pp. 103–177). Singapore: NUS Centre for Language Studies.
- Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus on form. *System*, 30, 419–432.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28, 339–369.
- Ellis, R., & Sheen, Y. (2006). Reexamining the role of recasts in second language acquisition. *Studies in Second Language Acquisition*, 28, 575–600.
- Ellis, R., & Shintani, N. (2013). *Exploring language pedagogy through second language acquisition research*. London: Routledge.
- Feskens, R. Keuning, J., Van Til, A. & Verheyen, R. (2014). *Performance Standards for the CEFR in Dutch secondary education*. Arnhem: Cito.
- Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18(3), 299–323.
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3(3), 215–47.
- Foster, P., & Skehan, P. (2013). Anticipating a post-task activity: The effects on accuracy, complexity and fluency of L2 language performance. *Canadian Modern Language Review* 69, 3, 249–273.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics* 21(3), 354–375.
- Fotos, S., & Ellis, R. (1991). Communicating about grammar: A task-based approach. *TESOL Quarterly*, 25, 605–628.
- Freed, B. F. (1995). Do students who study abroad become fluent? In B. F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123–148). Amsterdam: John Benjamins.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Erlbaum.
- Gass, S. M., & Mackey, A. (2006). Input, interaction and output: An overview. *AILA Review*, 19, 3–17.
- Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernandez-Garcia, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49, 549–581.
- Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory*. Aldine Publishing Company, Hawthorne, NY.
- Goo, J., & Mackey, A. (2013). The case against the case against recasts. *Studies in Second Language Acquisition*, 35, 127–165.
- Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly*, 36, 543–572.
- Hartley, J. (1998). *Learning and studying: A research perspective*. London: Routledge.
- Hawkes, M. L. (2012). Using task repetition to direct learner attention and focus on form. *ELT Journal*, 66(3), 327–336.
- Hulstijn, J. H., & De Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review*, 11, 97–113.
- Izawa, M. R., French, M. D., & Hedge, A. (2011). Shining new light on the Hawthorne illumination experiments. *Human Factors*, 53(5), 528–547.
- Janssen, F. (2012). *Uitdagend vakonderwijs voor alle leerlingen*. [Challenging content education for all students]. ICLON. Universiteit Leiden.
- Kim, Y. (2013). Effects of pretask modelling on attention to form and question development. *TESOL Quarterly*, 47, 8–35.
- Kim, Y., & McDonough, K. (2011). Using pretask modelling to encourage collaborative learning opportunities. *Language Teaching Research*, 15, 1–17.
- Krashen, S. D. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*, New York: Longman

- Larsen-Freeman, D. (1995). On the teaching and learning of grammar. In F. R. Eckman, D. Highland, P. W. Lee, J. Mileham, & R. Rutkowski Weber (Eds.), *Second language acquisition theory and pedagogy* (pp. 131-148). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Leeman, J. (2003). Recasts and second language development: Beyond negative evidence. *Studies in Second Language Acquisition*, 25, 37-63.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lewis, M. (1996). Implications of a lexical view of language. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 10-16). Oxford: Heinemann.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309-365.
- Lightbown, P. M. (2000). Anniversary article: Classroom SLA research and second language teaching. *Applied Linguistics*, 21, 431-462.
- Lightbown, P. M., & Spada, N. (1999). *How languages are learned*. Oxford: Oxford University Press.
- Littlewood, W. (2004). The task-based approach: Some questions and suggestions. *ELT Journal*, 58(4), 319-326.
- Littlewood, W. (2007). Communicative and task-based language teaching in East Asian classrooms. *Language Teaching* 40(3), 243-249.
- Lochtman, K. (2002). Oral corrective feedback in the foreign language classroom: How it affects interaction in analytic foreign language teaching. *International Journal of Educational Research*, 37, 271-283.
- Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 379-377). Oxford: Oxford University Press.
- Loewen, S., & Philp, J. (2006). Recasts in adults English L2 classrooms: Characteristics, explicitness, and effectiveness. *The Modern Language Journal*, 90, 536-556.
- Long, M. H. (1985). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input and second language acquisition* (pp. 377-93). Rowley, Mass.: Newbury House.
- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. De Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Philadelphia/Amsterdam: John Benjamins.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). New York: Academic Press.
- Long, M. H. (2007). Recasts in SLA: The story so far. In M. H. Long (Ed.), *Problems in SLA* (pp. 75-116). Mahwah, NJ: Lawrence Erlbaum.
- Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Oxford: Wiley-Blackwell.
- Long, M. H., & Crookes, G. (1993). Units of analysis in syllabus design: the case for the task. In G. Crookes & S. M. Gass (Eds.), *Tasks in a pedagogical context* (pp. 9-44). Cleveland, UK: Multilingual Matters.
- Long, M.H., Inagaki, S., & Ortega, L. (1998). The role of negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal*, 82, 357-371.
- Long, M. H., & Robinson, P. (1998). Focus on form: Theory, research and practice. In C. J. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). New York: Cambridge University Press.
- Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based learning. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123-167). Clevedon, England: Multilingual Matters.
- Lynch, T., & McLean J. (2000). Exploring the benefits of repetition and recycling of a classroom task. *Language Teaching Research*, 4(3), 21-250.
- Lynch, T., & McLean, J. (2001). A case of exercising: effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing*. (pp. 141-62). Harlow: Longman.
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 55-85.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-426.

- Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning*, 59, 453–498.
- Lyster, R., & Ranta, L. (1997). Corrective feedback and student uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37–66.
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265–302.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46, 1–40.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford: Oxford University Press.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82, 338–356.
- McDonough, K. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 323–338). New York: Oxford University Press.
- Macrory, G., & Stone, V. (2000). Pupil progress in the acquisition of the perfect tense in French: The relationship between knowledge and use. *Language Teaching Research*, 4, 54–82.
- Miller, P. C., & Pan, W. (2012). Recasts in the L2 classroom: A meta-analytic review. *International Journal of Educational Research*, 56, 48–59.
- Müller-Hartmann, A., & Schocker-von Ditfurth, M. (2011). *Teaching English: Task supported language learning*. Paderborn: Schöningh.
- Nassaji, H. (2009). Effects of recasts and elicitations in dyadic interaction and the role of feedback explicitness. *Language Learning*, 59, 411–452.
- Newton, J. & Kennedy, G. (1996). Effects of communication tasks on the grammatical relations marked by second language learners. *System*, 24, 309–322.
- Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language students. *Language Learning*, 51, 719–758.
- Norris, J., Brown, J. D., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments*. Honolulu: University of Hawaii.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417–528.
- Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.
- Nunan, D. (1999). *Second language teaching and learning*. Boston, MA: Heinle & Heinle.
- Nunan, D. (2004). *Task-based language teaching*. Cambridge: Cambridge University Press.
- Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 109–148.
- Park, S. (2010). The influence of pretask instructions and pretask planning on focus on form during Korean EFL task-based interaction. *Language Teaching Research*, 14, 9–26.
- Pica, T. (1987). Second-language acquisition, social interaction, and the classroom. *Applied Linguistics*, 8, 3–21.
- Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition*, 6(2), 186–214.
- Pienemann, M. (1998). *Language processing and second language development: Processability theory*. Amsterdam: John Benjamins.
- Prabhu, U. S. (1987). *Second language pedagogy*. Oxford University Press.
- Qian, L. (2014). The effects of post-task transcribing on learners' oral performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 129–154). London: John Benjamins.
- Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63(3), 595–626.
- Richards, J. C., Platt J., & Weber H. (1985). *Longman dictionary of applied linguistics*. London: Longman.
- Richards, J. C., & Rodgers, T. (2001). *Approaches and methods in language teaching* (2nd Ed.). Cambridge University Press.
- Rigenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, 14, 423–441.

- Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27–57.
- Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics*, 43, 1–32.
- Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning*, 61, 1–36.
- Robinson, P., & Gilabert, R. (2013). Task-based learning: Cognitive underpinnings. In Carol Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley-Blackwell.
- Rodgers, T.S. (2009). The methodology of foreign language teaching. In K. Knapp, & B. Seidlhofer (Eds.), *Handbook of foreign language communication and learning* (6 ed., pp. 401–438). Berlin: Mouton de Gruyter.
- Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133–164). Philadelphia/Amsterdam: John Benjamins.
- Samuda, V. (2001). Guiding relationships between form and meaning during task performance: the role of the teacher. In M. Bygate, P. Skehan, & M. Swain, (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp.119-140). London: Longman.
- Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. New York: Palgrave Macmillan.
- Sangarun, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 111–141). Amsterdam, the Netherlands: John Benjamins.
- Sato, R. (2010). Reconsidering the effectiveness and suitability of PPP and TBLT in the Japanese EFL classroom. *JALT Journal* 32(2), 189–200.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 127–158.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, *Proceedings of CLaSIC 2010*, Singapore, December 2-4 (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.
- Schoormann, M., & Schlak, T. (2012). Mündliche Korrekturstrategien im Zweit- und Fremdsprachenunterricht: Recasts und Prompts im Vergleich. [Oral corrective feedback in second and foreign language education: A comparison of recasts and prompts]. *Journal of Linguistics and Language Teaching*, 3, 15-59.
- Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research*, 57, 149–174.
- Schunk, D. H. (2007). *Learning Theories: An educational perspective*. 5th Ed. Upper Saddle River, NJ: Pearson Education.
- Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology*, 79, 54–61.
- Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist*, 32, 195–208.
- Sheen, R. (2003). Focus-on-form: a myth in the making. *ELT Journal*, 57, p. 225-33.
- Sheen, Y. (2011). *Corrective feedback, individual differences and second language learning*. Dordrecht, The Netherlands: Springer.
- Shehadeh, A., & Coombe, C. A. (2012). *Task-based language teaching in foreign language contexts: Research and implementation*. Philadelphia: John Benjamins.
- Sheppard, C. (2006). The effects of instruction directed at the gaps second language learners noticed in their oral production. Unpublished PhD thesis, University of Auckland, New Zealand. In R. Ellis (2009b). 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.
- Skehan, P. (1996a). A framework for the implementation of task-based instruction. *Applied Linguistics*, 17, 38–62.
- Skehan, P. (1996b). Second language acquisition research and task-based instruction. in J. Willis and D. Willis (eds.).
- Skehan, P. (1998a). *A cognitive approach to language learning*. Oxford : Oxford University Press.
- Skehan, P. (1998b). Task-based instruction. *Annual Review of Applied Linguistics*, 18, 268-286.

- Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain, (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 167–85). London, England: Longman
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36, 1-14.
- Skehan, P. (Ed.) (2014). *Processing perspectives on task performance*. London: John Benjamins.
- Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1(3), 185–211.
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49, 93-120.
- Skehan, P., Xiaoyue, B., Qian, L., & Wang, Z. (2012). The task is not enough: Processing approaches to task-based performance. *Language Teaching Research*, 16, 170–187.
- Sonnenschein, S., & Whitehurst, G. J. (1983). Training referential communication skills: The limits of success. *Journal of Experimental Child Psychology*, 35, 426–436.
- Sonnenschein, S., & Whitehurst, G. J. (1984). Developing referential communication: A hierarchy of skills. *Child Development*, 55, 1936–1945.
- Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in Second Language Acquisition*. Rowley, MA.: Newbury House.
- Swain, M. (1995) Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in the study of language* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. *Sociocultural theory and second language learning*, 97, 114.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). New York: Routledge.
- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82, 320–37.
- Swan, M. (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics* 26(3), 376-401.
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239 – 277). Amsterdam: Benjamins.
- Van den Branden, K., Bygate, M., & Norris, J. (Eds.) (2009). Introduction. In Task-based language teaching: A reader (pp.1-13). Amsterdam: John Benjamins Publishers.
- Van de Guchte, M., Braaksmā, M., Rijlaarsdam, G, & Bimmel, P. (2013a). *Effective task-based language learning: To prompt or to recast?* Paper presented at the fifteenth biennial EARLI conference for research on learning and instruction, 27-31 August, Munich, Germany.
- Van de Guchte, M., Braaksmā, M., Rijlaarsdam, G, & Bimmel, P. (2013b). *The effects of recasts and prompts in task-based language teaching*. Paper presented at the fifth biennial international conference on task-based language teaching (TBLT), 3-5 October 2013, Banff, Alberta, Canada.
- Van de Guchte, M., Braaksmā, M., Rijlaarsdam, G, & Bimmel, P. (2014). *The effects of task repetition as a post-task activity*. Paper presented at the 24th annual conference of the European second language association EUROSLA, 3-6 September 2014, York, United Kingdom.
- Van de Guchte, M., Braaksmā, M., Rijlaarsdam, G, & Bimmel, P. (2015a). Learning new grammatical structures in task-based language learning: The effects of recasts and prompts. *The Modern Language Journal*, 99(2), pp. 246-262.
- Van de Guchte, M., Braaksmā, M., Rijlaarsdam, G, & Bimmel, P. (2015b). *The effects of focusing on meaning and form in pre-task video model observations*. Paper presented at the sixth biennial international conference on task-based language teaching (TBLT), 16-18 September 2015, Leuven, Belgium.
- Van de Guchte, M., Braaksmā, M., Rijlaarsdam, G, & Bimmel, P. *Focus on form through task repetition* (in press). Language Teaching Research.
- Van de Guchte, M., Braaksmā, M., Rijlaarsdam, G, & Bimmel, P. *Focusing on meaning and form in pre-task video model observations: Effects on planning processes and task performance*. Manuscript submitted for publication.
- Van den Toorn, M., Pijnenburg, W., Van Leuvensteijn, J., & Van der Horst, J. (Eds.). (2007), *Geschiedenis van de Nederlandse taal, [History of the Dutch language]*. Amsterdam: Amsterdam University Press.

- VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12, 287-301.
- VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex Publishing.
- VanPatten, B. (2004). Input processing in SLA. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5-31). Mahwah, NJ: Lawrence Erlbaum.
- Wendel, J. N. (1997). *Planning and second-language narrative production*. Unpublished doctoral dissertation. Temple University, Philadelphia, PA.
- Whitehurst, G.J., & DeBaryshe, B.D. (1989). Observational learning and language acquisition. In G. E. Speidel & K. E. Nelson (Eds.), *The many faces of imitation in language learning* (pp. 251-276). New York: Springer-Verlag.
- Whitehurst, G.J., & Vasta, R. (1975). Is language acquired through imitation? *Journal of Psycholinguistic Research*, 4, 37-59.
- Widdowson, H. (2003). *Defining issues in English language teaching*. Oxford: Oxford University Press.
- Willis, J. (1993). Preaching what we practice - Training what we teach: Task-based language learning as an alternative to P.P.P. *The Teacher Trainer*, 8 (1), 17-20.
- Willis, J. (1996). *A Framework for Task-based Learning*. Harlow: Longman.
- Willis, D., & Willis, J. (1996). Consciousness-raising activities in the language classroom. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 63-76). Oxford: Heinemann.
- Willis, D., & Willis, J. (2001). Task-based language learning. In R. Carter, & D. Nunan, *The Cambridge guide to teaching English to speakers of other languages* (pp. 173-179). Cambridge: Cambridge University Press.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL students' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32, 235-263.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24, 1-27.
- Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology*, 94, 660-668.

AUTHOR INDEX

- Ahmadian, M. J., 19, 73, 74
Ammar, A., 50, 52, 53, 66, 80
Anderson, J. R., 13, 50, 87
Andringa, S., 55
- Bandura, A., 26, 30, 40
Bardovi-Harlig, K., 34
Basturkmen, H., 25
Bimmel, P., 3, 31, 49, 135
Bock, K., 40
Boersma, P., 34, 83, 102
Braaksma, M., 3, 27, 28, 31, 41, 49, 135
Breen, M., 8
Brown, J. D., 9
Bulté, B., 21
Bygate, M., 1, 2, 6, 9, 10, 19, 73, 74, 75, 88, 89, 95, 98, 99, 100, 104, 107, 121, 124, 125, 127, 131, 132
- Campbell, D. T., 97
Carroll, S., 52
Cook, T. D., 97
Coombe, C. A., 73
Council of Europe, 5, 23, 30, 56, 76, 101
Couzijn, M., 27
Criado, R., 5, 6
Crookes, G., 6, 17, 21
- De Bot, K., 20, 52, 66
De Graaff, R., 20, 54, 67
De Jong, N. H., 22, 34, 68, 83, 101, 102
DeKeyser, R., 13, 20, 22, 26, 50, 60, 87, 101
Dilans, G., 52
Dörnyei, Z., 26
Doughty, C., 11, 12, 13, 19, 21, 25, 41, 50, 51, 52, 73, 106, 108
- Egi, T., 52
Ellis, R., 1, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 25, 40, 42, 50, 52, 53, 54, 55, 60, 66, 73, 74, 80, 81, 99, 101, 106, 121, 122, 127, 129
- Erlam, R., 50
- Florijn, A., 101
Florijn, A. F., 101
Foster, P., 2, 17, 19, 21, 25, 29, 33, 35, 73, 88, 89, 103, 108
Fotos, S., 21
Freed, B. F., 102
French, M. D., 1, 27, 53, 56, 76, 80, 101
- Gass, S. M., 10, 19, 34, 73, 74, 75, 87, 104
Gilabert, R., 9
Glaser, B., 35
Goo, J., 50, 52, 57, 66
Griffin, Z. M., 40
- Hacquebord, H., 55
Han, Z., 52
Hartley, J., 5
Hawkes, M. L., 75
Hedge, A., 101
Housen, A., 21
Hudson, T., 9
Hulstijn, J. H., 20, 54, 56, 67, 78, 87, 102
- Inagaki, S., 52
Izawa, M. R., 101
Izquierdo, J., 50, 52, 53, 65
- Janssen, F., 106
- Kim, Y., 25, 27, 28, 29, 33, 34, 73, 104
Kitsantas, A., 27, 30, 40
Krashen, S., 9, 10, 13, 20, 49, 54, 122, 129
- Larsen-Freeman, D., 19
Leeman, J., 52
Levelt, W. J. M., 74, 75

- Lewis, M., 5
 Li, S., 18, 50
 Lightbown, P., 13, 50, 51
 Littlewood, W., 6, 7
 Loewen, S., 25, 50, 52, 54, 66
 Long, M., 5, 6, 7, 8, 10, 11, 12, 13, 16,
 17, 18, 49, 50, 51, 52, 58, 67, 122, 129
 Lynch, T., 19, 73, 75, 105, 107
 Lyster, R., 18, 50, 51, 52, 53, 54, 55, 65,
 66, 68, 73, 80, 98
- Mackey, A., 50, 52, 53, 57, 66, 73
 Maclean, J., 105, 107
 Macrory, G., 80
 McDonough, K., 25, 27, 52, 53, 104
 Miller, P. C., 52
 Müller-Hartmann, A., 6
- Nabei, T., 50, 54
 Nassaji, H., 53
 Nicholas, H., 51, 52
 Norris, J., 1, 2, 6, 9, 10, 34, 73, 121, 127
 Nunan, D., 1, 5, 7, 8, 15
- Ortega, L., 2, 17, 34, 52, 121, 127
- Pan, W., 52
 Park, S., 29, 40
 Philp, J., 52, 53, 54, 66
 Pica, T., 10
 Pienemann, M., 6, 20
 Platt J., 7
 Prabhu, U. S., 5, 13, 16, 105
- Qian, L., 19, 49, 73, 108
- Ranta, L., 18, 51
 Rebuschat, P., 21, 123, 129
 Richards, J. C., 6, 7
 Riggenbach, H., 34, 83
 Rijlaarsdam, G., 3, 27, 28, 31, 49, 78, 135
 Robinson, P., 9, 10, 11, 12, 23, 49, 51
 Rodgers, T., 6, 9
- Saito, K., 18, 50, 52, 73, 98
 Samuda, V., 6, 15, 17
 Samuda, V., 6
 Samuda, V., 15
 Samuda, V., 17
- Sangarun, J., 17, 25, 28, 29, 30, 33, 35,
 47
 Sato, M., 1, 50, 73, 74, 88, 98
 Sato, R., 1, 50, 73, 74, 88, 98
 Schlak, T., 50
 Schmidt, R., 10, 11, 13, 25, 26, 52, 122,
 129
 Schocker-von Ditfurth, M., 6
 Schoonen, R., 102
 Schunk, D. H., 26, 27, 31, 40, 41
 Sheen, R., 11, 25, 52, 54
 Sheen, Y., 11, 25, 52, 54
 Shehadeh, A., 73
 Sheppard, C., 74, 75, 99
 Shintani, N., 15
 Skehan, P., 2, 5, 6, 8, 9, 10, 11, 14, 15,
 16, 17, 19, 20, 21, 22, 23, 25, 26, 29,
 30, 35, 40, 43, 49, 50, 51, 55, 65, 66,
 67, 73, 74, 76, 77, 87, 88, 89, 93, 95,
 98, 99, 100, 103, 104, 108, 121, 122,
 124, 125, 128, 130, 131, 132
 Sonnenschein, S., 27
 Spada, N., 50, 51, 52, 53, 66
 Steinel, M. P., 101
 Stone, V., 80
 Strauss, A., 35
 Swain, M., 9, 10, 26, 27, 49, 52, 66, 100,
 104, 106, 122, 129
 Swan, M., 11, 12, 25, 73, 74, 97
- Tavakoli, M., 19, 22, 73
 Tavakoli, P., 19, 22, 73
 Tonkyn, A., 21, 33
- Van de Guchte, M., 31, 49, 78, 80, 81
 Van den Bergh, H., 27
 Van den Branden, K., 1, 2, 6, 10, 73, 121,
 127
 Van den Toorn, M., 31
 Van Hout-Wolters, B. H. A. M., 28
 VanPatten, B., 10, 11, 23, 25
 Varela, E., 21, 50, 51, 52
 Vasta, R., 40, 106
- Wang, Z., 50
 Weenink, D., 34, 83, 102
 Wempe, T., 22, 34, 83, 102
 Wendel, J. N., 28, 29
 Whitehurst, G.J., 27, 40, 106
 Widdowson, H., 9

- Wigglesworth, G., 21, 33
Williams, J., 11, 12, 13, 19, 25, 41, 73,
106, 108
Willis, D., 2, 5, 6, 9, 14, 15, 16, 17, 18,
42, 73, 121, 128
Willis, J., 2, 5, 6, 9, 14, 15, 16, 17, 18, 42,
73, 121, 128
Xiaoyue, B., 49
Yang, Y., 50, 52, 54, 55, 66, 80
Yoshioka, J., 9
Zimmerman, B. J., 26, 27, 30, 31, 40

SUMMARY

This dissertation presents research on the way in which learners focus on grammar in task-based language teaching (TBLT). Such a focus on grammar during meaningful task performance is called a Focus on Form (FonF). For this PhD research project we conducted three experimental studies which investigated the effect of different instructional focus on form strategies on learners' accurate use of German grammatical structures in a task-based language teaching environment. The thesis contains six chapters: an introduction (chapter 1), a theoretical framework on focus on form in task-based language teaching (chapter 2), followed by three reports of empirical studies (chapters 3, 4, 5), and a discussion chapter (chapter 6).

CHAPTER 1

In chapter 1 we describe how this dissertation arose from different views on language learning between two departments at the same Dutch secondary school: a task-based language department and a regular, more grammar-based, language department. In the task-based department learners acquire the second- or foreign language by performing communicative tasks in the target language. Some teachers of the regular language department feared that the *task-based* curriculum was overly focused on content and pleasure and not on the teaching of specific grammatical structures.

Subsequently, we discuss the role of grammar in TBLT. Literature on TBLT principles shows that although TBLT focuses primarily on meaning (content) it does not exclude grammar entirely. Moreover, TBLT 'even encourages a focus on form in view of optimizing the learning potential of task-based educational activities' (Van den Branden, Bygate & Norris, 2009, p. 6). In contrast to more traditional language methods in which isolated grammar exercises play an important role, a *focus on form* is always related to meaning-based activities. As a consequence, grammar teaching in TBLT is a *tool* for making meaning rather than an object to study (Ellis, 2003). In addition, a focus on form has proven to promote both the efficiency and effectiveness of language acquisition processes, especially during meaning-based activities (see review by Norris & Ortega, 2000).

Next, we describe the framework for task-based instruction (Skehan, 1996a, 1998a; Willis, 1996) which consists of a pre-task, during-task, and post-task stage. These stages follow the natural order of a task-based lesson. In each stage of the task, there is the possibility to direct learners' attention to form. In this PhD project we examined the effects of different focus on form strategies at each stage of the task-based teaching framework.

Study 1 focuses on the *pre-task* stage. In this study we examined the effects of learners focusing on either form or meaning during the observation of peer model videos, on measures of accuracy and complexity. Study 2 focuses on the *during-task* stage. For this study we examined the effects of two different feedback types, prompts and recasts, on the accurate use of two different German grammar structures, the dative case after a two-way preposition and the comparatives. Study 3 is aimed at the *post-task* stage. We examined the effects of learners repeating a task on the accurate use of the same grammar structures as in Study 2.

In addition, we investigated in Study 2 and 3 whether the effectiveness of the FonF strategies corrective feedback and task repetition depended on the grammatical structure involved. Furthermore, in all three studies, we examined whether learners' attention to an accurate task performance had negative consequences for the fluency or complexity of their performance. With this last question we investigated Skehan's (1996a, 1998a) Trade-off Hypothesis which claims that learners have limited attentional resources. As a consequence, learners' attention to the accuracy, complexity or fluency of their task performance would result in less attention being available to devote to the other dimensions and thus diminish their performance therein.

In this dissertation we investigated three major research questions:

- RQ1 Does the FonF strategy positively affect learners' declarative knowledge of the new grammatical structures and their oral task performance in terms of accuracy (Study 1, 2, 3) and complexity (Study1)?
- RQ2 Do the effects depend on the grammatical structure involved? (Study 2, 3)
- RQ3 Does learners' focus on accuracy come at the expense of oral fluency (Study 1, 2, 3) or oral complexity (Study 1)?

Finally, in chapter 1 we outline the organization of the thesis and the content of the chapters.

CHAPTER 2

Chapter 2 describes the theoretical framework of the studies carried out in this project. We review several definitions of a *task* in TBLT and conclude that Ellis' (2003) definition is the one that the studies in this dissertation will be based on. In all three studies the tasks 1) require an outcome in terms of content (meaning); 2) predispose learners to use particular grammatical structures; 3) are connected to real-world situations; and 4) engage learners in cognitive processing.

Subsequently we discuss the cognitive processes that are claimed to promote learning in TBLT. Tasks that provide learners with communicative and comprehensible input (Krashen, 1981), push learners to produce oral output (Swain, 1985), make learners interact (Long, 1981, 2015), and in addition direct learners' attention to the formal properties of the target language (Long, 1991, 2015; Schmidt, 1981) all claim to be beneficial to second and foreign language learning.

Although all supporters of TBLT see a role for grammar in terms of a focus on form during meaningful communication, opinions tend to differ on how such a FonF should be implemented in the TBLT lesson. We review several studies in which researchers have different views on FonF implementation, depending on their view on language learning. For the studies in this dissertation we used both design and methodological procedures to include a focus on form in the TBLT lessons. First, the tasks were designed to elicit the use of predetermined grammatical structures because we wanted to increase the chance that students would use the structures.

Second, attention to these structures was drawn by means of more implicit as well as explicit FonF strategies.

Furthermore, in chapter 2 we discuss the notion that a particular FonF strategy may be effective for one grammatical structure but not for another. The effectiveness may depend on the difficulty of the grammatical structure, which can be determined by the complexity, scope, and reliability of the grammar rule, the developmental readiness of the learner, and the contrast between L1 and L2. For that reason, Study 2 and 3 investigate whether the effectiveness of corrective feedback, in the form of prompts and recasts, and task repetition, depended on the grammatical structure involved. For the grammatical structures, we opted for German comparatives which we consider a simple structure, and the German dative case after a two-way preposition, which we consider a complex structure.

Finally, we discuss two knowledge types that are reflected in the measurements of learners' language knowledge in our studies. Since it is assumed that language learning involves both implicit and explicit knowledge (Rebuschat, 2013), both measures of implicit and explicit knowledge were taken into account. Learners' implicit knowledge was measured by means of meaning-based oral production tasks and assessed in terms of oral accuracy, complexity and fluency. Learners' explicit knowledge was measured through a metalinguistic knowledge test and a fill-in-the-gap exercise.

CHAPTER 3

Chapter 3 reports on the empirical study that investigated the effects of directing learners' attention to either form or meaning in the *pre-task* through guided observation of peer-model videos. Forty-eight ninth-grade students learning German as a foreign language were randomly assigned to two conditions: a focus on form (FonF) and a focus on meaning (FonM) condition. Students in both conditions watched two video clips of a peer student performing a similar task. Written instructions directed participants' attention to either form (FonF) or meaning (FonM). While observing, the FonF group focused on the use of two-way prepositions that require the dative case. The FonM group was asked to compare the rhetorical structure of both presentations and what the students they observed on video had done to make the presentation attractive and persuasive. Think-aloud methods and a communicative oral task were used to measure the effects of these two conditions on both planning processes and task performance. Task performance results showed that the FonF condition generated more (accurate) use of the target structure than the FonM condition, whereas the FonM condition outperformed the FonF condition on complexity in terms of clauses per analysis of speech unit (AS-Unit), ratio of coordinated sentences, and ratio of subordinated sentences. Although results showed a clear effect of the focus on task performance, this effect was not observed in the think-aloud planning protocols.

CHAPTER 4

Chapter 4 describes the empirical study that examined the effects of a focus on form at the *during-task* stage by comparing two types of oral corrective feedback. Sixty-four ninth-grade students learning German as a foreign language were randomly assigned to the two experimental conditions – one received prompts after an incorrect utterance (prompt condition) - the other recasts (recast condition). These experimental conditions were compared to a control condition, which was an intact class.

In the prompt condition, the teacher provided students with metalinguistic feedback on the target structure and then asked them to reformulate their answer. In the recast condition, the teacher reformulated the student's false utterance, minus the error. The control group followed the regular textbook-based curriculum. The study involved two subsequent interventions, which each consisted of a task with a focus on a particular language structure: The first targeted a complex structure, dative case after a preposition; the second a simple structure, comparatives. Pre-tests, immediate, and delayed post-tests included measures of written and oral accuracy as well as oral fluency.

Statistical comparisons on both written and oral post-tests showed that prompts and recasts were both effective when compared to the control group, with prompts producing better results than recasts. Furthermore, the effect of recasts depended on the structure involved: Recasts were more effective for the comparative than for the dative on written accuracy. In addition, the recast group performed the dative task more fluently than the prompt group did.

Finally, a negative correlation was observed for the prompt group in the delayed oral post-test of the dative structure. In other words, students who performed the task more accurately showed less fluency, a result that seems to confirm Skehan's (1996a, 1998b) Trade-off Theory, which claims that increased attention to accuracy will lead to a decrease in fluency.

CHAPTER 5

Chapter 5 reports on the empirical study that examined the effects of the repetition of a similar task as a *post-task* activity, after having directed learners' attention to form in the *during-task* stage. Forty-eight ninth-grade students learning German as a foreign language were randomly assigned to two conditions: One group repeated a similar task (R); the other group did not (NR). The study comprises two interventions, which each consisted of a task with a focus on a particular grammatical structure. The first intervention targeted the German dative case after a preposition; the second German comparatives. Pre-tests, immediate and delayed post-tests measured metalinguistic knowledge, written and oral accuracy as well as oral fluency.

Results showed that the R-condition outperformed the NR-condition on both written accuracy and metalinguistic knowledge. However, no statistically significant differences between conditions were found on oral accuracy. This may indicate that automatization of these structures in oral performance requires more practice on different versions of a task (Bygate, 2001). In addition, no interactions were observed between the effect of condition and task which indicates that the effect of task repe-

tion did not depend on the structure involved. Furthermore, this study confirms Skehan's (1996a, 1998b) Trade-off Theory which claims that people have limited attentional resources. The negative interaction we observed within the Repetition condition on the dative structure provides additional evidence that accuracy may come at the expense of fluency.

We conclude that the repetition of a similar task was effective in promoting learners' explicit knowledge on the grammar structures. This implies that previous knowledge on the target structure is available to learners to build on during subsequent task performance.

CHAPTER 6

Chapter 6 is the concluding chapter in which we first summarize the main results of the three separate studies in this thesis. Subsequently, we raise several methodological issues that need to be considered. Finally, for each stage of the TBLT framework directions for future research and implications for educational practice are presented.

Based on the findings of the three empirical studies carried out in this project, the overall picture seems to be that a focus on form is possible and effective in all three stages of the TBLT framework. Based on the results, we are led to the conclusion that the pre-task lends itself to a more implicit FonF that promotes learners' noticing of the grammatical structure with the aim to use it accurately during subsequent main task performance. Feedback at the during-task stage can be provided effectively in both implicit or explicit ways, depending on the complexity and L1-relatedness of the structure at hand. Post-task repetition, after having explicitly focused on the structures at the during-task stage, may raise or re-awaken learners' metalinguistic knowledge of the structures. However the repetition of a similar task did not lead to gains in oral accuracy, which may indicate that more practice of the same task type is needed to ensure such an effect (Bygate, 2001).

As for the second research question in our thesis, whether the effect of the FonF intervention depended on the target structure, we conclude that this was only the case for corrective feedback. Compared to prompts, recasts had a larger effect on the comparative structure than on the dative structure. We therefore conclude that simple structures, which are related to the L1, may be corrected effectively through recasts. The advantage of recasts is that they do not interrupt learners' flow of communication in an obtrusive way. Furthermore, the recast group performed the dative task more fluently than the prompt group did.

All three studies provided evidence for Skehan's (1996a, 1998b) Trade-off Theory which was the focus of our third research question. The more explicit foci on a complex grammar structure such as prompts and task repetition (preceded by rule deduction and prompts during the main task) came at the expense of fluency, whereas the more implicit foci such as recasts and the observation of models did not. In addition, an implicit focus on a complex structure during the observation of models led to lower scores on complexity measures.

Furthermore, we reflect on issues that concern the way in which the FonF strategies were operationalized and the treatment tasks were designed. In general, designing

tasks that elicit the use of a specific grammatical structure, while primarily focusing on meaning is difficult to achieve. Other relevant methodological issues we discuss are measurements, ecological validity and the generalizability of the effects.

Next, we describe how focus on form research could be extended for each stage of the TBLT framework. For the *pre-task* stage a challenging task for further research would be to investigate the effects of different focus on form and meaning instructions in relation to other grammatical structures and task types. As for the *during-task stage*, research on corrective feedback regarding German grammatical structures is scarce, additional research could specify more precisely which German grammatical structures benefit from which type of CF and in what way the L1 may contribute to this process. *Post-task* research could be extended by examining whether more practice of similar tasks would lead to improvement in oral accuracy.

Finally, we conclude the final chapter by stating that this thesis is not a plea to convince teachers to exchange their textbooks for a task-based syllabus. We suggest that teachers can also use communicative tasks alongside their textbooks. In addition, we make several suggestions on how to focus on grammar when working with communicative tasks. Furthermore, we hope that this thesis might help teacher educators, curriculum designers and teachers by making more informed decisions on task design and the selection of focus on form strategies.

SUMMARY (DUTCH)

Dit proefschrift gaat over hoe leerlingen aandacht voor grammatica kunnen hebben binnen taakgericht taalonderwijs. Een dergelijke focus op grammatica binnen betekenisvol taakgericht taalonderwijs wordt ook wel *Focus on Form* (FonF) genoemd. Voor dit onderzoeksproject hebben we drie experimentele studies uitgevoerd die het effect van verschillende focus on form instructiestrategieën op het correct gebruik van Duitse grammaticale structuren onderzoeken. Dit proefschrift omvat zes hoofdstukken: een inleidend hoofdstuk (hoofdstuk 1), een theoretisch kader over focus on form binnen taakgericht taalonderwijs (hoofdstuk 2), gevolgd door drie empirische studies (hoofdstuk 3, 4, 5) en een discussiehoofdstuk (hoofdstuk 6).

HOOFDSTUK 1

Hoofdstuk 1 begint met de totstandkoming van dit proefschrift. De auteur is docente Duitse taal en cultuur op een Nederlandse middelbare school waar twee verschillende afdelingen naast elkaar bestaan: een Quest en een reguliere afdeling. Het moderne vreemdetalenonderwijs in de Quest-afdeling is taakgericht en in de reguliere afdeling meer grammatica-gericht. In taakgericht taalonderwijs (TT) verwerven leerlingen de vreemde taal door het uitvoeren van communicatieve taaltaken in de doeltaal. Sommige docenten van de reguliere afdeling vreesden dat het taakgericht taalonderwijs binnen de Quest-afdeling te zeer gericht was op het overbrengen van inhoud en op plezier en te weinig op het aanleren van specifieke grammaticale structuren. Dit dilemma – aandacht voor het overbrengen van de inhoud versus aandacht voor grammatica – was het uitgangspunt voor de drie experimentele studies naar de effecten van focus on form strategieën binnen taakgericht taalonderwijs in dit proefschrift.

Vervolgens wordt in hoofdstuk 1 de rol van grammatica binnen taakgericht taalonderwijs besproken. Literatuur over TT laat zien dat, hoewel TT zich voornamelijk richt op het overbrengen van inhoud, er ook ruimte is voor grammatica. Sterker nog, TT moedigt aandacht voor grammatica binnen een betekenisvolle context juist aan, om op die manier de leermogelijkheden van TT te optimaliseren (Van den Branden, Bygate & Norris, 2009). Terwijl geïsoleerde grammaticale oefeningen binnen de meer traditionele methoden een belangrijke rol spelen, is een *focus on form* in TT altijd verbonden met betekenisvolle activiteiten. Het onderwijzen van grammatica binnen TT is daarom een middel om te kunnen communiceren en geen op zichzelf staand leerdoel (Ellis, 2003). Bovendien is bewezen dat een *focus on form* tijdens betekenisvolle activiteiten effectief is voor de bevordering van taalverwervingsprocessen (zie de review van Norris & Ortega, 2000).

Daarna beschrijven we de volgorde waarin een taakgerichte les kan verlopen, bestaande uit een *pre-*, *during-*, and *post-task* fase (zie *framework for task-based*

instruction, Skehan, 1996a, 1998a; Willis, 1996). In elk van deze fasen is het mogelijk om aandacht voor grammatica te hebben. De drie studies in dit proefschrift onderzoeken de effecten van verschillende FonF strategieën in zowel de *pre-task*, *during-task*, als *post-task* fase.

Studie 1 richt zich op de *pre-task* fase. In deze studie onderzochten we de effecten van leerlingen die hun aandacht richten op grammatica of de inhoud tijdens het observeren van modelvideo's waarin leeftijdgenoten eenzelfde soort taak uitvoeren. De effecten van deze observaties werden onderzocht op het correcte mondelinge gebruik van de Duitse grammaticale structuur datief (derde naamval) na een wisselvoorzetsel en de complexiteit van de opeenvolgende mondelinge taakuitvoering.

Studie 2 richt zich op de *during-task* fase. Voor deze studie onderzochten we de effecten van twee verschillende vormen van mondelinge feedback, *prompts* en *recasts* nadat leerlingen een fout hadden gemaakt. De effecten werden onderzocht op de mondelinge en schriftelijke correctheid van twee Duitse grammaticale structuren: de comparatief (de vergrotende trap) en het gebruik van de datief (derde naamval) na een wisselvoorzetsel.

Studie 3 richt zich op de *post-task* fase. Voor deze studie onderzochten we de effecten van leerlingen die een soortgelijke taak als in de *during-task* fase herhaalden. De effecten werden onderzocht op de mondelinge en schriftelijke correctheid van dezelfde twee grammaticale structuren als in studie 2.

Daarnaast hebben we in studie 2 en 3 onderzocht of de effectiviteit van de FonF strategieën correctieve feedback en het herhalen van een taak, afhankelijk was van de grammaticale structuur. Verder gingen we in alle drie de studies na of de aandacht van leerlingen voor de correctheid van de taakuitvoering ten koste zou gaan van de vloeiendheid of complexiteit van diezelfde taakuitvoering. Met deze laatste vraag wilden we Skehan's (1996a, 1998a) Trade-off Hypothese onderzoeken die stelt dat leerlingen maar een beperkte hoeveelheid aandacht tot hun beschikking hebben en dat correctheid, complexiteit en vloeiendheid met elkaar in competitie zijn.

In dit proefschrift onderzoeken we drie hoofdvragen:

- RQ1 Heeft de FonF strategie een positief effect op de declaratieve (feitelijke) kennis van de leerlingen over de grammaticale structuren (Studie 2,3), de mondelinge correctheid van diezelfde structuren (Studie 1, 2, 3) en op de complexiteit van de taakuitvoering (Studie 1)?
- RQ2 Hangt het effect van de FonF strategie af van de grammaticale structuur (Studie 2, 3)?
- RQ3 Gaat de aandacht van leerlingen voor de correctheid van de taakuitvoering ten koste van de vloeiendheid (Studie 1, 2, 3) of complexiteit (Studie 1) van diezelfde taakuitvoering?

Tenslotte beschrijven we in hoofdstuk 1 de indeling van het proefschrift en de inhoud van de hoofdstukken.

HOOFDSTUK 2

Hoofdstuk 2 omvat het theoretisch kader van dit proefschrift over aandacht voor grammatica binnen taakgericht taalonderwijs. Allereerst bespreken we welke verschillende definities er gebruikt worden voor een taaktaak binnen TT. Voor de taken die gebruikt zijn in dit proefschrift hanteren we de taakdefinitie van Ellis (2003). Alle taken in de drie studies: 1) resulteren in een communicatief eindproduct; 2) zijn gerelateerd aan levensechte situaties; 3) lokken het gebruik van een vooraf bepaalde grammaticale structuur uit; en 4) zetten cognitieve processen in gang.

Vervolgens bespreken we de cognitieve processen waarvan beweerd wordt dat deze bijdragen aan taalverwerving binnen TT. Allereerst kunnen taaltaken leerlingen voorzien van rijke en begrijpelijke input (Krashen, 1981). Daarnaast kunnen taken leerlingen ertoe aanzetten om zich mondeling in de vreemde taal te uiten (Swain, 1985) en met elkaar te laten communiceren (Long, 1981, 2015). Als laatste kan bij het uitvoeren van een taak de aandacht van de leerlingen gericht worden op de formele aspecten van taal zoals grammatica (Long, 1991, 2015; Schmidt, 1981).

Hoewel alle voorstanders van TT een rol voor grammatica zien in de vorm van een *focus on form* binnen betekenisvolle communicatie, verschillen de inzichten over hoe dit moet gebeuren en op welk moment. We bespreken in hoofdstuk 2 de voorstellen van verschillende onderzoekers, gebaseerd op hun ideeën over taalverwerving.

Voor de studies in dit proefschrift hebben we twee verschillende manieren benut om leerlingen te laten focussen op grammatica. Allereerst ontlocken de taken door het design het gebruik van een vooraf bepaalde grammaticale structuur. Zo verhoogden we de kans dat de structuur tijdens de taakuitvoering ook daadwerkelijk zou worden gebruikt. Daarnaast werd de aandacht van de leerlingen op deze structuren gevestigd door het gebruik van zowel impliciete als expliciete FonF-strategieën.

Verder komt in dit hoofdstuk aan de orde dat een bepaalde FonF-strategie effectief kan zijn voor de ene grammaticale structuur maar niet voor de ander. Die effectiviteit zou afhankelijk kunnen zijn van de moeilijkheid van de structuur. Om die reden onderzoeken we in studie 2 en 3 of het effect van correctieve feedback en het herhalen van taken afhangt van de grammaticale structuur. Als eenvoudige grammaticale structuur is gekozen voor de comparatief (vergroten trap) omdat deze structuur in het Duits veel overeenkomsten vertoont met het Nederlands. Als complexe structuur kozen we het gebruik van de datief na een wisselvoorzetsel, die juist geen overeenkomsten vertoont met het Nederlands.

Ten slotte beschrijven we de twee kennistypen die een rol spelen bij taalverwerving, te weten impliciete en expliciete kennis (Rebuschat, 2013). Beide kennistypen hebben we in onze studies onderzocht. De impliciete kennis van de leerlingen werd gemeten via betekenisvolle mondelinge productietaken. De mondelinge taken van leerlingen werden vervolgens beoordeeld op het gebied van mondelinge correctheid, complexiteit en vloeiendheid. De expliciete kennis van de leerlingen over de grammaticale structuren werd gemeten met een regelkennistest en een grammaticale inuloefening.

HOOFDSTUK 3

In hoofdstuk 3 rapporteren we over de empirische studie (voortoets-natoets-design) die de effecten onderzoekt van leerlingen die tijdens het observeren van *peer-model* video's hun aandacht ofwel op grammaticale structuren richten of op de inhoud van de taak. Achtenveertig leerlingen uit de derde klas havo/vwo die anderhalf jaar Duits als vreemde taal hadden gevolgd deden mee aan deze studie. De leerlingen werden aselekt toegewezen aan twee condities: een focus op grammaticale structuren (FG) of een focus op de inhoud (FI). Leerlingen in beide condities bekeken dezelfde twee videofragmenten van leeftijdgenoten die een rondleiding door de aula van de school gaven. Geschreven instructies zorgden ervoor dat de aandacht van de FG-conditie uitging naar de grammaticale structuur: wisselvoorzetsel-plus-datief (derde naamval). De FI-conditie richtte zich op de retorische structuur van de rondleidingen en wat de leeftijdgenoten deden om deze aantrekkelijk en overtuigend te maken. Een mondelinge test-taak waarbij de leerling nu zelf een rondleiding door de aula moest geven, werd gebruikt om de prestatie van de leerling te meten op het gebied van correctheid, complexiteit en vloeiendheid. Daarnaast onderzochten we (met behulp van hardop-denken-protocollen) of leerlingen tijdens het plannen van de taak ook al verschillen vertoonden in aandacht voor grammaticale correctheid of inhoudelijke aspecten.

Uit de resultaten van de metingen blijkt dat de FG-conditie meer algemeen en correct gebruik van de wisselvoorzetsel-plus-datief structuur laat zien dan de FI-conditie. De FI-conditie daarentegen laat een complexere taakuitvoering zien. Dit effect was echter niet zichtbaar in de hardop-denken-protocollen van de leerlingen tijdens het plannen van deze taak.

De *Trade-off Hypothese* van Skehan (1996a, 1998a) werd in deze studie bevestigd: het (correct) gebruik van de datief na een wisselvoorzetsel correleerde negatief met de complexiteit van diezelfde taakuitvoering.

HOOFDSTUK 4

Hoofdstuk 4 beschrijft de empirische studie (voortoets-natoets-design) die de effecten van een *focus on form* tijdens de *during-task* fase onderzoekt door twee vormen van mondelinge correctieve feedback met elkaar te vergelijken. Aan deze studie deden vierenzestig leerlingen uit de derde klas havo/vwo mee die anderhalf jaar Duits als vreemde taal hadden gevolgd. Deze leerlingen kwamen uit een ander cohort dan de leerlingen uit de studies in hoofdstuk 3 en 5. De leerlingen werden aselekt toegewezen aan twee experimentele condities: in de ene conditie ontvingen leerlingen een 'prompt' na het maken van een mondelinge fout – in de andere conditie een 'recast'. Deze twee experimentele condities werden vergeleken met een andere klas, die als controlegroep in het ontwerp fungeerde.

In de *prompt*-conditie gaf de docent de leerling metalinguïstische feedback (bijvoorbeeld de grammaticaregel) ten aanzien van de grammaticale structuur en vroeg hem of haar om het antwoord te herformuleren. In de *recast*-conditie herformuleerde de docent de incorrecte uiting van de leerling, zonder de fout. De controlegroep volgde tijdens de interventie het schoolboek *Na klar* voor het vak Duits, waarin zo-

wel schriftelijke als mondelinge oefeningen ten aanzien van de wisselvoorzetels plus datief als de vergrotende trap aan bod kwamen. Bij de experimentele condities vonden er twee opeenvolgende interventies plaats. De eerste betrof een taak waarbij leerlingen hun droomkamer moesten beschrijven. Leerlingen kregen hierbij feedback op de structuur 'datief na een wisselvoorzetel'. Tijdens de tweede interventie voerden leerlingen een taak uit waarbij ze twee mobiele telefoons moesten vergelijken. De feedback was hier gericht op de comparatief (vergotende trap).

De resultaten van de nametingen laten zien dat zowel de *prompt*-conditie als *recast*-conditie beter presteerde dan de controlegroep op zowel schriftelijke als mondelinge metingen. Daarnaast behaalde de *prompt*-conditie betere resultaten dan de *recast*-conditie op zowel de mondelinge als schriftelijke metingen. Verder bleek dat het effect van de *recasts* afhankelijk was van de grammaticale structuur: *recasts* waren effectiever voor de eenvoudige comparatief structuur dan voor de complexere datief structuur. Daarnaast voerde de *recast*-conditie de datief taak (beschrijving droomkamer) vloeiender uit dan de *prompt*-conditie. De *Trade-off Hypothese* van Skehan (1996a, 1998a) werd ook in deze studie bevestigd: leerlingen uit de *prompt*-conditie die hoger scoorden op mondelinge correctheid voerden de taak minder vloeiend uit.

HOOFDSTUK 5

In hoofdstuk 5 doen we verslag van de empirische studie (voortoets-natoets-design) naar de effecten van het herhalen van een soortgelijke taak als *post-task* activiteit. Achtenveertig leerlingen uit de derdeklas havo/vwo die anderhalf jaar Duits als vreemde taal hadden gevolgd, namen deel aan deze studie. De leerlingen waren afkomstig uit een ander cohort dan de leerlingen uit de studies in hoofdstuk 3 en 4. Binnen de studie waren er twee opeenvolgende interventies. De eerste betrof een taak waarbij leerlingen hun droomkamer moesten beschrijven. Alle leerlingen kregen hierbij feedback in de vorm van *prompts* op de structuur 'datief na een wisselvoorzetel'. Tijdens de tweede interventie voerden leerlingen een taak uit waarbij ze twee vakanties moesten vergelijken. De feedback, in de vorm van *prompts* was voor alle leerlingen gericht op de comparatief (vergotende trap).

Binnen elke klas waren de leerlingen aselekt toegewezen aan twee condities: de ene conditie (H) herhaalde een soortgelijke taak als de hierboven beschreven hoofdtak; de andere conditie deed dit niet (NH).

De resultaten van de nametingen lieten zien dat de H-conditie beter presteerde op regelkennis en schriftelijke correctheid. Dit was niet het geval voor mondelinge correctheid. Hieruit zouden we kunnen concluderen dat mondelinge correctheid van deze structuren om hogere automatisering van de structuren vraagt, wat bereikt zou kunnen worden door meer oefening met soortgelijke taken (Bygate, 2001).

Het effect van taakherhaling bleek niet afhankelijk van de structuur; taakherhaling bleek even effectief voor de comparatief als de datief. Verder werd ook in deze studie Skehan's (1996a, 1998a) *Trade-off Hypothese* bevestigd: leerlingen uit de herhaalconditie die hoger scoorden op mondelinge correctheid voerden de taak minder vloeiend uit.

HOOFDSTUK 6

In het slothoofdstuk vatten we eerst de voornaamste bevindingen uit de drie experimentele studies samen. Daarna reflecteren we op verschillende methodologische beslissingen die we hebben genomen. Vervolgens beschrijven we mogelijkheden voor vervolgonderzoek en noemen we verschillende implicaties voor de onderwijspraktijk.

Ten aanzien van de eerste onderzoeksvraag concluderen we dat een *focus on form* mogelijk én effectief is in alle drie de fasen van het *TBLT framework*. De *pre-task* leent zich voor meer impliciete FonF-strategieën die de aandacht van de leerder voor de structuren kunnen vergroten met als doel deze (correct) in de daaropvolgende taak te kunnen gebruiken. Het expliciet leren van grammaticale structuren lijkt ons hier minder geschikt omdat wij vinden dat de focus in de *pre-task* fase voornamelijk gericht moet zijn op de inhoud van de taak. In de *during-task* fase kan feedback op zowel impliciete als expliciete wijze worden gegeven. Alhoewel *recasts* en *prompts* beiden effect hadden, bleken de expliciete *prompts* effectiever te zijn dan de impliciete *recasts*. Het herhalen van eenzelfde type taak in de *post-task* bleek effectief te zijn voor het vergroten van regelkennis en van schriftelijke correctheid van grammaticale structuren. De herhaling leidde in onze studie echter niet tot effecten op mondelinge correctheid, wat kan betekenen dat voor effecten daarop meer oefening met soortgelijke taken nodig is (Bygate, 2001).

Ten aanzien van de tweede onderzoeksvraag concluderen we dat het effect van de FonF-strategie alleen bij het feedbacktype *recasts* verschillend was voor de verschillende grammaticale structuren. *Recasts* bleken effectiever voor de eenvoudige comparatiefstructuur dan voor de complexere datiefstructuur. Om die reden concluderen wij dat eenvoudige regels, die overeenkomsten vertonen met de moedertaal, ook effectief kunnen worden gecorrigeerd door middel van *recasts*. *Recasts* hebben bovendien het voordeel dat ze de ‘communicatieve flow’ van de leerling niet al te veel verstoren.

Met de derde onderzoeksvraag onderzochten we Skehan’s (1996a, 1998a) *Trade-off Hypothese*. Alle drie de studies lijken deze hypothese te ondersteunen. De meer expliciete FonF-strategieën zoals *prompts* en taakherhaling (voorafgegaan door *prompts* in de *during-task* fase) zorgden bij de complexe datiefstructuur voor een minder vloeiende taakuitvoering. Dit was niet het geval voor de impliciete FonF strategieën *recasts* en het observeren van modelvideo’s. Ook al leidde de impliciete focus op de complexe datief tijdens het observeren van model video’s niet tot lagere scores op vloeiendheid, het had wel een minder complexe taakuitvoering tot gevolg.

Verder reflecteren we in hoofdstuk 6 op de keuzes die we hebben gemaakt ten aanzien van het operationaliseren van de FonF-strategieën en het design van de interventietaken. Over het algemeen kunnen we concluderen dat het een lastige opgave is om taken te ontwerpen die als voornaamste doel het overbrengen van de inhoud hebben, maar ook het gebruik van een vooraf bepaalde grammaticale structuur uitlokken. Daarnaast gaan we in op kwesties over het meten van de variabelen, ecologische validiteit en generalisatie.

Suggesties voor vervolgonderzoek op het gebied van *focus on form* worden gedaan voor elke fase van het *TBLT framework*. Voor de *pre-task* stellen we voor om

te onderzoeken welke *focus on form* instructies tijdens observerend leren effectief zouden kunnen zijn voor het (correct) gebruik van grammaticale structuren en welke *focus on meaning* instructies een complexere taakuitvoering zouden bevorderen. Op het gebied van de correctieve feedback in de *during-task* fase kunnen we stellen dat er ten aanzien van het verwerven van Duitse grammaticale structuren weinig onderzoek is verricht. Aanvullend onderzoek zou meer duidelijkheid kunnen geven over welk type feedback geschikt zou zijn voor welke Duitse structuren en op welke manier de moedertaal kan bijdragen aan dit proces. Vervolgonderzoek in de *post-task* zou uitsluitel kunnen geven of meer oefening van hetzelfde taaktype zou leiden tot hogere scores op het gebied van mondelinge correctheid.

We eindigen het hoofdstuk door te stellen dat dit proefschrift geen pleidooi is om alle vreemdetalendocenten hun schoolboek te laten vervangen door een taakgericht curriculum. Wel bevelen we aan communicatieve taken te gebruiken zijn *naast* de bestaande methode. Het zou bijvoorbeeld een goed idee zijn om leerlingen aan het begin van een hoofdstuk een taaltaak te presenteren, zodat het communicatieve doel van het onderwijs dat volgt voor de leerlingen helder is en de uit te voeren leeractiviteiten betekenisvoller kunnen zijn. Daarnaast doen we voorstellen hoe docenten tijdens het taakgericht onderwijs aandacht voor grammatica kunnen genereren. Verder hopen we dat dit proefschrift eraan kan bijdragen dat lerarenopleiders, onderwijsontwikkelaars en leraren beter gefundeerde keuzes kunnen maken ten aanzien van het ontwerpen van taaltaken en de selectie van focus on form strategieën.

CURRICULUM VITAE

Marrit van de Guchte (1976) studied European Studies from 1994 to 1999 at the University of Amsterdam where she graduated on a thesis that dealt with cross-border healthcare in Germany. In this period she also studied German Language and Literature at the 'Freie Universität Berlin', Spanish at the 'Universidad Complutense de Madrid' and European and International Law. In 2000 she participated in the Erasmus program 'European Voluntary Service' for which she supervised German language courses for elementary students in Burgundy, France. Subsequently (2001) Marrit attended teacher training at the Graduate School of Teaching and Learning at University of Amsterdam (ILO) where she received her MA degree in teaching German Language and Literature in 2002. From 2001 to 2009 Marrit taught German at the Bonhoeffer College in Castricum and coached students with performance anxiety. Besides her job as a teacher of German she started working as a teacher educator at the Graduate School of Child Development and Education at the University of Amsterdam in 2007. In August 2011 she started her PhD research on focus on form in task-based language teaching at the Research Institute of Child Development and Education at the University of Amsterdam, under supervision of Gert Rijlaarsdam, Martine Braaksma, and Peter Bimmel. She presented the results of her PhD project at several conferences including EARLI, EUROSLA and TBLT. Furthermore Marrit is head of the advisory board for German as a foreign language at the Duitsland Instituut Amsterdam (DIA). Moreover, she is the editor for German for *Levende Talen Tijdschrift* (quarterly magazine for modern language teachers). Currently she designs task-based German language curricula and teaches task-based German courses at the A. Roland Holst College in Hilversum and works as a lecturer of German language education at the Graduate School of Teaching and Learning at the University of Amsterdam.