Language and Education

ISSN: 0950-0782 (Print) 1747-7581 (Online) Journal homepage: https://www.tandfonline.com/loi/rlae20

How secondary school students may benefit from linguistic metaconcepts to reason about L1 grammatical problems

Jimmy H.M. van Rijt, Astrid Wijnands & Peter-Arno J.M. Coppen

To cite this article: Jimmy H.M. van Rijt, Astrid Wijnands & Peter-Arno J.M. Coppen (2020) How secondary school students may benefit from linguistic metaconcepts to reason about L1 grammatical problems, Language and Education, 34:3, 231-248, DOI: 10.1080/09500782.2019.1690503

To link to this article: https://doi.org/10.1080/09500782.2019.1690503

Published online: 28 Nov 2019.

LANGUAGE AND EDUCATION
2020, VOL. 34, NO. 3, 231–248
https://doi.org/10.1080/09500782.2019.1690503

Routledge
Taylor & Francis Group

How secondary school students may benefit from linguistic metaconcepts to reason about L1 grammatical problems

Jimmy H.M. van Rijt[a,b], Astrid Wijnands[c,d] and Peter-Arno J.M. Coppen[d]

[a]Department of Teacher Education, Fontys University of Applied Sciences, Sittard, The Netherlands; [b]Department of Linguistics, Center for Language Studies, Radboud University, Nijmegen, The Netherlands; [c]Department of Teacher Education, Utrecht University of Applied Sciences, Utrecht, The Netherlands; [d]Radboud Teachers' Academy, Radboud University, Nijmegen, The Netherlands

ABSTRACT

In L1 grammar teaching, teachers often struggle with the students' conceptual understanding of the subject matter. Frequently, students do not acquire an in-depth understanding of grammar, and they seem generally incapable of reasoning about grammatical problems. Some scholars have argued that an in-depth understanding of grammar requires making connections between concepts from traditional grammar and underlying metaconcepts from linguistic theory. In the current study, we evaluate an intervention aiming to do this, following up on a previous study that found a significant effect for such an approach in university students of Dutch Language and Literature ($d = 0.62$). In the current study, 119 Dutch secondary school students' grammatical reasonings ($N = 684$) were evaluated by language teachers, teacher educators and linguists pre and post intervention using comparative judgement. Results indicate that the intervention significantly boosted the students' ability to reason grammatically ($d = 0.46$), and that many students can reason based on linguistic metaconcepts. The study also shows that reasoning based on explicit underlying linguistic metaconcepts and on explicit concepts from traditional grammar is more favored by teachers and (educational) linguists than reasoning without explicit (meta)concepts. However, some students show signs of incomplete acquisition of the metaconcepts. The paper discusses explanations for this incomplete acquisition.

ARTICLE HISTORY
Received 13 June 2019
Accepted 3 November 2019

KEYWORDS
L1 grammar teaching; linguistic metaconcepts; linguistic reasoning; comparative judgement; secondary education

Introduction

In recent years, explicit grammar teaching in L1 contexts has repositioned itself on the pedagogical agenda (Chen and Myhill 2016; Locke 2010). In spite of this renewed educational interest in grammar education, there is a limited amount of empirical research into the topic (Boivin et al. 2018). In most cases, if there is research into grammar teaching, then it focuses on the relationship between explicit grammar teaching and writing development

CONTACT Jimmy H.M. van Rijt j.vanrijt@fontys.nl

(e.g., Myhill et al. 2012; Camps and Fontich 2019). However, important this relationship may be, much less research is being conducted into grammatical learning and instruction in itself. This can be considered problematic, because in many cases, grammar education has a traditional image (cf. Lefstein 2009; Van Gelderen 2010; Watson 2015), in which sentences are analyzed out of context, using mostly rules of thumb (Berry 2015) and an outdated body of conceptual knowledge (Hulshof 1985; Van Rijt and Coppen 2017; Van Rijt, De Swart and Coppen 2019), addressing mostly lower order thinking skills (Van Rijt, Wijnands and Coppen 2019). Such grammar teaching arguably does not enhance the students' language awareness, and it fails to provide students with an instrumentarium with which they can adequately analyze language (see Myhill 2000, 2003). One of the greatest challenges in grammar teaching therefore deals with the question how teachers can make sure their students acquire an in-depth understanding of the grammatical subject matter.

Many researchers and teachers have attempted to tackle this issue by implementing Systemic Functional Linguistics (SFL) into the classroom, drawing on the work of the linguist Michael Halliday (cf. Halliday and Matthiessen 2004; McCabe 2017), in which grammar is presented as a (rhetorical) choice, rather than as a strict body of rules that has to be obeyed by the learner (Myhill, Jones and Lines 2018). However, in spite of educational ideology being tilted towards teaching SFL, in practice, most grammar teaching is of a far more traditional nature, even in SFL strongholds, such as Australia (Jones and Chen 2012). In a recent survey, Macken-Horarik, Love and Horarik (2018) discovered that Australian teachers possess far less knowledge on SFL and its practical application than they should according to the official curriulum, which leads to more traditional grammar teaching. Such traditional forms of grammar teaching can be found throughout the world, even in the face of alternatives, such as SFL. The systematic literature review of Van Rijt, De Swart and Coppen (2019) concludes that the field of L1 grammar teaching can be characterized as 'traditional-Hallidayan'—'traditional' being mostly applicable to the classroom practice of grammar teaching (cf. Lefstein 2009; Watson 2015), whereas Halliday dominates educational ideology (Fearn and Farnan 2007; Myhill 2018). Given this mismatch and the difficulties of implementing SFL in traditional grammar education, it might be beneficial for grammar education if approaches to grammar are considered that do not involve making a choice between either teaching traditional grammar or adopting SFL**1**. Instead, there are proposals in which grammar teaching (of whatever signature) can be enriched in other ways.

One such possible proposal is to restore the bond between linguistics and L1 grammar teaching (e.g., Hudson 2010; Hudson and Walmsley 2005; Giovanelli 2016), for example by enriching traditional grammar with (meta)concepts from modern linguistic theory (Hulshof 2013; Van Rijt and Coppen 2017; Van Rijt, De Swart and Coppen 2019).

Even though modern linguistics has evolved from traditional grammar, with Saussurian Structuralism as an important step in this evolution (Seuren 1998), there is still a large gap: modern linguistics has generated large amounts of new insights that could greatly enrich traditional grammar teaching in schools. The question which of these conceptual insights may be useful for enriching grammar teaching has been explored before by Van Rijt and Coppen (2017). In Van Rijt and Coppen (2017), a general agreement on 26 metaconcepts that linguistic experts deemed of great importance for both linguistic theory and L1 grammar teaching was reached, drawing on a very broad spectrum of linguistic theories such as structuralism, generativism, cognitivism, construction grammar and SFL. Following Lipman (2003, 181) and Van Rijt et al. (2019), metaconcepts can be defined as 'the mind's generalized

representation of one or more concepts'. In other words: 'metaconcepts are concepts with an overarching value, used for conceptualization or rational deduction and mediating the students' understanding of secondary subject-specific concepts' (Van Rijt et al. 2019). For example, the metaconcept of *valency* (cf. Perini 2015) could be used to enhance the students' understanding of the difference between objects and adverbials (secondary concepts), and could help students understand why some sentences contain (in)direct objects, whereas others do not. In traditional grammar teaching, teachers mostly lack the means to adequately explain such things, because relevant underlying metaconcepts are not incorporated into traditional grammar. Hence, restoring the bond between linguistic theory and grammar teaching could provide teachers with opportunities for enrichment, making principled understanding much more likely than in traditional grammar teaching. Moreover, there are good indications that an approach to grammar learning and instruction in which active connections are made between an underlying metaconcept (e.g., 'valency') and related secondary concepts (e.g., 'object') have a positive impact on grammatical reasoning quality.

In a previous study, university students of Dutch Language and Literature showed a significant increase in their ability to tackle grammatical problems after they followed an intervention that fostered such connections (see Van Rijt, De Swart, Wijnands and Coppen 2019), with a very reasonable effect size (cohen's $d = 0.62$). Additionally, a multiple regression analysis revealed that referring to explicit linguistic metaconcepts, such as *valency, predication, modification,* or *complementation,* was a significant predictor for grammatical reasoning ability, especially in combination with using explicit secondary concepts from traditional grammar. Metaconcepts appear to have the ability to mediate the students' understanding of secondary concepts—a finding that aligns with the function of metaconcepts in historical reasoning in history education (Van Drie and Van Boxtel 2008). These findings show that the use of linguistic concepts and metaconcepts is a crucial factor in metalinguistic activity, a subject that is more and more researched in recent decades (cf. Andrews 1997; Camps and Milian 1999; Myhill and Jones 2015).

However, in spite of the theoretical benefits of such an approach and some empirical evidence in favor of a metaconceptual approach in university students, no research to date has investigated to what extent a metaconceptual approach might improve secondary school students' grammatical reasoning ability. Is it possible that students from secondary education can benefit from a metaconceptual approach, and if so, which design principles might be beneficial for fostering their linguistic metaconcept use? To answer these questions, the current study discusses findings from a metaconceptual intervention in lower secondary education that was underpinned by theoretical design principles. The study is part of a larger design-based research cycle (cf. Plomp and Nieveen 2007), and it aims to identify relevant design principles that can be used to inform interventions aimed at enhancing the students' grammatical reasoning ability.

Theoretical design principles for grammatical understanding

When reviewing the educational literature on (grammar) teaching, a handful of useful design principles can be discerned that could enhance students' grammatical reasoning. In total, we will discuss five promising design principles that underpinned our intervention.

The first design principle states that linguistic metaconcepts should be actively related to the secondary concepts associated with them. In such an approach, teachers would first

aim to generate a basic understanding of a linguistic metaconcept, before refining the students' understanding of the metaconcept with secondary concepts from traditional grammar. For example, it would be easier for students to first learn about the concept of *valency* (i.e., the idea that the main verb serves out grammatical and semantic roles), and later about the concepts from traditional grammar that constitute these roles (e.g., subject, direct, and indirect objects). In traditional school grammar, the underlying metaconcept of valency is mostly ignored or left out (Van Rijt, Wijnands, and Coppen 2019).

This first design principle follows from the theoretical position defended in Van Rijt and Coppen (2017), Van Rijt, De Swart, and Coppen (2019) and Hulshof (2013), and from an exploratory empirical study (Van Rijt et al. 2019). Hence, explicit metalinguistic knowledge was promoted in the intervention. According to Watson and Newman (2017), students struggle with articulating metasyntactic choices in writing, although this struggle can be diminished when explicit syntax is addressed (Gombert 1992; Van Rijt et al. 2019). Given the difficulties students experience when reflecting on syntactic choices, it can be expected that they will experience similar difficulties in reasoning about grammatical problems. The intervention aims to provide students with more adequate means to talk and think about syntax.

The second design principle we adhered to, states that grammar education should make use of the students' own intuitions about language. In traditional grammar teaching, labeling parts of speech is the central activity, but linguistic analysis is rarely linked to the students' intuitions (Coppen 2009). Establishing such links could be of vital importance, since every individual has his or her own intuitions about language, and these intuitions can differ among people (De Hoop 2016; Van Rijt, Wijnands and Coppen 2019; Wijnands 2017). Traditional grammar teaching immediately zooms in on cognitive activities, neglecting the stage of 'experiencing' altogether (Coppen 2012). For an adequate cognitive understanding, fostering such experiences is essential (Moseley et al. 2005). In the current intervention, most assignments were therefore designed to elicit an experience in students, triggering their language intuitions, before moving forward to cognitive (Moseley et al. 2005), reflective (King and Kitchener 2004) or creative thinking (Ritter 2012). An effective way to make students more aware of their own language intuitions, is to employ (guided) inductive assignments (cf. Prince and Felder 2006), that are aimed at making the students discover an underlying grammatical pattern (cf. Haight, Herron and Cole 2007).

A third design principle relates to dealing with uncertainties in grammar. Analyzing language can be labeled an ill-structured problem (King and Kitchener 2004), for which clear-cut answers hardly exist (Van Rijt, Wijnands and Coppen 2019; Wijnands 2017; Kuiper and Nokes 2014). It is generally accepted in the literature that dealing with ill-structured problems requires critical and reflective thinking (King and Kitchener 2004; Moseley et al. 2005). In the case of grammar learning, this means that students would need the ability to argue in favor of or against a particular grammatical analysis, deliberating which possible solution to deal with a grammatical problem is the best. For example, consider the sentence *He walked one round*. Simple as this sentence might seem, a fundamental question arises while analyzing it: what is the grammatical function of 'one round'? Upon closer inspection, the constituent has characteristics of both a direct object and an adverbial, but the answer to this and related questions are not clear-cut, even among linguists. In fact, contrary to what traditional grammar suggests, most sentences in real-life cannot be analyzed unambiguously (Coppen 2009), partly because conceptual categories themselves are sometimes

fuzzy (Kuiper and Nokes 2014). Hence, students that are faced with such issues should learn to deliberate which of these options is best, which requires critical thinking. According to a recent survey among 110 Dutch Language teachers, most teachers feel that in grammar teaching, students should learn to deal with such uncertainties more (Van Rijt, Wijnands, and Coppen 2019). The assignments in the intervention in this study were designed in such a manner that they exposed students to a limited degree of uncertainty. Such uncertainties can hardly be dealt with in traditional grammar teaching with its overt focus on identifying one correct answer. Rather, employing metaconcepts into grammatical thinking presents the learner with the ability to tackle a grammatical problem more adequately (Van Rijt et al. 2019). In the case given, the valency pattern of the verb *to walk* (which is normally monovalent, requiring only a subject) could be a good indication that we may not be dealing with a direct object. Such an observation could be a good starting point for reasoning about grammatical problems.

A fourth design principle connects to the previous one. Since dealing with uncertainties requires students to take multiple ways to deal with a grammatical problem into account, stimulating discussion could facilitate such multiperspectivity. Indeed, several scholars working on L1 grammar teaching have argued that verbalization supports metalinguistic thinking (e.g., Fontich and Camps 2014; Myhill and Jones 2015; Ribas, Fontich, and Guasch 2014). In particular, embedding students into a sociocultural approach and stimulating them to engage in exploratory talk (cf. Mercer 2004, 2013), could have positive benefits for their grammatical reasoning. According to Mercer, exploratory talk can be considered the most powerful form of group discussion, since in such talks, 'knowledge is made more publicly accountable, and reasoning is made more visible' (Mercer 2004, 146). Exploratory talk is characterized by participants critically taking multiple perspectives into account, in which statements and suggestions are offered for joint consideration.

In other research fields, such as history education, exploratory talk has proven a very successful way to enhance critical thinking on historical subjects (Havekes 2015). Moreover, in L2 grammar teaching, sociocultural approaches are frequently employed successfully (Lantolf, Thorne, and Poehner 2015). To stimulate students to engage in exploratory talk, the teachers in the intervention were encouraged to employ ground rules of good exploratory talk before every discussion activity (Mercer 2013, 13). In the intervention, students were first asked to discuss a problem in pairs, before tackling other problems in groups of four. This was done to gradually introduce students to exploratory talk.

The fifth design principle, finally, deals with the role of the teacher in effectively scaffolding the students' reasoning and discussion on grammatical problems. It is known from history education that complex reasoning does not emerge on its own (Havekes 2015), but that it has to be carefully guided by a teacher. Such guidance is only effective if it extends beyond simply encouraging the students to search for more possible answers. Students are commonly known to believe that whenever multiple answers to a question are possible, any one of them is valid as long as there is some sort of argument to support it (Havekes 2015, 88). This is however not the case, since good domain specific reasoning in ill-structured knowledge domains such as grammar is strongly context depended, requiring active deliberation (Moore 2004). Therefore, teachers must also 'guide the use of specialized language' (i.e., secondary grammatical concepts and linguistic metaconcepts) and 'discuss the criteria with which to judge the given answers' (Havekes 2015, 88). Moreover, the quality of any reasoning is strongly determined by its *coherence*, i.e., the way relationships between

different concepts (especially between concepts and metaconcepts) are expressed (Havekes 2015, p.69). To facilitate this, teachers were asked to make the students reflect on both the (meta)conceptual content and the processes of cooperation and discussion. A description of the intervention is given in Appendix 1.

The current research is the first to explore the potential of these design principles by evaluating an intervention that was implemented in Dutch secondary schools. In particular, we aimed to answer the question: 'To what extent do secondary school students reason better about unknown grammatical problems after a metaconceptual intervention, and how does this relate to their use of linguistic metaconcepts?'

Method

The intervention

The intervention consisted of four lessons of 50 minutes each, the normal time for a Dutch secondary school lesson. It was enrolled in five different secondary schools in 3 vwo (the highest level of lower pre-university education).

The intervention focused on four somewhat related metaconcepts: *predication, valency, complementation* and *modification* (see Table 1 for explanations of these metaconcepts), and it repeated the students' knowledge of the related secondary concepts from traditional grammar (e.g., subject, object, and adverbial). See Table 1 for brief explanations of the metaconcepts from the intervention.

In each of the four lessons, one of the metaconcepts played a central role, although valency and predication were considered to be the central metaconcepts, receiving attention in every lesson. Each teacher received assignment booklets for their students, and additional teaching materials, such as plasticized cards that were used for the inductive assignments. In addition, an elaborate the teachers' manual was written, in which teachers received detailed instruction on how to handle each lesson according to the design principles. Prior to the intervention, teachers were extensively briefed on the goals and materials of the intervention. Specific attention was paid to the design principles that underpin the intervention (notably that students had to engage in exploratory talk), and teachers were told that they could deviate from the teachers' manual if the occasion called for it, so long as they acted in the spirit of the design principles. They were also encouraged to register any changes they had made to the intervention, and for what reason, by returning fidelity measures. To ensure that the participating teachers had sufficient knowledge on the linguistic metaconcepts from the intervention, brief background articles on these concepts were provided, and if necessary, detailed briefing was given to them.

Teachers gave one lesson each week, finishing the intervention in four weeks. The intervention was enrolled between November and December 2018.

Table 1. Metaconcepts targeted in the intervention and their descriptions (based on Van Rijt and Coppen 2017, 379 and 380).

Metaconcept	Explanation
Predication	Elements can be linked to *do* or *to be* meaning
Valency	The verb chooses a number of arguments
complementation	Some parts of speech are very closely related to the verb
modification	Some parts of speech are very lossly related to the verb

Participants

Five teachers of Dutch Language and Literature from different secondary schools in the Netherlands participated in the study (1 male). Their teaching experience in Dutch Language and Literature varied, ranging from 3 to >25 years. Teachers also differed in terms of their qualifications. Three of the teachers held a grade one certification (masters' degree), whereas two others held a grade 2 certification (bachelor's degree).

Teachers volunteered for participation in the research. Their students signed a consent form, in which it was stated that their data would be used anonymously for scientific research. No students withheld their consent. In addition, the schools and parents were notified. In total, 119 secondary school students (mean age = 14.0 years, $SD = 0.46$ years) participated in the study, 56 of whom were male and 63 female. In the pretest, data from 116 students were collected. In the posttest, with drop-outs due to illness, data from 112 students were collected.

Pre- and posttest

In the pre- and posttest, students were confronted with three grammatical problems they individually had to tackle. These problems were of a type they had never encountered before, nor were they in any way part of the intervention. Students were for example asked to explain the ambiguity of sentences like 'He saw the man with binoculars', and explain how this ambiguity could be accounted for by investigating the grammatical structure, or they had to explain why one sentence of a pair could be considered grammatical, whereas another one could not (e.g., Jan regent nat ('Jan is raining wet') versus Jan regent'* ('*Jan rains'). The students' responses to these problems were considered as 'reasonings' for the analysis. Students were encouraged to tackle the grammatical problems as elaborately as possible. Two of the items were designed in such a way that they could easily elicit the metaconcepts that were covered in the intervention, although they could also be tackled using concepts from traditional grammar. These were the target items. One of the items was a filler item, and it related to a grammatical problem the students did not receive any education in ('binding' and anaphoric reference). See Appendix 2 for a short overview of the grammatical problems that students had to tackle.

Filler items were meant to measure whether any difference in reasoning ability could not simply be attributed to a general increase in their cognitive development, or to a testing effect. To nullify any effects related to the difficulty or order of the tasks, the groups were randomly divided into two and the pre- and posttest items were counterbalanced (Shadish, Cook, and Campbell 2002, 109). Students were given 10 minutes to tackle the set of grammatical problems, using a pen-and-paper task. Although exploratory talk was an important design principle in the intervention, in the data collection of the current study students was not actually talking to each other. However, the individual writing assignment in which they had to reason about an unfamiliar linguistic problem is thought to share some important characteristics with the exploratory talk the students were engaged in during the intervention (such as the need to speculate and the possibly 'messy' character of the reasoning), so that it can be considered a form of 'exploratory writing' (cf. Mercer (2003, 16)). In any case, the writing assignment is meant to assess the quality of individual reasoning students achieved as a result of the intervention. Most students easily managed to complete the task

within the given time. They were encouraged to make a real effort, because it was explained that the researchers would use the data to help their teachers become better at teaching grammar. In total, 684 student reasonings were collected this way; 35 questions were left blank because students did not manage to come up with a solution to the grammatical problem (attrition rate: 5.1%).

Rating the quality of the students' grammatical reasoning

The quality of the individual student reasonings was rated using an online platform for comparative judgement, D-PAC (http://www.d-pac.be). In comparative judgement, raters repeatedly compare two grammatical reasonings (with pre- and posttest reasonings mixed), judging which of the two is the best. Performances are compared multiple times to various other reasonings by multiple raters. This results in a scale, ranking all of the reasonings from worst to best (Lesterhuis et al. 2016). Comparative judgment has been proven a much more effective way of rating performances than other methods, such as ranking on a Likert scale or using a rubric (Sadler 2009). This means that comparative judgment is a more valid way of assessing. This method also eliminates complications resulting from sequential effects and differences in the severity of raters (Pollitt 2012).

The grammatical reasonings were evaluated by 16 experienced raters, consisting of a mixture of linguists, language teacher educators and secondary school language teachers. On average, reasonings were compared 21.1 times, resulting in a strong reliability of .83 (see Verhavert et al. 2018). The raters were unaware that an intervention had taken place, nor were they given other information related to the aim or design of the study.

Analyzing the students' grammatical reasoning

When analyzing the grammatical reasoning of students, we noticed that two aspects of the use of grammatical concepts seemed of vital importance. First, their explicit use of secondary concepts and metaconcepts, and second, the conceptual coherence that was expressed in the reasoning of a student, operationalized as the quality of the relationship between concepts. No relation between concepts was considered 'incoherent', and a correctly articulated relationship was considered 'coherent'. In history education, such a characterization has proven effective in analyzing (meta)concept use in historical reasoning (Havekes 2015). Therefore, we adopted a similar method here. Schematically, this leads to four analyzing categories, consisting of two axes: *linguistic metaconcepts versus no linguistic metaconcepts*, and *no coherence* versus *coherence*. See Table 2 for an overview.

The first type of answer provided by students was labeled *no grammatical concepts*. In this category, students used no metalinguistic terminology whatsoever, nor was there any coherence in the students' response. In most cases, this simply meant that a student would try to explain in his own words what the meaning of an utterance from the question was. A typical response for this category related to grammatical problem 3 from Appendix 2 (translations from Dutch by the authors):

> Because sentence a does not contain a location, so you don't know why he blew the paper. In sentence b it does say how he did it. You won't just blow against a paper for no reason. (Reasoning #2)

Table 2. Characterization of the students' grammatical reasoning.

	No linguistic metaconcepts		
No coherence	**No grammatical concepts (1)** The answer consists of no explicit grammatical concepts, and the answer shows limited coherence	**Traditional concepts related to each other (2)** The answer contains concepts from traditional grammar, which are meaningfully related to one another	**Coherence**
	Blind linguistic metaconcepts (3) The answer contains metaconcepts, but these serve no clear function	**Linguistic metaconcepts related to traditional concepts (4)** The answer consists of metaconcepts and traditional concepts, which are meaningfully intertwined	
	Linguistic metaconcepts		

A typical example for the second category, labeled *traditional concepts related to each other*, is Reasoning #34. In such a reasoning, students use traditional grammatical terminology to indicate how the problem could be tackled, relating these concepts to the grammatical problem and to one another:

> The direct object could be 'the man'. Then you have seen a man by means of binoculars ('with binoculars' is an adverbial in that case). The direct object could also be 'the man with binoculars'. In that case, you have seen a man who has binoculars. (Reasoning #34)

The third category, labeled *blind linguistic metaconcepts*, is characterized by reasonings in which linguistic metaconcepts are used by students, but without any clear function of the metaconcept within the reasoning. In most cases, it would basically amount to a case of name dropping in the hopes of saying something that their teacher would want to hear. Therefore, such metaconcept use was dubbed 'blind' metaconcept use. In such cases, the students' metaconcept use indicates an incomplete acquisition or understanding of the metaconcept. A prime example is reasoning #408, related to grammatical problem 6 from Appendix 2, in which the addition of the concept of predication has no explanatory value:

> In sentence (1), the father of Jan is photographing, so it could be anyone he photographs. In sentence (2) Jan is photographing, so that could also be anyone. Predication. (Reasoning #408)

This leaves us with the fourth category, *linguistic metaconcepts related to traditional concepts*. In this category, students make active connections between metaconcepts and the related concepts from traditional grammar. A typical example is constituted by reasoning #572, related to problem 3 from Appendix 2:

> The valency of 'to blow' is one. SOMEONE blows SOMETHING is impossible, except when an adverbial is present, as is the case in sentence (b) ('Jan blows the paper off the table').

In this characterization of the students' reasoning, categories 2 and 4 are the most preferred, because they are most indicative of an in-depth understanding of the grammatical subject matter.

All reasonings were scored by the principal investigator, and in any cases of doubt the third author of the present paper was consulted. Any differences in opinion were resolved through discussion. A small random sample of reasonings from the posttest (10%) was rated independently by the first and third author of this paper. Cohen's Kappa revealed substantial initial interrater agreement ($\kappa = .80$, $p < .001$).

Statistical analysis of grammatical reasonings

In order to ascertain whether students' reasoning improved after the intervention, various T-tests were conducted. Additionally, an ANCOVA analysis examined whether there were quality differences associated with the four categories from the previous section.

Results

Descriptives

On average, students wrote 26.54 words per reasoning ($SD = 15.13$). In the pretest, students wrote an average of 25.45 words ($SD = 13.87$); in the posttest, the number of words averaged at 27.67 ($SD = 16.32$). There was no significant difference in the number of words written between pre- and posttest ($t(682) = 1.91$, $p = .056$).

Table 3 shows the distribution of the four reasoning categories over both measurements.

As can be inferred from Table 3, categories 3 and 4 do not occur in the pretest. The table reveals that the Category 1 reasonings have decreased in the posttest, as well as the Category 2 reasonings.

Analysis of grammatical reasonings

Table 4 summarizes the differences between pre- and posttest for reasoning quality. Scores in comparative judgement are centered around 0, with 0 as the average reasoning score at both tests.

Students managed to improve on the target items. 105 students participated in both measurements. On average, they scored −.40 in the pretest, and .32 in the posttest. A paired samples T-test revealed that their improvement was significant: $t(104) = -4.36$, $p = < 0.01$. On the target items, no significant change could be detected. When the analysis focuses on the differences in the students' reasoning, an independent samples T-test revealed that the students' reasoning improved significantly on the target items, with an effect size approaching the medium threshold ($t(430) = 4.75$, $p = <.001$, $d = 0.46$). The students' reasoning remained constant on the filler items ($t(215) = 0.47$, $p = .64$).

Table 3. Distribution of reasoning categories.

	Pretest		Posttest		
Category	*N*	% of pretest	*N*	% of posttest	Total
1	272	83.2	198	61.5	470
2	55	16.8	32	9.9	87
3	0	0.0	37	11.5	37
4	0	0.0	55	17.0	55
Total	327	100	322	100	649

Table 4. Reasoning differences between pre- and posttest.

	Pretest				Posttest			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Reasoning score	−0.11	1.51	−5.58	4.97	0.29	0.74	−5.51	5.01
Target items	−0.33	1.52	−5.58	4.97	0.32	1.29	−5.51	5.01
Filler items	0.33	1.40	−2.12	4.93	0.24	1.42	−3.71	3.22

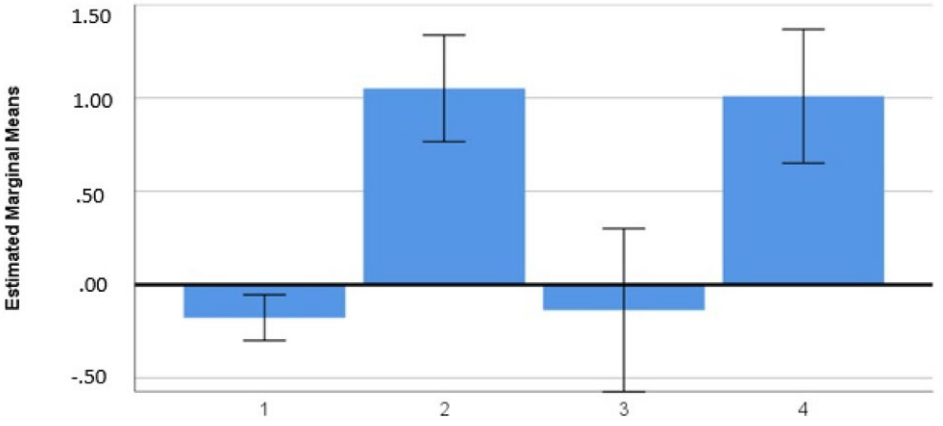

Figure 1. Quality differences associated with reasoning categories. Error bars 95% CI.

An ANCOVA analysis controlling for the effect of the teacher, with *reasoning quality* as the dependent variable and with *reasoning category* (1–4) as a fixed factor, revealed significant differences in quality between the four reasoning categories: $F(3,644) = 29.53$, $p = <.001$). In Figure 1, these differences are depicted.

Additional contrast analyses revealed significant differences in quality between categories 1 and 2 ($p = <.001$), 2 and 3 ($p = <.001$), 3 and 4 ($p = <.001$) and 1 and 4 ($p = <.001$). Categories 2 and 4 ($p = .86$) and 1 and 3 ($p = .86$) did not significantly differ in quality.

Discussion

The current study set out to evaluate an intervention in L1 grammar teaching aiming to increase the students' grammatical reasoning ability by relating concepts from traditional grammar to underlying metaconcepts from linguistic theory. To this end, differences in the quality of their reasoning prior to and after an intervention were examined.

From the data, it can be inferred that the intervention successfully managed to improve the quality of the students' reasoning on the target items, whereas the quality of the filler items remained constant. This suggests that the improvement is not simply due to a testing effect.

We suggest that, generally, the intervention seems to stimulate the students' ability to reason grammatically. The significant increase can arguably be attributed to two changes in students' reasoning in the posttest. First, the relative decrease of Category 1 reasonings, in which no grammatical concepts are used and in which no coherence is present. This finding aligns with those of Watson and Newman (2017), who noticed that students find it easier to talk about semantics (meaning) than about syntax. The findings also echo Myhill (2005), who found that the teachers' lexical subject knowledge is much stronger than their syntactic knowledge. In line with previous research, students seem to shift toward more syntactic reasoning after having received explicit instruction (Gombert 1992; Van Rijt et al. 2019).

A second change in the students' reasoning is their increased ability to reason in a way in which underlying metaconcepts are related to concepts from traditional grammar (Category 4 reasonings). This finding resonates with those from our previous study

(Van Rijt et al. 2019), in which university students' reasoning significantly increased due to their ability to relate metaconcepts to concepts from traditional grammar in their grammatical reasoning.

However, a few caveats are in order when interpreting the results. First, after the intervention, some students showed an incomplete acquisition or understanding of the metaconcepts from the intervention, given the number of reasonings from Category 3, in which metaconcepts were used 'blindly'. The question arises what could have caused blind concept use. The most likely explanation for this is that the intervention was too short (four lessons), and too congested to achieve its full potential. Some evidence for this was found in the fidelity measures. Teachers reported they had often skipped reflecting on the processes of effective exploratory talk—a finding that was later confirmed in a focus group interview. Therefore, the sociocultural nature of the intervention may have been somewhat compromised. In particular, leaving out the reflection in classroom discussion could have had a negative impact on the effect of the intervention (Van Drie and Van Boxtel 2011), since this stage is crucial for teachers to discover and deal with any cases of 'blind' (meta)concept use.

From the ANCOVA analysis, it could be deduced that reasoning with 'blind' concepts (Category 3) is on par with reasoning from Category 1 in terms of quality. Neither of these categories is preferred in education, since they indicate a limited understanding of the grammatical subject matter. Interestingly, for Category 3 reasonings, the standard deviation was much higher than for Category 1 reasonings (cf. Figure 1). This may be explained by the fact that some raters may have appreciated students using metaconcepts, even if the metaconcepts did not fulfill a clear function within the argument, whereas others could not appreciate such 'name dropping'. Notwithstanding these differences between raters, there appears to be a general agreement among them that metaconcept use is not favored per se; it is only truly valued if such metaconcepts have a clear and relevant function within the students' reasoning. The results from Figure 1 are also a good indicator that the categorization of the students' reasoning into four categories is an adequate way to analyze such data.

The above-mentioned factors are likely to have had an impact on the effect of the intervention, that in spite of the reported caveats revealed an effect size that was close to being moderate. This effect may also have been larger if all of the interventions' metaconcepts had received an equal amount of attention. *Modification* and *complementation* were less well represented, and as a consequence, students struggled with those metaconcepts more.

Research limitations and future prospects

The current study is the first to explore relevant design principles for interventions in which students learn to actively relate concepts from traditional grammar to underlying linguistic metaconcepts. Given the study's exploratory nature, however, it is not without limitations. First, other than the teachers' views on how the students performed in class and their self-reported fidelity measures, there is no observational data of how teachers actually implemented the intervention, nor is there any evidence about the students' reasoning in the classroom settings. This means that this study focuses on the assessment of the result of the intervention rather than on investigating the intervention itself. However, since a clear improvement of reasoning quality was noted, the study can be regarded as a proof of principle for the effectivity of the design principles in general, although it remains unclear

exactly which of the principles are most effective. More research into the (implementation of) these design principles is crucial in gaining a deeper understanding of their effectivity.

Second, the intervention was tested with teachers who had volunteered to try it out into their classrooms. Therefore, the positive results from this paper cannot be completely generalized to a broader teacher audience. More research is needed on the implementation of such an intervention in classrooms with teachers who hold different views on grammar teaching. Given teachers' limited subject knowledge on grammar (cf. Chen and Myhill 2016; Sangster, Anderson, and O'Hara 2013; Alderson and Hudson 2013), special attention might be given to investigating the metalinguistic knowledge that teachers should possess in order to effectively take advantage of the benefits that a metaconceptual approach can offer. Moreover, teachers initial beliefs and their existing practice can be a real hindrance when they are trying to teach grammar this way (Van Rijt, Wijnands and Coppen 2019). In the present study, highly motivated teachers participated, and even they had expressed that they experienced difficulties in implementing the intervention that were caused by their normal classroom practice. Even though the teachers stated that they felt it was important for students to deal with grammatical uncertainties, they felt somewhat uncomfortable in dealing with uncertainties themselves, given remarks they made on their minutely preparation. Future research would therefore do well to also take teachers (initial) beliefs into account, and to address their own stance towards dealing with uncertainties. Finally, follow-up research should involve a stronger methodological design (e.g., a switching replications design, cf. Shadish, Cook, and Campbell 2002) to allow for more rebost conclusions.

In spite of its limitations, the present study has provided positive evidence in favor of a metaconceptual approach to grammar in L1 classrooms, showing that the students' grammatical reasoning can significantly improve when linguistic metaconcepts are meaningfully incorporated into grammar education.

Note

1. In what follows, we will focus on enriching traditional grammar teaching, since not only international grammar teaching, but also in particular the Dutch context, seems to exhibit strong characteristics of traditional grammar teaching (cf. Van Gelderen 2010; Van Rijt and Coppen 2017).

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Netherlands Organisation for Scientific Research (NWO) under grant number 023.009.034.

References

Alderson, J., and R. Hudson. 2013. "The Metalinguistic Knowledge of Undergraduate Students of English Language or Linguistics." *Language Awareness* 22 (4): 320–337.

Andrews, S. 1997. "Metalinguistic awareness and teacher explanation." *Language Awareness* 6 (2–3): 147–161.
Berry, R. 2015. "Grammar Myths." *Language Awareness* 24 (1): 15–37. doi:10.1080/09658416.2013.873803.
Boivin, M.-C., X. Fontich, R. Funke, M.-J. García-Folgado, and D. Myhill. 2018. "Working on Grammar at School in L1 Education: Empirical Research across Linguistic Regions." *L1 Educational Studies in Language and Literature* 1–6. doi:10.17239/L1ESLL-2018.18.04.01.
Camps, A. & Milian, M. (Eds). 1999. *Metalinguistic Activity in Learning to Write*. Amsterdam: Amsterdam University Press.
Camps, A., and X. Fontich. 2019. "Teachers' Concepts on the Teaching of Grammar in Relation to the Teaching of Writing in Spain: A Case Study." *L1 Educational Studies in Language and Literature* 19: 1–36. doi:10.17239/L1ESLL-2019.19.02.02.
Chen, H., and D. Myhill. 2016. "Children Talking about Writing: Investigating Metalinguistic Understanding." *Linguistics and Education* 35: 100–108. doi:10.1016/j.linged.2016.07.004.
Coppen, P.-A. 2009. *Leren Tasten in Het Duister. [Learning to Grope in the Dark]* Inaugural Address. Nijmegen: Radboud University.
Coppen, P. A. 2012. "Sturen zonder handen." In *Zesentwintigste conferentie Het Schoolvak Nederlands* edited by S. Vanhooren and A. Mottart, 207–213. Gent: Academia Press.
De Hoop, H. 2016. "Woordvolgordevariatie: theorie versus Empirie? [Word Order Variation: theory versus Empiricism?]." *Nederlandse Taalkunde* 21 (2): 275–284. doi:10.5117/NEDTAA2016.2.HOOP.
Fearn, L., and N. Farnan. 2007. "When Is a Verb Using Functional Grammar to Teach Writing." *Journal of Basic Writing* 26 (1): 1–26.
Fontich, X., and A. Camps. 2014. "Towards a Rationale for Research into Grammar Teaching at Schools." *Research Papers in Education* 29 (5): 598–625. doi:10.1080/02671522.2013.813579.
Giovanelli, M. 2016. "Developing Beginning Teachers' Linguistic Awareness: Issues and Practice in Initial Teacher Education." In *Knowing about Language. Linguistics and the Secondary English Classroom*, edited by M. Giovanelli and D. Clayton, 186–197. London: Routledge.
Gombert, E. 1992. *Metalinguistic Development*. Hemel Hempstead: Harvester Wheatsheaf.
Haight, C., C. Herron, and S. Cole. 2007. "The Effects of Deductive and Guided Inductive Instructional Approaches on the Learning of Grammar in the Elementary Foreign Language College Classroom." *Foreign Language Annals* 40 (2): 288–310. doi:10.1111/j.1944-9720.2007.tb03202.x.
Halliday, M., and C. Matthiessen. 2004. *An Introduction to Functional Grammar*. London: Arnold.
Havekes, H. 2015. *Knowing and Doing History. Learning Historical Thinking in the Classroom*. Doctoral dissertation. Radboud University Nijmegen.
Hudson, R. 2010. "Linguistic Theory." In *The Handbook of Educational Linguistics*, edited by B. Spolsky and F. Hult, 53–65. Sussex: Wiley-Blackwell.
Hudson, R., and J. Walmsley. 2005. "The English Patient: English Grammar and Teaching in the Twentieth Century." *Journal of Linguistics* 41 (3): 593–622. doi:10.1017/S0022226705003464.
Hulshof, H. 1985. *C.H. Den Hertog Als Grammaticus [C.H. Den Hertog as a Grammarian] Doctoral Dissertation*. Coutinho, Muiderberg: KU Nijmegen.
Hulshof, H. 2013. "Onderzoek Naar Taalkundeonderwijs: Geloof en Empirie." In *Neerlandistiek in Beeld*, edited by T. Janssen and T. van Strien, 263–272. Amsterdam: Neerlandistiek VU.
Jones, P., and H. Chen. 2012. "Teachers' Knowledge about Language: Issues of Pedagogy and Expertise." *Australian Journal of Language and Literacy* 35 (2): 147–168.
King, P.M., and K.S. Kitchener. 2004. "Reflective Judgement: Theory and Research on the Development of Epistemic Assumptions through Adulthood." *Educational Psychologist* 39 (1): 5–18. doi:10.1207/s15326985ep3901_2.
Kuiper, K., and J. Nokes. 2014. *Theories of Syntax. Concepts and Case Studies*. New York: Palgrave Macmillian.
Lantolf, J., S. L. Thorne, and M. Poehner. 2015. "Sociocultural Theory and Second Language Development." In *Theories in Second Language Acquisition*, edited by B. van Patten and J. Williams, 207–226. New York: Routledge.

Lefstein, A. 2009. "Rhetorical Grammar and the Grammar of Schooling: Teaching 'Powerful Verbs' in the English National Literacy Strategy." *Linguistics and Education* 20 (4): 378–400. doi:10.1016/j.linged.2009.04.002.

Lesterhuis, M., S. Verhavert, L. Coertjens, V. Donche, and S. De Maeyer. 2016. "Comparative Judgement as Promising Alternative to Score Competences." In *Innovative Practices for Higher Education Assessment and Measurement,* edited by G. Ion and E. Cano, 120–140. Hershey: IGI Global.

Lipman, M. 2003. *Thinking in Education.* 2nd ed. Cambridge University Press.

Locke, T. (Ed.). 2010. *Beyond the Grammar Wars. A Resource for Teachers and Studentson Developing Language Knowledge in the English/Literacy Classroom*. New York: Routledge.

Macken-Horarik, M.,. K. Love, and S. Horarik. 2018. "Rethinking Grammar in Language Arts: Insights from an Australian Survey of Teachers' Subject Knowledge." *Research in the Teaching of English* 52(3): 288–316.

McCabe, A. 2017. "Systemic Functional Linguistics and Language Teaching." In *The Routledge Handbook of Systemic Functional Linguistics*, edited by T. Bartlett and G. O'Grady, 615–628. London and New York: Routledge.

Mercer, N. 2003. "Helping Children to Talk and Think Together More Effectively." *POLIFONIA* 7: 1–26.

Mercer, N. 2004. "Sociocultural Discourse Analysis: analyzing Classroom Talk as a Social Mode of Thinking." *Journal of Applied Linguistics* 1 (2): 137–168. doi:10.1558/japl.2004.1.2.137.

Mercer, N. 2013. "The Social Brain, Language, and Goal-Directed Collective Thinking: A Social Conception of Cognition and Its Implications for Understanding How We Think, Teach, and Learn." *Educational Psychologist* 48 (3): 148–168. doi:10.1080/00461520.2013.804394.

Moore, A. 2004. *The good teacher: Dominant discourses in teaching and teacher education*. London and New York: Routledge.

Moseley, M., et al. 2005. *Frameworks for Thinking. A Handbook for Teaching and Learning*. Cambridge: Cambridge University Press.

Myhill, D. 2000. "Misconceptions and Difficulties in the Acquisition of Metalinguistic Knowledge." *Language and Education* 14 (3): 151–163. doi:10.1080/09500780008666787.

Myhill, D. 2003. "Principled Understanding? Teaching the Active and Passive Voice." *Language and Education* 17 (5): 355–370. doi:10.1080/09500780308666856.

Myhill, D. 2005. "Ways of Knowing: Writing with Grammar in Mind." *English Teaching: Practice and Critique* 4 (3): 77–96.

Myhill, D. 2018. "Grammar as a Meaning-Making Resource for Improving Writing." *L1 Educational Studies in Language and Literature* 18: 1–21. doi:10.17239/L1ESLL-2018.18.04.04.

Myhill, D., and S. M. Jones. 2015. "Conceptualising Metalinguistic Understanding in Writing." *Cultura y Educación* 27 (4): 839–867. doi:10.1080/11356405.2015.1089387.

Myhill, D., S. Jones, and H. Lines. 2018. "Supporting Less Proficient Writers through Linguistically Aware Teaching." *Language and Education* 32 (4): 333–349. doi:10.1080/09500782.2018.1438468.

Myhill, D., S. Jones, H. Lines, and A. Watson. 2012. "Re-Thinking Grammar: The Impact of Embedded Grammar Teaching on Students' Writing and Students' Metalinguistic Understanding." *Research Papers in Education* 27 (2): 139–163. doi:10.1080/02671522.2011.637640.

Perini, M. 2015. *Describing Verb Valency. Practical and Theoretical Issues*. New York: Springer.

Plomp, T., and N. Nieveen. 2007. *Educational Design Research*. Enschede: SLO.

Pollitt, A. 2012. "Comparative Judgement for Assessment." *International Journal of Technology and Design Education* 22 (2): 157–170. doi:10.1007/s10798-011-9189-x.

Prince, M., and R. Felder. 2006. "Inductive Teaching and Learning Methods: Definitions, Comparisons, and Research Bases." *Journal of Engineering Education* 95 (2): 123–138. doi:10.1002/j.2168-9830.2006.tb00884.x.

Ribas, T., X. Fontich, and O. Guasch. eds. 2014. *Grammar at School. Research on Metalinguistic Activity in Language Education*. Brussels: Peter Lang.

Ritter, S. 2012. *Creativity: understanding and Enhancing Creative Thinking*. Doctoral dissertation. Nijmegen: Radboud University Nijmegen.

Sadler, D. 2009. "Transforming Holistic Assessment and Grading into a Vehicle for Complex Learning." In *Assessment, Learning and Judgement in Higher Education,* edited by G. Joughin, 1–19. Dordrecht: The Netherlands.

Sangster, P., C. Anderson, and P. O'Hara. 2013. "Perceived and Actual Levels of Knowledge about Language Amongst Primary and Secondary Student Teachers: do They Know What They Think They Know?." *Language Awareness* 22 (4): 293–319. doi:10.1080/09658416.2012.722643.

Seuren, P. 1998. *Western Linguistics. An Historical Introduction*. Malden: Blackwell Publishers.

Shadish, W., T. Cook, and D. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont: Wadsworth.

Van Drie, J., and C. Van Boxtel. 2008. "Historical Reasoning: Towards a Framework for Analyzing Students' Reasoning About the Past." *Educational Psychology Review* 20 (2): 87–110. doi:10.1007/s10648-007-9056-1.

van Drie, J., and C. van Boxtel. 2011. "In Essence I'm Only Reflecting: teaching Strategies for Fostering Historical Reasoning in Whole-Class Discussion." *History Education Research Journal* 10 (1): 55–66. doi:10.18546/HERJ.10.1.05.

Van Gelderen, A. 2010. "Does Explicit Teaching of Grammar Help Students to Become Better Writers? in-Sights from Empirical Research.". In T. Locke (Ed.), *Beyond the Grammar Wars. A Resource for Teachers and Students on Developing Language Knowledge in the English/Literacy Classroom,* 109–128. New York, NY: Routledge.

Van Rijt, J., and P.-A. Coppen. 2017. "Bridging the Gap between Linguistic Theory and L1 Grammar Education — Experts' Views on Essential Linguistic Concepts." *Language Awareness* 26 (4): 360–380. doi:10.1080/09658416.2017.1410552.

Van Rijt, J., P. De Swart, and P.-A. Coppen. 2019. "Linguistic Concepts in L1 Grammar Education: A Systematic Literature Review." *Research Papers in Education* 34 (5): 621–648. doi:10.1080/02671522.2018.1493742.

Van Rijt, J., P. De Swart, A. Wijnands, and P.-A. Coppen. 2019. "When Students Tackle Grammatical Problems: Exploring Linguistic Reasoning with Linguistic Metaconcepts in L1 Grammar Education." *Linguistics and Education* 52: 78–88. doi:10.1016/j.linged.2019.06.004.

Van Rijt, J., A. Wijnands, and P.-A. Coppen. 2019. "Dutch Teachers' Beliefs on Linguistic Concepts and Reflective Judgement in Grammar Teaching." *L1 Educational Studies in Language and Literature* 19: 1–28.. doi:10.17239/L1ESLL-2019.19.02.03.

Verhavert, S., S. De Maeyer, V. Donche, and L. Coertjens. 2018. "Scale Separation Reliability: What Does It Mean in the Context of Comparative Judgement." *Applied Psychological Measurement* 42 (6): 428–445. doi:10.1177/0146621617748321.

Watson, A. 2015. "Conceptualisations of Grammar Teaching: L1 English Teachers' Beliefs about Teaching Grammar for Writing." *Language Awareness* 24 (1): 1–14. doi:10.1080/09658416.2013.828736.

Watson, A., and R. Newman. 2017. "Talking Grammatically: L1 Adolescent Metalinguistic Reflection on Writing." *Language Awareness* 26 (4): 381–398. doi:10.1080/09658416.2017.1410554.

Wijnands, A. (2017, June). *The use of Reference Grammars in secondary education*. Paper presented at the ARLE PhD Preconference in Tallinn, Estonia.

Appendix 1. Description of the intervention

	Lesson 1	Lesson 2	Lesson 3	Lesson 4
Central linguistic metaconcept	Predication	Predication + valency	Valency + complementation	Complementation + modification
Central secondary concepts from traditional grammar	• Verbs • Verbal predicate • Subject complement	• Verbs (main vs. auxiliary) • Objects (direct, indirect) • subject	• Main verbs • Objects (direct, indirect) • Subject • Adverbials	• Main verbs • Adverbials • Adjectivals • Objects
Typical assignment related to design principles 1. **Relate metaconcepts to traditional concepts** 2. **Using students' own intuitions to grasp a linguistic metaconcept** 3. **Dealing with uncertainty (ill-structuredness)** 4. **Exploratory talk** 5. **Appropriate scaffolding**	Students receive an enveloppe with plasticized cards containing two word utterances from toddlers, such as *mama lief* ('mommy sweet') and *papa slapen* ('daddy sleeping'). In a group of four, they have to unanimously decide based on which criterium they will devide the cards into two even groups. The best categorization involves cards that are dealing with 'being' (e.g. mommy sweet) versus doing (e.g. daddy sleeping). This distinction aims to inductively didacticize the metaconcept of *predication*, and it serves as a stepping stone for explaining the distinction between verbal predicates and subject complements. Most cards fit into the categories nicely, except for one—this is meant to invoke critical thinking, and illustrates the ill-structuredness of language. The teacher reflects with the students on what they have learned from this task and guides their thinking about categorizing the sentences. (1,2,3,4, 5)	Students receive yellow cards, containing different verbs, and blue cards, containing the words *iemand* ('*someone*') or *iets* ('*something*'). They are then asked to determine how many blue cards can fit into the pattern of the verb. This creates categories of verbs needing one, two or three roles ('valency'): the verb *to give* requires three syntactic roles (*someone gives something to someone*), whereas a verb like *to walk* only requires one: *someone walks*, and not: **someone walks something*. After students have categorized the verb patterns into three categories, they need to replace the blue cards with green ones, which contain noun phrases, thus forming 'normal' sentences. The teacher stimulates students' reflection on what they have discovered from this excercise (1,2,4, 5).	Students are presented with a set of four verbs. Using the sentence 'X is the odd one out, because the other three verbs …', they have to decide which of the verbs is the odd one out. In principle, each of these verbs could be the odd one out, but the argumentation is crucial in deciding which one. Students could use different arguments for their choice (e.g. 'the valency pattern of verb X is different from the other verbs', or 'verb X can be made passive, whereas the other verbs cannot). Students tackle this assignment in a group of four, and they have to reach a general agreement on which verb is the best one to single out. The teacher has an important task in helping the students to reason—he/she needs to guide (meta)concept use and stimulate a deeper argument, without giving 'the answer'. (1, 3, 4, 5)	Four students are presented with a grammatical discussion, in which Tom says that in sentence X, Y is a direct object, whereas Bert says that Y is an adverbial in that sentence. Students are then presented with several arguments that are either in favor of Tom's point of view or Bert's. The students then have to decide which argument strengthens either Tom's or Bert's case, or which argument refutes either case. At the end, the group will have categorized all arguments and they will see that there is evidence for both Tom's and Bert's position. It will be up to them to decide which point of view can best be defended based on the available evidence/ argumentation. This way, students learn to deal with ill-structuredness, and they will find that knowledge on metaconcepts can help in their joint discussion. The teacher crucially guides the students' discussion (1, 2, 3, 4, 5).

Appendix 2. Grammatical problems

Target items	Filler items
1. Review the sentence Ik heb de man met de verrekijker gezien ('I have seen the man with binoculars'). Explain that this sentence has two possible meanings, and explain how this is possible using grammatical terminology.	5. Review sentences a and b below. a. Rinus zegt tegen Joop dat Wim zichzelf heeft gefotografeerd ('Rinus tells Joop that Wim photographed himself') b. Rinus zegt tegen Joop dat Wim hem heeft gefotografeerd ('Rinus tells Joop that Wim photographed him') A student from abroad that came to the Netherlands asks you what the word *zichzelf* ('himself') can refer to in these sentences. Explain how it is possible that *zichzelf* in a. cannot refer to Joop or Rinus, but it can to Wim, whereas *hem* ('him') in b. cannot refer to Wim, but it can to the other two.
2. Review the newspaper headline Stichting wil ouders in narigheid helpen ('Foundation wishes to help parents in trouble')* Explain that this sentence has two possible meanings and explain how this is possible using grammatical terminology.	6. Review sentences a and b below. a. De vader van Jan fotografeert hem ('Jans father photographs him') b. Jan fotografeert hem ('Jan photographs him') A student from abroad that came to the Netherlands asks you what the word *him* in these sentences can refer to. Explain how it is possible that *hem* in a. cannot refer to *vader*, but that it can to *Jan*, whereas *him* cannot be used to refer to Jan.
3. Below you will see two sentences. How is it possible that the a-sentence is ungrammatical, wehereas the b-sentence is grammatical? Try to explain this as good as possible using grammatical terminology. a. *Jan blies het blaadje. ('Jan blew the paper') b. Jan blies het blaadje van tafel. ('Jan blew the paper off the table')	
4. Below you will see two sentences. How is it possible that the a-sentence is ungrammatical, wehereas the b-sentence is grammatical? Try to explain this as good as possible with using grammatical terminology. a. *Jan regent ('Jan rains') b. Jan regent nat ('Jan rains wet' i.e. Jan becomes wet due to the rain).	

Note Translations by the authors. Target items 3 and 4 were designed as ill-structured grammar problems, for which traditional school grammar does not offer clear solutions. All of the grammatical problems were of a kind student has never seen before.