Reading and Writing (2023) 36:2605–2625
https://doi.org/10.1007/s11145-022-10405-z

The relationship between grammatical understanding and writing skills in Finnish secondary L1 education

Jenni Marjokorpi[1]

Accepted: 16 December 2022 / Published online: 9 January 2023

Abstract

Previous studies have indicated that students' writing skills benefit from contextualized L1 grammar teaching, in which language structures are observed and analyzed in authentic texts and as embedded into teaching of reading and writing. The contextualized approach is also promoted by the current Finnish curriculum for basic education. This study investigates the relationship between grammatical understanding and writing skills using statistical methods as well as a complementary qualitative analysis of student texts. The data are derived from a large cross-sectional assessment of L1 learning achievement in Finnish year 9 students (N = 6,044). Linear regression analyses indicate that grammatical understanding is a significant predictor of writing skills and correlates strongly with the syntactic, stylistic, genre-related, and orthographical quality of the students' argumentative texts. Weaker writers use less complex vocabulary and sometimes "lose control" of syntactic structures. Learning grammar is related to metalinguistic understanding which, in turn, helps writers to analyze and control their language use, and thus produce better texts.

Keywords Grammar · Writing · Metalinguistic understanding · First language · Secondary education

Introduction

It has been a topic of long debate, whether the teaching of first language (L1) grammar benefits the students' writing skills. Some empirical studies, including systematic reviews, have come to critical conclusions, indicating that L1 grammar teaching cannot be justified from the viewpoint of writing development, at least in the ways it is traditionally taught (Andrews et al. 2004; Graham & Perin, 2007; Hillocks, 2006; Wyse, 2001; Wyse & Torgerson, 2017). In contrast, other studies have indicated that such benefit can be achieved if grammar is taught in a *contextualized* way (Collins

✉ Jenni Marjokorpi
jenni.marjokorpi@helsinki.fi

[1] Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

& Norris, 2017; Costa, 2019; Derewianka, 2012; Fearn & Farnan, 2007; Fontich, 2016; Jones, Myhill, & Bailey, 2013; Myhill, Jones, Lines, & Watson, 2012). This approach stems from systemic-functional grammar which connects language forms and their functions in grammatical description, as well as emphasizes the social nature of language use (Derewianka & Jones, 2010; Halliday & Matthiessen, 2004; Myhill, 2021). On a pedagogical level, the contextualized and functionally oriented grammar teaching is actualized in, for example, applying grammar as a tool for linguistic reflection in reading and writing, and using authentic text examples as a starting point for metalinguistic discussion (Derewianka & Jones, 2010; Fontich & Camps, 2014; Myhill et al. 2020; Schleppegrell, 2010).

The functionally oriented and contextualized approach is also promoted by the current Finnish National Core Curriculum for Basic Education (NCC, 2016), in which L1 teaching is based on "a social and functional perception of language: the structures of language are studied in age-appropriate language-use situations and while working with text genres suitable for the age group". The orientation is also seen in the curricular placement of grammar within the content areas of interpreting and producing texts. It also has historical roots, as Finnish linguists and educators have long emphasized that school grammar needs to build on the practical basis of reading textbook chapters (Cygnaeus 1861 in Koskinen, 1988, 96), and it should not appear as irrelevant and dull (Hakulinen, 1925; Tainio, 2020). Despite this practical orientation and the faith in the potential benefits of grammar teaching on, for example, text interpretation skills (Leino & Sääskilahti, 2010; Pallaskallio, 2016; Paukkunen, 2011) and cognitive skills in general (Rättyä, 2017; Rättyä & Kulju, 2018; Tainio, 2020), there is little research on the benefits of grammar for writing in the Finnish context.

The debate about the benefits of grammar teaching forms the starting point of this paper. I investigate the statistical relationship between grammatical understanding and writing skills in Finland, where the curriculum advocates contextualized grammar teaching. To better understand the possible connection between the two sets of skills, I also analyze the linguistic structures in student texts and compare them to the writer's level of grammatical understanding. The data are derived from a nationwide assessment of L1 learning achievement in year 9 students ($N = 6{,}044$), originally conducted by the Finnish Education Evaluation Centre (FINEEC). The assessment evaluates the students' reading and writing skills and knowledge about language at the end of basic education. In this study, I focus on grammar and writing

Setting the context: teaching of grammar in Finland

Rättyä, Awramiuk, and Fontich (2019) note that the question about the benefits of grammar on writing skills cannot be resolved without understanding how grammar is taught and perceived in different educational contexts. Therefore, I first describe the educational context of L1 (Finnish) grammar teaching in Finland (see also Nupponen et al. 2019).

Knowledge about language (KAL) is understood in Finland as an umbrella concept which encompasses different areas of language study, such as pragmatics,

variation and change, stylistics and discourse studies, sociolinguistics, language identities, and grammar, defined as the study of language structure (Tainio, 2020). The concept of KAL is therefore more educational than purely linguistic by nature, and it has also been used internationally in the context of language education (e.g. Ellis, 2008; Hudson, 2015; Matruglio, 2020). The concept of grammar is not limited to syntax but also includes phonology, morphology, and word formation, as well as the differences between spoken and written language (Carter & McCarthy, 2006; Karlsson, 2008). Both KAL and grammar (as its subarea) are taught to foster language awareness, which is understood as a broad concept, concerning not only knowledge and competences but also attitudes and understanding the importance of language in human life (Kalliokoski et al. 2015; Svalberg, 2012). Already for decades, Finnish L1 teachers (cf. Koskinen, 1990) and researchers (Savolainen, 1998; Varis, 2012) have argued that prescriptive teaching should not form the core of KAL or grammar. Instead, school grammar should—alike academic linguistics—focus on describing *what language is like*, not what it *should be like*.

The Finnish curriculum (NCC, 2016) presents reading and writing instruction in the context of multiliteracies. This perspective emphasizes learning to (critically) interpret and produce multimodal texts (Cope & Kalantzis, 2000; New London Group, 1996). Multiliteracies are connected to understanding how texts are used to construct meanings (Unsworth, 2001), and this requires metalinguistic understanding, especially the kind that is functional and considers the social contexts of language use. The NCC sees grammatical norms as flexible and varying, and linguistic competence is valued more than remembering grammatical concepts (Kalliokoski et al., 2015). The grammar curriculum is implemented in schools in a rather contextualized way: 60% of Finnish upper comprehensive school L1 teachers report teaching KAL together with other L1 content areas (Harjunen & Rautopuro, 2015). According to the teachers, the most typical contents of KAL are grammatical concepts and written language conventions, and regular teaching methods vary from reading textbooks and doing grammar exercises to writing, linguistic inquiry, and discussion (Harjunen & Rautopuro, 2015). When it comes to literacy skills, Finnish students have often succeeded in international assessment studies: they have reached the top five of OECD countries in all PISA studies (2003; 2004; 2007; 2010; 2014; 2016; 2019) so far and performed well also in PIRLS studies (*Progress in International Reading Literacy Study,* Mullis, Martin, Foy, & Drucker 2012; Mullis, Martin, Foy, & Hooper, 2017). However, these studies have concentrated on reading literacy, and the Finnish 'PISA fever' seems to have overridden the academic interest on writing skills, which has been lesser (Harjunen, Marjanen, & Karlsson, 2019). Based on national assessment studies, Routarinne & Absetz (2013) have even noted that while the results in reading literacy have been outstanding, the same students struggle with writing literacy, which is in downhill slope.[1] Similar observations have been made, for example, in Australia (Chen, Lewis, & Myhill, 2021).

[1] The gap between reading and writing skills may also have historical roots, as the Finnish society has required basic reading—but not writing—skills of the whole population already since the 17th century. In the 1890s, nearly all adult Lutheran Finns were literate, but only 20% of them were also able to write. (Sinnemäki & Saarikivi, 2019.)

How might grammar enhance writing quality?

For a layperson, grammar often equals error-centered prescriptivism, which has also been the core of traditional school grammar (Hancock, 2009; Hudson, 2004). From a linguistic perspective, grammar especially means description of language structures (e.g. syntax, morphology, and phonology), often following a certain linguistic tradition, such as generative, structural, or functional linguistics. The third meaning of grammar is the implicitly acquired set of norms that conduct the use of language by its speakers. Grammar in this third sense is rather indisputably needed in writing, as in any language use. In contrast, teaching detached, prescriptive, and error-centered grammar has appeared useless for writing in multiple studies (Andrews et al., 2004; Graham & Perin, 2007; Hillocks, 2006; Koster et al. 2015; Wyse, 2001). Some other studies have adopted an updated conception of grammar which is more contextualized, functional and discussive in nature, which Myhill et al. (2012) consider being closer to the understanding of grammar within modern linguistics. In studies which have conceptualized grammar in this way, its teaching has appeared beneficial (e.g., Collins & Norris, 2017; Costa, 2019; Derewianka, 2012; Fearn & Farnan, 2007; Fontich, 2016; Myhill et al., 2012). However, the debate is not yet settled, and more research about the issue has been called for (e.g., Chatterjee & Halder, 2022; Wyse & Torgerson, 2017).

The grammar debates essentially include the question that, if grammar enhances writing skills, how does this enhancement take place. The observed benefits of grammar for writing have been explained using the concept of *metalinguistic understanding,* defined by Myhill (2012, 250) as "the explicit bringing into consciousness of an attention to language as an artifact, and the conscious monitoring and manipulation of language to create desired meanings, grounded in socially shared understandings". Metalinguistic understanding enables the students to see how texts construct meanings. Studying grammar also teaches the skill and the habit of taking language under conscious reflection (Wijnands, van Rijt, & Coppen, 2021). This reflective metalinguistic activity does not only mean analyzing grammatical structures, but also taking under consideration viewpoints such as rhetorical devices, argumentative structures, and social discourses. For example, syntactic concepts lead the students to observe the functions of different words and phrases within a clause or sentence (Chipere, 2003), whereas morphological instruction raises awareness of lexical structures, which has appeared beneficial for reading comprehension (Bowers, Kirby, & Deacon, 2010; Goodwin & Ahn, 2010; Kirby, Deacon, Bowers, Izenberg, Wade-Woolley, & Parrila, 2012).

Secondly, when students learn to observe and analyze language they read, they can also learn to observe and analyze the language that they produce. Within writing instruction, "going meta" refers to setting language as a topic of explicit observation and discussion (Myhill et al., 2020), where the systematics of grammar provide a widely shared and useful metalanguage. Watson, Newman, and Morgan (2021) distinguish *declarative* metalinguistic knowledge, which is conscious, explicit, and verbalizable, from *procedural,* which can be either conscious or unconscious implicit knowledge taking place in language use. According to Fontich (2016), these types

of metalinguistic knowledge should be, yet insufficiently are, linked for grammar teaching to be effective from the viewpoint of writing development.

When it comes to assessing writing quality, more developed writers generally produce texts with higher lexical sophistication, syntactic complexity, and textual cohesion (Crossley, 2020), whereas weaker writers prefer structures that are more common in speech and focus on content rather than the forms which communicate it (Myhill, 2008). The quality of student writing increases by age in lexical, syntactical, morphological, and macro-structural ways, as well as in relation to content, argumentation, and following of standard language conventions (Hankala et al. 2015; Pajunen, 2012; Pajunen & Honko, 2021). Skilled or older writers also tend to use more concise language (Schleppegrell & Christie, 2018), although this is also a matter of personal style as well as genre (Pajunen & Honko, 2021). However, growing up does not automatically mean advancing in linguistic skills, but interaction, linguistic input, and teaching play a crucial role (Pajunen, 2012; Tomasello, 2003). The role of metalinguistic understanding in writing has been examined by Lappalainen (2004), according to whom the students' knowledge about language structures, metalinguistic concepts, writing conventions, vocabulary, and idioms explains 39% of their writing performance. Harjunen and Rautopuro (2015) make a similar observation, with the explanation rate as high as 55%. Both studies have been based on a similar national assessment data as I will analyze in this study, but from previous age cohorts, which leads to hypothesize that a similar connection might be found again this time. In addition, I aim at further describing and explaining the possible connection in a more nuanced way, using both quantitative and qualitative analysis methods.

Present study

Research objectives

The aim of this paper is to examine the statistical relationship between grammatical understanding (GU) and writing skills in Finnish year 9 students. The second aim is to describe this potential relationship and the benefits of GU for student writing in a qualitative way. In this study, the concept of GU is understood as a type of metalinguistic understanding which is especially related to language structures.

Data collection and participants

The research data have been collected by FINEEC in 2019 for the purposes of a national assessment study of learning outcomes (Hellgren & Kauppinen, 2020; Kauppinen & Marjanen, 2020). The data comprise the responses of 6,044 year 9 students (ca. 15-year-olds) from 118 Finnish-language secondary schools all around the country. In the sampling, all Finnish schools were first categorized according to region and municipality type (city, town, rural area), to ensure that all regions and municipality types are represented according to the real population. Within

these categories, the sampling was randomized on the levels of school and, in bigger schools (with more than 50 year 9 students) also on the individual level (see Kauppinen & Marjanen, 2020, 34–35). The data illustrate the linguistic competences acquired during basic education, of which year 9 is the final stage. The data were collected in schools in 2019 using FINEEC's digital evaluation system which assigned the students both open-ended and multiple-choice tasks. The students completed the assignments individually and under teachers' supervision using computers but not Internet or other materials. They took the assignments in two parts, both within a 90-minute time frame. The first part included two text production items and a background survey (e.g., attitudes towards studying L1 and literature, parental education and support levels, completing homework, screen time, home language, receiving of support due to learning difficulties, reading of books, library visits, and L1 grades). The second part measured multimodal text interpretation skills and KAL. In this paper, I focus on the tasks measuring writing and GU, which I understand as a component of KAL.

As a researcher, I have no affiliation to FINEEC, but the Centre has granted me official permission to use the assessment data for the purposes of this study. The data was transferred to me on a visit to FINEEC, during which I discussed my research plans and was given guidance in using the data. The officials of FINEEC have also commented on my research, but I am responsible for all the analyses and interpretations.

The measurement of writing skills

Writing skills were assessed using two text production assignments (Table 1). First, the students produced an *argumentative text* (AT) about one of the five preassigned topics (e.g. nature conservation) that were designed to be familiar and thought-provoking to them. Second, they wrote a *reflective text* (RT) in which they discussed their own AT writing processes and themselves as writers. Both texts were written on a computer and keyboard without any visual or typographic formatting nor multimodalities; for this reason, this paper prefers the term *writing* instead of the multimodal *text production* (Jakobs & Perrin, 2014). The FINEEC assessment defined the skills of text production in a holistic way, in which only 4 out of 26 points came from orthography whereas reflective abilities were highly emphasized (Table 1). The advised word count for ATs ranged from 200 to 250. Both texts were anonymously assessed by teachers according to pre-established criteria (see Table 1). For example, the teachers gave 0 to 4 points about the structure of the text, comparing the text properties to the assessment criteria for each point. To evaluate inter-rater reliability, a random 10% of the texts were re-assessed by one or two FINEEC's evaluators, and the conformity of the teacher and evaluator ratings was analyzed statistically by Kauppinen and Marjanen (2020, 36–37; see also Marjanen, 2020) who conclude that the overall reliability was good (intra-class correlation coefficient, ICC > 0.75).

Table 1 The items measuring writing skills and GU selected from the FINEEC data (see also Kauppinen & Marjanen 2020, 33, 61, 73)

Content areas	Assignments and assessment criteria	Max. score	Item type
Writing	**Argumentative text (AT)**	14	Assessment criteria, applied by teachers and evaluators
	Structure	4	
	Content and genre-specific qualities	4	
	Spelling, punctuation, and grammar	4	
	Style	2	
	Reflective text (RT)	12	
	Self-assessment of AT quality	3	
	Reflection on AT writing process	3	
	Reflection about oneself as a writer	3	
	Overall quality of the RT	3	
Grammatical understanding (GU)	Forms and functions of language structures	11	9 multiple-choice items, 2 open-ended items
	Standard language norms	2	1 open-ended item
	Morphology	2	2 multiple-choice items

The measurement of grammatical understanding

While the measures of writing come straight from the original assessment, I created a new compound variable of grammatical understanding (GU). The original data had 22 KAL items, the topics of which ranged from the linguistic landscape of Finland to language variation and change, and from morpho-syntax to word formation. This is in concert with the curricular contextualization of KAL. For the purposes of this study, I chose 14 items that were related to grammar, or the morpho-syntactic forms of language and their functions; I excluded the items that measured more factual and general knowledge about language as a phenomenon.

The multiple-choice items mostly measured remembering (e.g., what word in a text belongs to a certain category), understanding (e.g. understanding certain morpho-phonological phenomena), and analyzing (e.g. analyze word formation types), whereas one of the open-ended items required applying grammatical knowledge (Kauppinen & Marjanen, 2020). In one of the open-ended items, the students had to correct the spelling, punctuation, and grammar (SPaG) of a short erroneous text. The multiple-choice items were scored automatically by FINEEC's digital evaluation system, and the open-ended responses were scored by teachers; I made no changes to the scoring. The assignments cannot be reprinted due to their possible use in the future evaluations.

The mean GU score was 8.64 (SD = 3.12, N = 5,454). The maximum score of 15 was obtained by 0.8% of the students. The distribution of scores was approximately normal (skewness − .274, SES = .033, kurtosis − .599, SEK = .066), even though Kolmogorov-Smirnov's test failed to confirm its normality (.097, $p < .001$). The suitability of the GU compound variable was further tested using principal

component analysis (PCA) with Quatrimax rotation. The 14 items loaded onto three components, the eigenvalue of which was greater than 1. The first component (λ = 3.419) consisted of three items measuring knowledge about word formation. The second component (λ = 1.386) was related to knowledge about phonology, morpho-syntax, and written language conventions, and the third one (λ = 1.209) measured semantic understanding of derived words that differed from one another in small morphological detail. One item belonged to both the second and the third component; the first one was more clearly separate. Thus, the compound variable measured different kinds of skills, which is in accordance with the theoretical assumption that grammatical understanding consists of different kinds of subareas: for example, syntactical, lexical, phonological, and morphological. Nevertheless, those areas tend to overlap, and in Finnish language, the boundaries between, for example, morphology and syntax are not always clear (Karlsson, 2008). Cronbach's alpha for GU compound variable was .735, suggesting acceptable internal consistency, and therefore I decided to proceed by using the GU compound variable, and not the components extracted in the PCA.

Methods of data analysis

I analyze the data in both quantitative and qualitative way. The quantitative analyses focus on the statistical relationship between GU and writing, as well as some selected background variables, using stepwise linear regression analysis in SPSS 27. The background variables included: attitudes towards studying L1, reading books, completing homework, and parental education level. For a more detailed analysis, I used the random selection tool in SPSS 27 to extract from the randomly arranged data file of all student responses a subsample of 150 student ATs (of which 12 ATs were blank or missing, and therefore N = 138). I used the *Analyzemywriting.com* online software to count the basic syntactic statistics, such as word and sentence lengths and their variations, for each text (headlines excluded). The results were quantitatively and qualitatively compared to the students' GU scores and the morpho-syntactic features of their texts. In the qualitative part, I performed a linguistic text analysis to examine the morphological and syntactical structures that are used by students of different GU levels.

The relationship between GU and writing skills

Statistical analyses

Table 2 presents the correlations between writing, GU, and the selected background variables. The strongest correlation was between GU and writing. The relationships were further tested with a stepwise linear regression analysis (Table 3). The regression analysis shows that grammar explains 32.1% of the variance of writing skills, whereas attitudes, homework and reading habits, and home background have

Table 2 Bivariate correlations of writing skills and related variables

	Writing	1.	2.	3.	4.	5.	6.	7.
1. GU	.565**							
2. Usefulness of L1	.442**	.343**						
3. Self-efficacy	.303**	.282**	.128**					
4. Liking L1	.368**	.225**	.535**	.352**				
5. Book reading	.274**	.249**	.239**	.184**	.279**			
6. Homework	.415**	.346**	.387**	.249**	.385**	.221**		
7. Maternal EL	.193**	.215**	.131**	.116**	.074**	.066**	.105**	
8. Paternal EL	.187**	.210**	.141**	.093**	.074**	.103**	.085**	.505**

** All the correlations are significant at the .01 level (2-tailed).

Note. "Usefulness of L1", "self-efficacy", and "liking L1" are each derived from mean scores of five Likert-scaled questions assessing attitudes towards L1 teaching. "Book reading" refers to the N of books (apart from textbooks) read per month (a multiple-choice question; the highest value being "4 or more books per month"); "Homework" refers to the frequency of doing L1 homework; "Maternal/paternal EL" is the educational level of the student's parents—according to self-report.

Table 3 Stepwise regression of predictors of writing skills

Dependent: Writing	Adjusted *R* square	B	Standard error	Beta	Significance of *t*
(Constant)		−5.22			< .001
GU	.321	.764	4.83	.403	< .001
Usefulness of L1	.379	1.09	4.62	.162	< .001
Homework	.423	.667	4.54	.118	< .001
Self-efficacy	.433	.857	4.49	.095	< .001
Book reading	.438	.467	4.47	.074	< .001
Liking L1	.441	.438	4.46	.074	< .001
Maternal EL	.443	.163	4.45	.035	< .001
Paternal EL	.443	.147	4.45	.032	< .001

a smaller explanatory power even altogether. The approximately linear relationship between grammar and writing is visualized in Fig. 1.

The relationship between grammar and specific measures of writing

Table 4 shows significant correlations of grammar with all the assessment criteria of ATs and RTs, especially AT spelling, punctuation, and grammar (following the norms of standard written language), and structure (see also Table 1 above). However, GU also correlates with content (what is said), genre-specific features (following the conventions of argumentative texts), and style (e.g., no spoken language style).

AT quality correlated strongly with text length ($r = .650$, $p < .001$) and with GU (Table 4). A regression analysis confirmed that the effect of grammar on text

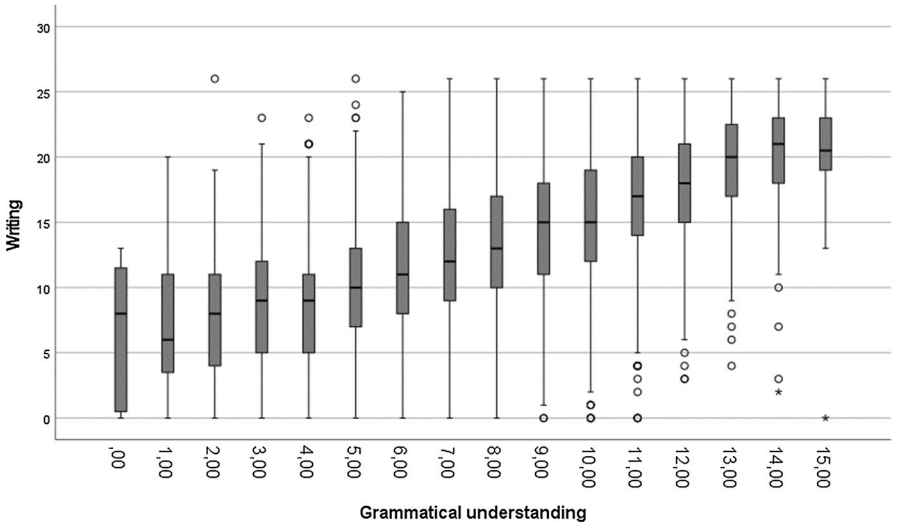

Fig. 1 The relationship between writing skills and grammatical understanding

Table 4 Assessment criteria of writing tasks and their bivariate correlations with grammatical understanding (GU)

	Max	Mean	SD	r with GU
Grammatical understanding	16	8.64	3.12	
AT: Structure	4	2.13	1.32	.484**
AT: Content and genre-specific qualities	4	2.18	1.19	.466**
AT: Spelling, punctuation, and grammar	4	2.20	1.10	.520**
AT: Style	2	1.30	.632	.420**
Total AT score	14	7.81	3.66	.553**
AT length (words)[1]	400	169.0	83.5	.354**
RT: AT self-assessment	3	1.41	.797	.352**
RT: Writing process	3	1.61	.809	.405**
RT: Me as a writer	3	1.67	.851	.426**
RT: Text as a whole	3	1.59	.797	.458**
Total RT score	12	6.27	2.79	.479**
Total writing score (AT + RT)	26	14.0	5.99	.565**
Time on task (min)[2]	105	58.9	17.3	.381**

** $p < .001$. [1]Texts longer than 400 words ($N = 13$) were excluded due to violation of word count guidelines. [2]Time scores longer than 105 minutes were excluded due to violation of time limit.

quality remains significant even when word count has been controlled (Table 5). Longer texts thus tend to be better (see also Kauppinen & Marjanen, 2020, 99–101), but they may be even better, if they are written by a student with high

Table 5 Stepwise regression of the effect of word count and grammatical understanding for AT quality

Dependent: AT	Adjusted *R* square	B	Standard error	Beta	Significance of *t*
(Constant)		−.544			< .001
Word count	.413	.027	2.77	.345	< .001
Grammatical understanding	.511	.402	2.53	.499	< .001

Table 6 Stepwise regression of the effect of time on task and grammatical understanding for writing scores

Dependent: writing	Adjusted *R* square	B	Standard error	Beta	Significance of *t*
(Constant)		.835			< .001
Grammatical understanding	.314	.852	4.90	.449	< .001
Time on task	.381	.100	4.62	.281	< .001

GU. I excluded texts longer than 400 words ($N = 13$) from the analysis due to failure in following the word count instructions (200–250 words).

Time on task was measured automatically as the time range between opening and closing the online test window, in which the students wrote both ATs and RTs; it therefore does not reveal the actual typing time. The correlation between time on task and writing score was positive and rather strong ($r = .475$, $p < .001$, $N = 5512$). Kauppinen and Marjanen (2020, 100) have also observed that after a certain point, the benefit of spending more time decreases. Durations longer than 105 minutes were excluded. According to a regression analysis (Table 6), Time on task has a small explanatory power when GU has been controlled.

In general, writing time therefore indicates the effort that the students put on the task, which often resulted in higher scores. The positive correlation between time on task and GU (see Table 4) could either suggest that they both are affected by attitudes, motivation, or some personality trait, such as conscientiousness, which would lead the student to work more diligently. On the other hand, time on task had a minor negative correlation with doing homework ($r = -.077$, $p < .001$), which counters this explanation, even though the smallness of the correlation indicates that general diligence had no significant effect on the time the student spent writing. Perhaps a more accurate interpretation would be that of Fontich's (2016), that a student with high metalinguistic awareness is seldom satisfied with first draft but is able to shape and reshape the text, which is time-consuming but usually improves text quality. Metalinguistic awareness also broadens the selection of linguistic structures among which the writer can choose (Myhill, 2012), and this kind of choice-making can take time. The data did not include key logs, which would have provided information about the editing processes. Nonetheless, the difference between editing and non-editing writers is illustrated in the following examples, drawn from student RTs. The writer A both planned (a little)

Table 7 Some further text characteristics and their bivariate correlations with AT and grammatical understanding (GU) ($N = 138$)

N = 138	Word count	Sentence count	Word length		Sentence length	
			M	SD	M	SD
Min	3	1	3.80	2.13	3.00	1.00
Max	329	33	8.00	20.8	41.00	21.93
Mean	181.96	14.43	6.588	3.279[2]	13.40	5.965[2]
SD	64.137	6.170	0.586[1]	1.554	4.7615[1]	2.824
r with AT	.587**	.589**	.358**	.215*	-.237**	-.155
r with GU	.323**	.373**	.255**	.100	-.294**	-.217*

** $p < .01$ * $p < .05$ (two-tailed). [1]The standard deviation of the mean word/sentence length, i.e., how much the mean word/sentence length varies among the sample of ATs. [2]The mean of the standard deviations of word/sentence lengths.

and edited the text, resulting in an AT score of 14 (GU = 15) whereas B decided not to correct even the obvious errors (AT = 4, GU = 5).

> A: Usually I would plan, the text on paper, but now it was not possible, so I thought a little about the content and started to write straight onto the screen. Afterwards I edited some clause structures and made additions.[2]
> B: I wrote text straight from my head with *what I came up with I did not plan the text at all I just wrote what crossed my mind. I read though my *text once and noticed multiple *errors in writing, but I decided that I won't stay to correct even them.[3]

The examples above also portray how different levels of GU can be seen in student texts. A, who has high GU, is seemingly able to write better in terms of all assessment criteria, but especially when it comes to SPaG and sentence structures. Next, I will examine the differences between high and low GU students in more detail.

Syntactic features of argumentative texts: a mixed-methods analysis of subsample

In this section, I will focus on the linguistic features of a subsample that consists of 138 student ATs. Table 7 shows the basic characteristics of these texts for an analytical starting point. Perhaps surprisingly, the average word length of student ATs was slightly higher than in Finnish written language in general, in which it varies

[2] The Finnish original: "Yleensä suunnittelisin, tekstiä paperilla, mutta nyt se ei ollut mahdollista, joten mietin hieman sisältöä ja aloin kirjoittaa suoraan ruudulle. Jälkikäteen muokkasin joitain lauserakenteita ja tein lisäyksiä." Misspellings are indicated with an asterisk in the English translations.

[3] "Kirjoitin tekstiä suoraan päästä sillä mit keksin en suunnitellut tekstiä yhtään kirjoitin vain mitä mieleen tuli. Luin tekstni kerran läpi ja huomasin useamman kirjoitusvirhee, mutta päätin etten edes niitä jää korjaamaan."

between 7.4 and 7.9 characters per word (Pajunen & Palomäki, 1984). This might be explained by the argumentative genre as well as the societal nature of the preassigned topics (such as environmental protection), in which a formal register with longer words was expectable. The average sentence length in student ATs can be considered typical in Finnish written language (A. Hakulinen, Karlsson, & Vilkuna, 1996; Salmi, 2010).

Finnish language is highly inflectional and rich in the use of suffixes (Karlsson 2008). From this viewpoint it is obvious that word length was related to both AT quality and GU. According to Pajunen and Vainio (2021), average word length increases according to age from primary school through early adulthood. The increase is related to vocabulary development, specification of meanings, inflectional complexity, and the developing preference of written language forms over spoken ones (Pajunen, 2012; Pajunen & Vainio, 2021). For example, older students use more complex derivatives. On the other hand, more mature writers tend to use conjunctives and anaphoric words to increase textual cohesion, and this skill decreases the mean word length in skilled writers because of the shortness of those words (Pajunen, 2012). In my subsample, the more skilled writers used more complex words like *kierrättäisimme* ('we would recycle'), *historiallisestikin* ('even historically') and *selvittämättä* ('without investigating'), all of which consist of four morphemes, whereas the vocabulary was simpler in less skilled writers. The significant contribution of vocabulary on literacy skills has been observed previously in different linguistic contexts (e.g. Carlisle, 2000; Kusnetsoff, 2017; Wagner et al. 2007).

Word length can also be increased by the use of compound words, as the Finnish orthography often prefers closed compounds over open ones (e.g. *vaellusreitti* 'hiking trail', cf. Hyvärinen, 2019). Thus, the ability to write correctly closed compounds might result in longer words and therefore explain some of the effect of word length on AT quality. According to Pajunen (2012), year 6 students and younger write compounds rather seldom; partly because compounds tend to be used less in spoken than in written language, the conventions of which the youngest students are still quite unfamiliar with. The use of compound words increases by age, but non-standard writing remains common: as many as 40% of the best-achieving group of students in matriculation examination (end of year 12) made compound word errors, and in the lowest-achieving group, everyone made them (Lyytikäinen, 2000). In Lauri's study (2018), secondary-level students struggled with compound word norms in 38.5% of the instances, with closed compounds being more difficult (error rate 39.5%) than open compounds (28.5%). In my own subsample, 34.7% of the students made at least one closed compound error[4], whereas 32.6% of them had written open compounds as closed ones[5]. Either one or both types of errors were made by

[4] Examples of incorrectly open compounds: *saastuttamis tavoista ('about means of pollution'), *paperi työt ('paper work')

[5] Examples of incorrectly closed compounds: *kokoajan ('all the time'), *niinkuin ('like', 'as'), *15-vuotta ('15 years'), *seksuaalisestasuuntautumisesta ('of a sexual orientation'), *köyhäperhe ('a poor family')

52.7%. The incidence of compound word errors was related to neither AT quality ($r = -.017$, $p = .879$, $N = 138$) nor GU ($r = .017$, $p = .887$, $N = 133$). Therefore, the effect of word length on AT quality is not explained by the student's ability to follow Finnish compound word norms.

Sentence length and AT quality correlated negatively; weaker writers tended to write longer sentences. A comparison can be made to a Dutch analysis (van Rijt et al. 2021) which examined the syntactic characteristics of ATs written by year 10 students ($N = 125$, average age 15.5). The Finnish students wrote less wordy sentences (M = 13.40, SD = 4.76) than their Dutch peers (M = 20.33, SD = 5.96, van Rijt et al., 2021). The difference is explained by language typologies, as the Finnish language tends to use inflectional endings whereas the Dutch language uses more prepositions as well as definite and indefinite articles, which (written standard) Finnish does not have. In contrast with my own results, van Rijt et al. (2021) did not find any significant effect of sentence length on text quality. Neither did Myhill (2008) in 13 to 15-year-old English-speaking students, although she observed increase in sentence length in year 10 compared to year 8 students (cf. Loban, 1976 for similar results). Lu (2011), on the other hand, observed higher mean sentence length in more skilled writers in college-level Chinese-speaking learners of ESL. What, then, could explain my contradictory results? Myhill (2008, 277) explains that weaker writers sometimes "lose control" of the syntactic structure or the meaning content of the sentence, and thus tend to stretch it unnecessarily. In my data, the longest sentences are written by weaker writers and consist of multiple subordinate and coordinate clauses that eventually compromise meaningful argumentation. The following sentence is one of the longest:

> You don't have to keep the lights on all day, but if it is still to some degree bright outside then you can easily open the curtains and let the light come in, and only then switch the lights on when it is no *more bright enough outside, as in *finland it is not usually bright.[6]

A more advanced writer might have expressed the same ideas in multiple sentences. Similar kind of 'losing control' of sentence structures appears typical in low-achieving writers: in the subsample ($N = 138$) this phenomenon was qualitatively detected in ten writers (7.24%) whose GU scores ranged from 0 to 10 (M = 4.9). These students seem to possess difficulties in syntactic understanding; they struggle to build sentences that form meaningful semantic units. On some occasions, the problem lies in adding too many subordinate or coordinate clauses (example above), whereas in others, the writer does not frame sentences with full stops:

[6] The Finnish original: "Valojakaan ei tarvitse koko päivää päällä pitää, mutta jos ulkona on vieläkin jollakin asteella kirkasta niin voi helposti avata verhot ja antaa valon tulla sisään, ja vasta pistää valot päälle kun ulkona ei enään ole tarpeeksi kirkasta, niin kuin suomessa ei yleensä ole kirkasta."

> Usually many people throw garbage and all *whatever else in the *nature, many people are not able to value the nature and its wonders for example how *all the forests have been born --[7]

The observation is in accordance with previous studies. For example, Juvonen (2010) has noted that certain rhetorical structures are more prevalent in skillful writers whereas less proficient writers tend to use less established and more ambiguous syntax. It thus seems that lower GU students have more difficulties in controlling and framing sentences, which leads to unclear writing as seen in the example above. In contrast, the following example is produced by a sophisticated writer (AT = 14, KAL = 14) who builds complex sentences but also uses short ones to create effect (cf. Myhill, 2008):

> The children, whose parents do not support them need, someone who does, and if this *kind of a person does not exist, it might be difficult for the child to concentrate on school or the relationships with friends. For these reasons the children who are *born into bad conditions will stay in bad conditions.This has to change.[8]

However, variation in sentence length had no effect on AT quality, which means that writers of all levels can possibly (make) use (of) shorter sentences. This corresponds to previous findings (e.g. van Rijt et al., 2021). The small negative correlation between GU and the standard deviation of sentence length implies that students with higher GU tended to vary their sentence lengths less than the others. This is mainly explained by the long, incoherent sentences written by low GU writers. The sophisticated ability to create rhythm by varying sentence lengths is thus not seen in the quantitative analysis, even though it can be qualitatively traced.

Discussion

The study explored the relationship of grammatical understanding (GU) and writing skills in a large and comprehensive cross-sectional data of Finnish year 9 students (N = 6,044). The connection emerged as positive, approximately linear, and rather strong. Among the examined predictors of writing skills, GU was the most significant one. It was related to all assessment criteria of writing, and the correlation was at its strongest in SPaG and text structure. These findings, along with the results of additional mixed-methods analyses, can be interpreted in support of the assumption that the literacy-related benefits of grammar teaching lie in the development of metalinguistic awareness, which in turn helps the student to analyze and control language use as they learn “go meta” (Myhill et al.,

[7] ”Yleensä moni ihminen heittää roskia ja kaikkea tiesmuuta luontoo, monet ihmiset eivät osaa arvostaa luontoa ja sen ihmeitä esimerkiksi miten kiakki metsät ovat syntyneet --”

[8] ”Lapset, joiden vanhemmat eivät tue heitä tarvitsevat, jonkun joka tukee ja jos tälläistä henkilöä ei ole niin lapselle voi olla vaikeaa jaksaa keskittyä koulunkäyntiin tai ystävyyssuhteisiin. Näistä syistä usein lapset jotka syntyät huonoihin oloihin jäävät huonoihin oloihin.Tähän pitää tulla muutos.”

2020) when writing their own texts. Some students with low GU indicated shaky understanding about sentence formation, which resulted in structures that followed the norms of speech rather than writing. In addition, grammatical understanding was related to following genre-specific written language conventions, but this relationship need not be interpreted solely in terms of avoiding errors but also as a sign of understanding how linguistic choices can be made to create the desired meanings. This conclusion is supported by the data observation that the relationship between GU and word length was not explained by compound word errors. Instead, the high GU students tended to use more complex derivative words, which is not only a sign of broader vocabulary but also of morphological understanding of the complex ways of Finnish word formation (Pajunen & Honko, 2021). Grammar therefore contributes to understanding of language both *as structure* and *as choice* (Carter & McCarthy, 2006), as it is related to a growing repertoire of linguistic forms, from which the writer can choose.

Future studies could add grammar teaching methods and approaches, as reported by teachers (e.g. Harjunen & Rautopuro, 2015), as variables in the regression analyses. This could reveal differences in both grammar and writing skills depending on the teachers' pedagogical approaches. Finnish teachers have a high degree of autonomy, and not all of them habitually connect grammar with teaching of reading and writing (cf. Harjunen & Rautopuro, 2015). Their work could benefit from the numerous pedagogical rationales and practices of contextualized grammar teaching, developed in the field of educational linguistics: for example, grammar instructional sequences (e.g. Fontich, 2016), LEAD principles (Myhill et al., 2020), fostering a reflective attitude in discussing language (Wijnands et al., 2021), as well as the methods of languaging and visualizing in building conceptual metalinguistic knowledge (Rättyä, 2013). Large-scale teacher surveys, such as the ones conducted by FINEEC, could aim at mapping the prevalence of these kinds of pedagogical approaches, as well as the more traditional ones, such as doing identification and categorization exercises in textbooks.

Even though an elaborate analysis of the reflective texts (RTs) was beyond the scope of this study, I observed some differences in the practices of process writing according to GU scores and text quality. This supports Fontich's (2016) observation that metalinguistic understanding is seen in the ability to shape and reshape texts. However, as Calil and Myhill (2020) point out, a finished piece of writing does not present sufficient information about the writer's metalinguistic decision-making, and other types of studies (e.g., ones applying textual genetics) are needed to reveal more about the writing process. Another possibility could be to analyze what the students reported about their editing processes in their RTs. It is also important to note that the characteristics of a good text differ according to genres, topics, disciplines (Crossley, 2020; van Rijt et al, 2021); what is rhetorically effective in an argumentative text may not be ideal in, for example, a holiday narrative. Also, the use of vocabulary is genre-specific already in school-age children (Olinghouse & Wilson, 2013). The FINEEC data could be used to analyze and compare the genre-specific structures of student ATs and RTs and to compare the structures written by students of different grammar levels.

Limitations

The natural limitation of a cross-sectional study is its inability to examine causation. Thus, it cannot be concluded that learning grammar *enhances* writing skills, or vice versa, or that they are both caused by other factors, such as general cognitive ability or the level of conscientiousness that steers the amount of effort put in the task. The study was, however, able to control for important factors such as family background, learning attitudes, reading habits, and time on task, and conclude that GU remains a strong predictor of writing skills. Additionally, the study findings do not encompass all the possible ways in which GU can affect writing skills, but they do shed light on some of them.

Conclusions

This study has brought a novel viewpoint to the long-debated question about grammar and writing by comparing these sets of skills, as well as other variables, in a large group of students in the linguistic and educational context of Finnish and Finland. The quantitative analyses are also elaborated qualitatively. Previous studies with a similar setting have been rare, and there has been especially little information about the correlations of a functionally oriented grammar teaching, which is required by the Finnish curriculum, and writing skills.

The findings may be interpreted in support of Costa's (2019) view, according to which the relationship between language use, linguistic development (e.g. writing skills), and metalinguistic reflection is triangular. Learning to write requires explicit attention to language, and this attention-paying is facilitated by teaching of grammar as it provides necessary conceptual tools and teacher-modeled practice for metalinguistic reflection. Thus, as the students learn to understand how language works and to use linguistic concepts (in a wider sense than mere terminology)—in other words, become more aware of language—they become more aware of what they are writing and able to control it. The increased mastery and enjoyment of language can also encourage the students to read and write more, which gives them more practice and thus further develops their skills in both understanding language and using it.

Acknowledgements The author wants to thank the Finnish Education Evaluation Centre (FINEEC) and its Counsellors of Evaluation for granting the access to the national evaluation data as well as for guidance in its use.

Funding Open Access funding provided by University of Helsinki including Helsinki University Central Hospital. This work was supported by the Emil Aaltonen Foundation [Grant Number 210133].

Declarations

Conflict of interest None

References

Andrews, R., Torgerson, C., Beverton, S., Locke, T., Low, G., Robinson, A., & Zhu, D. (2004). *The effect of grammar teaching (syntax) in English on 5 to 16 year olds' accuracy and quality in written composition*. University of York, UK.

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research, 80*(2), 144–179. https://doi.org/10.3102/0034654309359353

Calil, E., & Myhill, D. (2020). Dialogue, erasure and spontaneous comments during textual composition: What students' metalinguistic talk reveals about newly-literate writers' understanding of revision. *Linguistics and Education*. https://doi.org/10.1016/J.LINGED.2020.100875

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal, 12*(3), 169–190.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English: a comprehensive guide*. Cambridge University Press.

Chatterjee, A., & Halder, S. (2022). Teaching grammar in the context of writing: A critical review. *Journal of Education*. https://doi.org/10.1177/0022057422107430

Chen, H., Lewis, H., & Myhill, D. (2021). Fostering critical reasoning: Developing argumentative competence in early and middle primary years. *Australian Journal of Language and Literacy, 44*(2), 46–61.

Chipere, N. (2003). *Understanding complex sentences. Native speaker variation in syntactic competence.* Palgrave Macmillan.

Collins, G., & Norris, J. (2017). Written language performance following embedded grammar instruction. *Reading Horizons: A Journal of Literacy and Language Arts, 56*(3), 16–30.

Cope, B., & Kalantzis, M. (Eds.). (2000). *Multiliteracies: literacy learning and the design of social futures*. Taylor: Routledge.

Costa, A. L. (2019). A contribution to the implicit/explicit debate on grammar learning: the case of contrast connectors. *Crossroads A Journal of English Studies, 24*(1), 45–64. https://doi.org/10.15290/CR.2019.24.1.04

Crossley, S. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research, 11*(3), 415–443. https://doi.org/10.17239/jowr-2020.11.03.01

Derewianka, B. (2012). Knowledge about language in the Australian curriculum: English. *Australian Journal of Language and Literacy, 35*(2), 127–146.

Derewianka, B., & Jones, P. (2010). From traditional grammar to functional grammar: bridging the divide. *Faculty of Social Sciences Papers., 10*, 1001.

Ellis, N. (2008). Implicit and explicit knowledge about language. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education* (2nd ed., pp. 110–132). London: Springer.

Fearn, L., & Farnan, N. (2007). When is a verb? Using functional grammar to teach writing. *Journal of Basic Writing, 26*(1), 63–87.

Fontich, X. (2016). L1 grammar instruction and writing: Metalinguistic activity as a teaching and research focus. *Language and Linguistics Compass, 10*(5), 238–254.

Fontich, X., & Camps, A. (2014). Towards a rationale for research into grammar teaching in schools. *Research Papers in Education, 29*(5), 598–625.

Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia, 60*(2), 183–208.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476.

Hakulinen, A., Karlsson, F., & Vilkuna, M. (1996). *Suomen tekstilauseiden piirteitä: kvantitatiivinen tutkimus*. University of Helsinki, Helsinki.

Hakulinen, L. (1925). Äidinkielen kieliopin opetuksesta ala-asteilla. *Virittäjä, 29*, 31–39.

Halliday, M.A.K. & Matthiessen, C. (2004). *An introduction to functional grammar*. Arnold.
Hancock, C. (2009). How linguistics can inform the teaching of writing. In R. Beard, D. Myhill, J. Riley, & M. Nystrand (Eds.), *The SAGE handbook of writing development* (pp. 194–207). SAGE.
Hankala, M., Kauppinen, M., & Kulju, P. (2015). Tutkimuksesta tukea kirjoittamisen opettamiseen. In E. Harjunen (Ed.), *Tekstit puntarissa: ajatuksia äidinkielen ja kirjallisuuden oppimistuloksista perusopetuksen päättöarvioinnissa 2014 ja 2010* (pp. 73–84). FINEEC.
Harjunen, E., Marjanen, J. & Karlsson, J. (2019). *Äidinkielen pieni pitkittäisarviointi 2014–2017. Perusopetuksen päätöstä lukion päättöön*. FINEEC.
Harjunen, E. & Rautopuro, J. (2015). *Kielenkäytön ajattelua ja ajattelun kielentämistä. Äidinkielen ja kirjallisuuden oppimistulokset perusopetuksen päättövaiheessa*. FINEEC.
Hellgren, J. & Kauppinen, M. (Eds.) (2020). *Kirjoitan, siis ajattelen: näkökulmia kirjoittamisen opetukseen. Jag skriver, alltså tänker jag: perspektiv på skrivpedagogik*. FINEEC.
Hillocks, G. (2006). Two Decades of Research on Teaching Writing in the Secondary Schools in the US. *L1 Educational Studies in Language and Literature*. *06*(2): 29–51
Hudson, R. (2015). Grammar instruction. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 288–300). Guilford.
Hudson, R. (2004). Why education needs linguistics? *Journal of Linguistics, 40*(1), 105–130.
Hyvärinen, I. (2019). Compounds and multi-word expressions in Finnish. In B. Schlücker (Ed.), *Complex Lexical Units* (pp. 307–336). De Gruyter.
Jakobs, E.-M. Perrin, D. (Eds.) (2014). *Handbook of writing and text production*, De Gruyter.
Jones, S., Myhill, D., & Bailey, T. (2013). Grammar for writing? An investigation of the effects of contextualised grammar teaching on students' writing. *Reading and Writing, 26*(8), 1241–1263.
Juvonen, R. (2010). Evaluoiva että-yhdyslause ja retoriset rakenteet. *Virittäjä, 114*(1), 39–70.
Kalliokoski, J., Kumenius, J., Luukka, M.-R., Mustaparta, A.-K., Nissilä, L., & Tuomi, M. (2015). Kielitiedon opettamisesta. In A.-K. Mustaparta (Ed.), *Kieli koulun ytimessä, Näkökulmia kielikasvatukseen* (pp. 40–54). Finnish National Agency for Education.
Karlsson, F. (2008). *Finnish: an essential grammar* (Transl. A. Chesterman). Routledge.
Kauppinen, M. & Marjanen, J. (2020). *Millaista on yhdeksäsluokkalaisten kielellinen osaaminen?* FINEEC.
Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., Wade-Woolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing, 25*(2), 389–410. https://doi.org/10.1007/s11145-010-9276-5
Koskinen, I. (1988). Hyvästit kieliopille. In *Äidinkielen kielioppi suomenkielisessä oppikoulussa ja kansakoulussa vuoden 1843 koulujärjestyksestä peruskoulu-uudistukseen*. University of Helsinki, Department of Teacher Education
Koskinen, I. (1990). *"Kieliopin" paluu. Äidinkielen ns. kielentuntemuksen opetus suomenkielisessä peruskoulussa*. University of Helsinki, Department of Teacher Education
Koster, M., Tribushinina, E., de Jong, P. F., & van den Bergh, H. (2015). Teaching children to write: A meta-analysis of writing intervention research. *Journal of Writing Research, 7*(2), 249–274.
Kusnetsoff, T. (2017). *Suomenkielisten nuorten morfologinen tietoisuus ja sen yhteys luetun ymmärtämiseen*. Master's thesis, University of Tampere.
Lappalainen, H.-P. (2004). *Kerroin kaiken tietämäni: perusopetuksen äidinkielen ja kirjallisuuden oppimistulosten arviointi 9. vuosiluokalla 2003*. Finnish National Agency for Education.
Lauri, J. (2018). *Ammattikoululaisten ja lukiolaisten yhdyssanaosaaminen osittaissanelutestillä mitattuna*. Master's thesis, University of Oulu.
Leino, J., & Sääskilahti, M. (2010). Kieliopin keinoin asenteita analysoimaan. In T. Törmälehto (Ed.), *Maailma tuli tunnille: kansainvälisyyskasvatus osana äidinkielen ja kirjallisuuden opetusta* (pp. 71–84). Liitto: Äidinkielen opettajain liitto.
Loban, W. (1976). *Language development: kindergarten through grade twelve*. Champaign: National Council of Teachers of English.
Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36–62.
Lyytikäinen, E. (2000). Ylioppilasaineiden kielivirheet - lisääntymän päin? *Virittäjä, 4*, 602–609.
Marjanen, J. (2020). *Suomen kielen ja kirjallisuuden perusopetuksen oppimistulosarviointi. Menetelmäliite*. Finnish Education Evaluation Centre.
Matruglio, E. (2020). What two teachers took up: metalanguage, pedagogy and potentials for long-term change. *Language and education, 35*(5), 463–478. https://doi.org/10.1080/09500782.2020.1825477

Mullis, I.V.S., Martin, M.O., Foy, P. & Drucker, K.T. (2012). *PIRLS 2011: International results in reading*. International Association for the Evaluation of Educational Achievement.
Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. T. (2017). *PIRLS 2016: International results in reading*. International Association for the Evaluation of Educational Achievement.
Myhill, D. (2008). Towards a linguistic model of sentence development in writing. *Language and education, 22*(5), 271–288. https://doi.org/10.1080/09500780802152655
Myhill, D. (2012). The ordeal of deliberate choice: Metalinguistic development in secondary writers. In V. W. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 247–272). Psychology Press.
Myhill, D. (2021). Grammar re-imagined: foregrounding understanding of language choice in writing. *English in Education, 55*(3), 265–278. https://doi.org/10.1080/04250494.2021.1885975
Myhill, D., Jones, S., Lines, H., & Watson, A. (2012). Re-thinking grammar: the impact of embedded grammar teaching on students' writing and students' metalinguistic understanding. *Research Papers in Education, 27*(2), 139–166. https://doi.org/10.1080/02671522.2011.637640
Myhill, D., Watson, A., & Newman, R. (2020). Thinking differently about grammar and metalinguistic understanding in writing. *Bellaterra Journal of Teaching & Learning Language & Literature, 13*(2), 870. https://doi.org/10.5565/rev/jtl3.870
NCC. (2016). *National core curriculum for basic education 2014*. Finnish National Agency for Education.
New London Group 1996 A pedagogy of multiliteracies: designing social futures. *Harvard Educational Review. 66*(1): 60–92
Nupponen, A.M., Jeskanen, S. Rättyä, K. (2019). Finnish student language teachers reflecting on linguistic concepts related to sentence structures: Students recognising linguistic concepts in L1 and L2 textbooks. *L1 Educational Studies in Language and Literature*. *19*(1), 1–25.
OECD (2003). *Literacy skills for the world of tomorrow. Further Results from PISA 2000*.
OECD (2004). *Learning for tomorrow's world. First Results from PISA 2003*.
OECD (2007). *PISA 2006*.
OECD (2010). *PISA 2009 Results: what students know and can do*.
OECD (2014). *PISA 2012 Results: what students know and can do (Volume I)*
OECD (2016). *PISA 2015 Results (Volume I)*.
OECD (2019). *PISA 2018 Results (Volume I)*.
Olinghouse, N. G., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading and Writing, 26*(1), 45–65. https://doi.org/10.1007/s11145-012-9392-5
Pajunen, A. (2012). Kirjoittamistaitojen kehitys 8–12-vuotiailla. Alakoululaisten unelmakirjoitelmat. *Virittäjä*, *116*(1), 4–32
Pajunen, A., & Honko, M. (Eds.). (2021). *Suomen kielen hallinta ja sen kehitys*. Peruskoululaiset ja nuoret aikuiset.
Pajunen, A. & Palomäki, U. (1984). *Tilastotietoa suomen kielen rakenteesta 1*. Kotimaisten kielten tutkimuskeskus.
Pajunen, A. Vainio, S. (2021). Kielen rakenteiden hallinta peruskoululaisilla ja nuorilla aikuisilla. In A. Pajunen & M. Honko (Eds.), *Suomen kielen hallinta ja sen kehitys. Peruskoululaiset ja nuoret aikuiset* (pp. 367–420). Finnish Literature Society SKS.
Pallaskallio, R. (2006). Georgialaisnaiset median pyörityksessä: tempukset uutistekstin analysoinnin apuna. In M. Harmanen & M. Siiroinen (Eds.), *Kielioppi koulussa. Äidinkielen opettajain liiton vuosikirja XLX* (pp. 137–154). Äidinkielen opettajain liitto.
Paukkunen, U.-M. (2011). *Lauseiden virrassa. Peruskoulun yhdeksäsluokkalaiset lauseiden tulkitsijoina*. University of Oulu: Oulu.
Rättyä, K. (2013). Languaging and visualisation method for grammar teaching: a conceptual change theory perspective. *English Teaching: Practice and Critique, 12*(3), 87–101.
Rättyä, K. (2017). *Kielitiedon didaktiikkaa. Kielentäminen ja visualisointi sanaluokkien ja lauseenjäsenten opetusmenetelminä*. University of Helsinki: Helsinki.
Rättyä K, Awramiuk E, Fontich, X. (2019). What is grammar in L1 education today? *L1 Educational Studies in Language and Literature*. 19: 1–8.
Rättyä, K., & Kulju, P. (2018). Kielitietoisuutta kielentämällä–kieliopin tehtävätyyppien kehittäminen. *Ainedidaktiikka, 2*(1), 59–74.
Routarinne, S., & Absetz, P. (2013). Writing performance and writing self-efficacy: the case of Finnish fifth graders. In S. Tavares (Ed.), *Why do we write as we write?* (pp. 51–66). Inter-Disciplinary Press.

Salmi, N. (2010). *Virkkeiden ja lauseiden piirteitä edistyneiden suomenoppijoiden teksteissä* – Master's thesis, University of Turku.

Savolainen, K. (1998). *Kieli ja sen käyttäjä äidinkielen oppikirjasarjan tuottamana*, Joensuun yliopisto.

Schleppegrell, M., & Christie, F. (2018). Linguistic features of writing development: a functional perspective. In C. Bazerman, A. N. Applebee, V. W. Berninger, D. Brandt, S. Graham, J. V. Jeffery, P. K. Matsuda, S. Murphy, D. W. Rowe, M. Schleppegrell, & K. C. Wilcox (Eds.), *The lifespan development of writing*. National Council of Teachers of English. pp. 111–150

Schleppegrell, M. (2010). Supporting a 'reading to write' pedagogy with functional grammar. In C Coffin (Ed), *Language support in EAL contexts Why systemic functional linguistics? (Special Issue of NALDIC Quartelry)*. NALDIC. pp. 26–31

Sinnemäki, K., & Saarikivi, J. (2019). Sacred language: reformation, nationalism, and linguistic culture. In K. Sinnemäki, A. Portman, J. Tilli, & R. H. Nelson (Eds.), *On the legacy of lutheranism in Finland: societal perspectives* (pp. 39–68). Finnish Literature Society SKS.

Svalberg, A.M.-L. (2012). Language awareness in language learning and teaching: A research agenda. *Language Teaching, 45*(3), 376–388. https://doi.org/10.1017/S0261444812000079

Tainio, L. (2020). Kielitieto. In L. Tainio, M. Ahlholm, S. Grünthal, S. Happonen, R. Juvonen, U. Karvonen, & S. Routarinne (Eds.), *Suomen kieli ja kirjallisuus koulussa* (pp. 81–123). Suomen ainedidaktinen tutkimusseura.

Tomasello, M. (2003). *Constructing a language*. Harvard University Press.

Unsworth, L. (2001). *Teaching multiliteracies across the curriculum Changing contexts of text and image in classroom practice*. Open University Press.

Varis, M. (2012). *Kielikäsitys yläkoulun äidinkielen oppikirjoissa*. Oulu: University of Oulu.

van Rijt, J., van den Broek, B., & de Maeyer, S. (2021). Syntactic predictors for text quality in Dutch upper-secondary school students' L1 argumentative writing. *Reading and Writing, 34*, 449–465. https://doi.org/10.1007/s11145-020-10079-5

Wagner, R. K., Muse, A. E., & Tannenbaum, K. R. (Eds.). (2007). *Vocabulary acquisition. Implications for reading comprehension*. Guilford.

Watson, A. M., Newman, R. M. C., & Morgan, S. D. (2021). Metatalk and metalinguistic knowledge: the interplay of procedural and declarative knowledge in the classroom discourse of first-language grammar teaching. *Language Awareness, 30*(3), 257–275. https://doi.org/10.1080/09658416.2021.1905655

Wijnands, A., van Rijt, J., & Coppen, P. A. (2021). Learning to think about language step by step: a pedagogical template for the development of cognitive and reflective thinking skills in L1 grammar education. *Language Awareness, 30*(4), 317–335. https://doi.org/10.1080/09658416.2021.1871911

Wyse, D. (2001). Grammar for writing? A critical review of empirical evidence. *British Journal of Educational Studies, 49*(4), 411–427. https://doi.org/10.1111/1467-8527.t01-1-00185

Wyse, D., & Torgerson, C. (2017). Experimental trials and 'what works?' in education: the case of grammar for writing. *British Educational Research Journal, 43*(6), 1019–1047. https://doi.org/10.1002/berj.3315