A Statistical Model of Learning Descriptive Grammar[*]

Richard Madsen

Abstract

The study presented here describes the development of a statistical model of the learning of descriptive grammar by students of English. The purposes of the model are: identifying the areas of descriptive grammar, such as recognising parts of speech, clause constituents, phrase constituents, etc., whose knowledge best predicts success in learning descriptive grammar, and determining whether there is an implicational relationship between areas of descriptive grammar such that the learning of one area depends on having learnt another area. The motivation for the second part of the study is to assess assumptions that textbooks and teachers seem to have concerning the best order in which to teach the various areas of descriptive grammar, implying that some of the fields of grammar are more basic than others and therefore need to be learnt before the other fields can be acquired.

Keywords: *language acquisition, learning descriptive grammar, statistical model*

1. Introduction

The purpose of the study presented here is to develop a statistical model that describes the learning of descriptive grammar by students of English whose mother tongue is not English. Descriptive grammar is understood as the knowledge of grammatical phenomena as defined and described by theories of grammar, that is the ability to recognise parts of speech, clause constituents, phrase constituents, etc.

The motivation for developing such a statistical model is twofold. One reason is to identify the areas of descriptive grammar the knowledge of which best predicts success in learning descriptive grammar. For students of English Business Communication at Aalborg University, Denmark, success in learning descriptive grammar is measured in terms of the results of the grammar exam. The goal of identifying such predictors is the ability to detect students early in the grammar course who are in danger of failing the exam, as inspired by the works of Elbro and Scarborough (2003) and Lyon and Moats (1997).

The other reason is to ascertain whether there is an implicational relation between certain areas of descriptive grammar. The term 'implicational relation' describes the situation where the knowledge of one area depends on the knowledge of another area, i.e. it is necessary to have learnt area A, say parts of speech, before area B, say clause constituents, can be learnt. The motivation for searching for implicational relations between areas of descriptive grammar

[*] My gratitude goes to my supervisor, Kim Ebensgaard Jensen, to those participants of CECIL'S 4 who viewed my poster presentation, and to the anonymous reviewer of my paper for their invaluable comments and suggestions.

is to test assumptions that textbooks and teachers seem to have concerning the best way of sequencing the various areas of descriptive grammar, implying that some of the fields of grammar are more basic than others and therefore need to be learnt before the other fields can be acquired. For example, Hjulmand and Schwarz (2009), whose textbook has been used for years in my department, seem to imply by the very order of the chapters that learning parts of speech is more elementary than learning clause constituents. One of the purposes of this study is thus to test whether there is any basis in acquisition patterns for such assumptions.

To sum up, here are the two research questions that the present study seeks to answer:

1) What are the best predictors of success in learning descriptive grammar?
2) Is there an implicational relationship between areas of descriptive grammar?

2. Theory

The present paper fits within the general framework of the study of second language acquisition. However, it does not have the same focus as other studies within SLA tend to have since it does not deal with the students' command of the English language in communication, but with their mastery of the concepts of theoretical grammar.

Thus, this study does not concern itself with the students' ability to use grammar in actual communication, i.e. the level of correctness they demonstrate in speaking or writing English. Nor does it discuss the usefulness of studying descriptive grammar for the sake of improving one's linguistic correctness as has been debated since Krashen (1981) (see also Ellis 2009). Consequently, the terms learning and acquisition are used interchangeably in this study. For a discussion of the relationship between a knowledge of descriptive grammar and the level of grammatical correctness demonstrated in writing English, see Madsen (2014).

However, no matter what the relationship between knowing descriptive grammar and being able to communicate using grammatically correct language might be, it remains a fact that students studying a language at a university must study and learn the concepts of descriptive grammar. Imparting the nomenclature of descriptive grammar to students has been a standard part of the curriculum of university language studies for centuries. Therefore, it is not uninteresting or irrelevant to investigate how such theoretical concepts are acquired. I am not aware of any paper that has explored this aspect of SLA.

Having a different focus from conventional SLA studies notwithstanding, this paper does use concepts that are well-established in mainstream SLA research. Most notable among these concepts is the assumption that a language cannot be acquired in any order. No matter whether one subscribes to Krashen's (1981) dismissal of teaching descriptive grammar with the purpose of facilitating language acquisition or not, there is a general agreement (Clahsen, Meisel and Pienemann 1983; Ellis 2012; Lund 1997; Pienemann 1998) that the acquisition of a language follows a certain order, although there is considerable disagreement as to the precise order. It effectively means that certain elements in a language must be internalised by the language learner before other elements can be acquired. This is the very same notion as the one underlying the hypothesis about an implicational relation between certain concepts of

descriptive grammar, which is one of the research objectives of this paper. Besides being anchored in SLA research, this paper draws heavily on inferential statistics in order to answer the research questions. The statistical calculations are shown in the section on methodology.

3. Methodology and data

In this section, the data that have been used for the analyses are first described, and then the statistical calculations that were used to answer the research questions outlined in the introduction are presented.

3.1. Description of the data

The model presented in this paper is constructed on the basis of an analysis of the performance in a grammar exam of altogether 323 freshmen from the academic years 2009-2013 in my department, which is the department of English Business Communication within the School of Culture and Global Studies at Aalborg University, Denmark. The grammar exam, to be passed at the end of the first semester, tests the students' knowledge of up to 14 different areas of descriptive grammar out of a set of altogether 17 areas. Henceforth, for the sake of brevity, the knowledge of a certain area or field of descriptive grammar will be referred to as a skill.

The exam has always consisted of 100 questions, the vast majority of which belong to the gap-filling or multiple-choice type. Every year the students have been given 120 minutes to complete the exam and have never been allowed to use any aid during the exam. However, as hinted at in the previous paragraph, there has been some variation in the exact form and scope of the exam through the years. Here are some illustrative examples of questions from the grammar exam:

(1) a. Determine which part of speech the underlined word belongs to.
This exam features John Maynard Keynes, who is one of the most famous economists.

b. Determine the semantic relation between the expressions below.
Linguistics vs. science

c. Divide the word below into root and affixes, and describe each morpheme.
Invariably

d. Determine what clause constituent the underlined sequence of words is.
This exam features John Maynard Keynes, who is one of the most famous economists.

e. Decide whether the underlined sequence of words is a phrase or clause.
Having attended Eton College, Keynes studied at Cambridge University.

f. Determine which phrase type the sequence of words below belongs to.
most famous

g. Determine what phrase constituent the underlined sequence of words is.
someone special

h. Determine what kind of pronoun the underlined word is.
I wonder who has developed the concept of money.

i. Determine the type of the underlined subclause.
Keynes married a woman who had a low social background.

j. Determine how many matrix clauses the period below consists of.
However well you think you know someone, you might still be surprised.

k. Relative clauses: Are commas necessary with the relative clause below?
Statistics(,) which is the study of probability and distribution(,) is an important part of economics.

The various skills have always had a varying and uneven representation with respect to the number of questions that test them. The nomenclature taught has also varied somewhat over the years, most notably in 2012 when Madsen (unpublished) was used as an experimental textbook instead of Hjulmand and Schwarz (2009). Madsen (unpublished) is a much more detailed description of English than Hjulmand and Schwarz (2009). Consequently the (number of) possible answers too have varied within the same areas of grammar. Tables 1a through 1h give an overview of the terminology used during the period studied. With these details, one can easily compare the years with each other, and this research with other studies within the field.

Table 1a. *Parts of speech nomenclature*

Years	**Parts of speech**									
2009	adj	adv	art	n	prep	pron	v			
2010-2011	adj	adv	art	n	prep	pron	v	conj		
2012	adj	adv	art	n	prep	pron	v	conj	subjunction[1]	infinitive marker
2013	adj	adv	art	n	prep	pron	v	conj		

Table 1b. *Clause constituents*[2]

Years	**Clause constituents**													
2009	V	S	DO	IO	SC	OC	A	preliminary S (PS)						
2010-2011	V	S	DO	IO	SC	OC	A	PS	preliminary DO (PDO)					
2012	V	S	DO	IO	SC	OC	A	PS	PDO	conjunctional const. (CC)	oblique DO	oblique IO	oblique SC	oblique OC	agentive const.
2013	V	S	DO	IO	SC	OC	A	PS	PDO	CC					

Table 1c. *Phrase types*

Years	**Phrase types**							
2009-2011	NP	VP	AdjP	AdvP	PP			
2012	NP	VP	AdjP	AdvP	PP	Conjunctional phrase	Subjunctional phrase	Prenominal phrase
2013	NP	VP	AdjP	AdvP	PP			

Table 1d. *Subclause types*

Years	**Subclause types**		
2009-2013	nominal	adverbial	modifying[3]

[1] A portmanteau word combining ***sub****ordinating* and *con****junction***.

[2] A: adverbial constituent, DO: direct object, IO: indirect object, OC: object complement, S: subject, SC: subject complement, V: verb(al constituent)

Table 1e. *NP constituents*

Years	NP constituents			
2010-2011	determiner (det)	premodifier (prem)	head	postmodifier (pom)

Table 1f. *Phrase constituents (italicized constituents apply to NPs only)*

Years	Phrase constituents							
2012	*predeterminer*	*det*	*ordinator*	*quantifier*	*descriptor*	head	pom	degree
2013	*det*	prem	head	pom				

Table 1g. *Pronoun types*

Years	Pronoun types									
2010 -2011	demonstr.	interrog.	personal	possessive	reflexive	relative	indef.			
2012	dem.	inter.	pers.	poss.	refl.	rel.	reciprocal	indef.	negative	universal
2013	dem.	inter.	pers.	poss.	refl.	rel.	rec.	indef.		

Table 1h. *Semantic relations*

Years	Semantic relations				
2012-2013	synonymy	antonymy	homonymy	hypo-/hyperonymy	holo-/meronymy

Moreover, up until 2011, the students were provided with a list of possible answers in each task; thus, for example, in the task on clause constituents, the names of all the possible clause constituents were listed for the students' convenience. These lists were scrapped in 2012 and onwards in order to make the exam a little more challenging. However, there is still an example question-answer pair provided for each skill, and, of course, some questions necessarily and unavoidably reveal the possible answers by their very nature, for instance the questions concerning finiteness, which is a binary category.

On the other hand, up until 2011, the exam was (fine) graded whereas in 2012 and 2013 it was marked as simply pass/nonpass (fine grading was reinstated for 2014). Nonetheless, in all the years the minimum requirement for passing (for a passing grade) has been 60 correct answers out of 100, with all the questions having equal weight. Because of all the above-mentioned differences between the exams over the years, all the calculations have been made for each year separately.

3.1.1. Overview of the results of the grammar exam

Table 2 shows a summary of all the years with the skills tested, displaying the average efficacy and the standard deviation thereof in each skill together with the number of questions within the exam testing the given skill and the number of possible answers to the questions. Efficacy is calculated as the number of correct answers divided by the number of questions testing a given skill, or all the questions when computing the overall efficacy. In this way, the individual skills and the overall knowledge of descriptive grammar can be compared despite the fact that they

[3] The students are in principle required to distinguish between relative and appositive clauses; however, all the modifying clauses in the exams have been relative clauses.

are tested by different numbers of questions. The averages of individual skills in green are higher than the overall average, and the averages in red are lower. This gives a rough estimate of which tasks are more (red) or less (green) challenging for the students (Corder 1987).

Table 2. *Overview of the results of the grammar exam*

		2009	2010	2011	2012	2013
Number of students in database[4]		56	63	54	58	92
Percentage of fails		5.36	15.87	12.96	29.31	26.09
Overall average efficacy (failure: efficacy<0.60)		0.751	0.746	0.742	0.678	0.644
Standard deviation of overall efficacy		0.100	0.140	0.147	0.163	0.143
Parts of speech	Number of questions:Possible answers/question	13:7	15:8	15:8	11:10	10:8
	Standard deviation	0.153	0.175	0.146	0.192	0.190
	Average	0.76	0.71	0.81	0.73	0.74
Clause constituents	Questions:Possibilities	21:8	20:9	20:9	20:15	20:10
	Standard deviation	0.153	0.213	0.196	0.238	0.220
	Average	0.72	0.73	0.66	0.62	0.58
Phrase vs. subclause	Questions:Possibilities	12:2	10:2	10:2	10:2	10:2
	Standard deviation	0.174	0.196	0.197	0.186	0.182
	Average	0.84	0.76	0.75	0.86	0.59
Phrase types	Questions:Possibilities	15:5	10:5	10:5	10:8	10:5
	Standard deviation	0.185	0.172	0.180	0.167	0.220
	Average	0.76	0.78	0.75	0.58	0.69
Subclause types	Questions:Possibilities	12:3	7:3	7:3	7:3	7:3
	Standard deviation	0.169	0.249	0.222	0.266	0.232
	Average	0.63	0.69	0.71	0.62	0.59
Subclause finiteness	Questions:Possibilities	5:2	7:2	7:2	7:2	7:2
	Standard deviation	0.178	0.177	0.230	0.159	0.199
	Average	0.82	0.83	0.72	0.88	0.73
Number of matrix clauses in a paragraph	Questions:Possibilities	7:2	5:2	5:2	5:∞	5:∞
	Standard deviation	0.193	0.225	0.191	0.256	0.265
	Average	0.75	0.81	0.78	0.58	0.58
Comma	Questions:Possibilities	4:∞	4:2	4:2	4:2	4:2
	Standard deviation	0.222	0.239	0.250	0.187	0.208
	Average	0.71	0.69	0.78	0.88	0.79
Inserting relative pronoun	Questions:Possibilities[5]	4:6	1:6	1:6	1:6	1:6

[4] The reason for the substantial increase in the number of participants in 2013 is that I was granted access to the exam results of the students of a colleague of mine. For all the other years, only my own students were included in the database. They represented two-thirds of all the freshmen in the Department of English Business Communication in the years 2009 and 2010, and half of them in the years 2011 and 2012. In 2013, three quarters of the freshmen were included in the database.

[5] Due to the fact that all the actual sentences ever used in the exams contained only physical objects as antecedents, be they animate or inanimate, only the forms *who, whom, whose, which, that* and zero pronoun could be considered viable options. Danish does not distinguish between animate and inanimate antecedents, for which reason choosing the right relative pronoun in English tends to be a challenge for Danes. Hence the number of possible answers is not further restricted by the animacy of the antecedent (Jarvis and Pavlenko 2008; Jarvis 2011).

	Standard deviation	0.120	0.405	0.487	0.326	0.483
	Average	0.93	0.79	0.39	0.88	0.63
Finding the subclause	Questions:Possibilities	5:∞				
	Standard deviation	0.144				
	Average	0.75				
Finding the correct sentence	Questions:Possibilities	2:4				
	Standard deviation	0.267				
	Average	0.75				
NP constituents	Questions:Possibilities		10:4	10:4		
	Standard deviation		0.197	0.207		
	Average		0.76	0.85		
Phrase constituents	Questions:Possibilities				10:7	10:5
	Standard deviation				0.230	0.163
	Average				0.61	0.69
Pronoun types	Questions:Possibilities		10:7	10:7	10:10	10:10
	Standard deviation		0.260	0.252	0.299	0.256
	Average		0.76	0.73	0.68	0.69
Verb form finiteness	Questions:Possibilities[6]		1:32	1:32	1:32	1:32
	Standard deviation		0.498	0.494	0.485	0.465
	Average		0.46	0.57	0.38	0.32
Semantic relations	Questions:Possibilities				3:8	3:8
	Standard deviation				0.330	0.295
	Average				0.56	0.66
Morphological analysis	Questions:Possibilities				1:∞	2:∞
	Standard deviation				0.469	0.369
	Average				0.33	0.30

The number of possible answers to questions testing a given skill is determined by the nomenclature taught in the given year, i.e. how many different terms are distinguished within a given area of descriptive grammar (see Tables 1a through 1h). The symbol ∞ as the number of possible answers indicates an open question. The number of possible answers is not truly infinite, but very large and in any case impossible to calculate generally because it depends on the actual word or clause to be analysed, which of course varies from year to year. In a manner of speaking, the number of possible answers is indicative of the chance for the students to score a point by choosing an answer randomly.

Although no exact measurement of the phenomenon has been made, it seems that the number of empty or nonsensical answers (e.g. giving a part of speech when asked about a clause constituent) has increased since 2012 compared to previous years possibly due to the fact that the students are no longer furnished with complete lists of the possible answers. The noticeable decrease in student efficacy in 2012 and 2013 as compared with the previous years is likely to be the result of both the above mentioned measures to increase the difficulty of the exam and the lack of fine-grading, which possibly provides a less powerful motivation for the students to maximise their efforts. The results of the exam in 2014 will probably make it possible to determine how extensive the effect of (the lack of) fine-grading is as the grammar exam in this year will be fine-graded again.

[6] In this task, the students have to determine the finiteness of the five verb forms provided. For this reason, the total number of possible answer patterns is $2^5 = 32$.

For the sake of putting the results of the grammar exam into perspective, a test was administered to a group of freshmen in their very first class of grammar at the university in order to see what level of knowledge of descriptive grammar they have from high-school. Only the knowledge of clause constituents and parts of speech was tested because these are the only skills that every Danish high-school student can be expected to have been exposed to. The average efficacy of clause constituents was 0.41 and that of parts of speech 0.58. This result suggests that the students do improve their knowledge of descriptive grammar during their first term at the university. Yet the relation between the skills concerning clause constituents and parts of speech remains the same, namely the knowledge of clause constituents consistently lags behind that of parts of speech year after year (see Table 2 to compare the results of this test with the results of the grammar exams).

3.1.2. Normalised database

As can be seen in Table 2, the testing of clause constituents has always been the singularly most dominant task constituting 20% of the entire battery of test questions. In order to compensate for this bias, a normalised database has been made. In this database, the number of questions concerning clause constituents has been reduced by ten by randomly deleting answers to such questions from all students, and thereby reducing the total number of questions to 90.

Of course, this normalisation does not eliminate all the differences among the sets of questions. However, it does reduce the bias towards the testing of clause constituents by reducing its weight to the same level as that of all the other major subtests, e.g. parts of speech, which are tested by 10 questions each. All the statistical calculations have been made on both the original and the normalised databases. Table 3 shows the differences between the two databases with respect to the overall averages and other statistics.

Table 3. *Differences between the original and normalised databases*

		2009	2010	2011	2012	2013
Original database	Overall averages	0.751	0.746	0.742	0.678	0.644
	Averages of questions about clause constituents	0.718	0.733	0.661	0.622	0.580
	Average number of questions/skill	9.09	8.33	8.33	7.14	7.14
	Standard deviation of the numbers of questions/skill	5.63	5.27	5.27	5.05	4.93
p-value for the paired t-test of the difference between the databases overall		0.102	0.313	0.001	0.014	0.000
p-value for the paired t-test of the difference between the databases with respect to the questions about clause constituents		0.649	0.630	0.477	0.341	0.237
Normalised database	Overall averages	0.755	0.748	0.749	0.683	0.653
	Averages of questions about clause constituents	0.713	0.738	0.650	0.610	0.593
	Average number of questions/skill	8.18	7.50	7.50	6.43	6.43
	Standard deviation of the numbers of questions/skill	4.28	3.99	3.99	3.72	3.54

It can be seen from the standard deviations in Table 3, which are smaller for the normalised database than for the original database, that the distribution of skills in the normalised database is indeed more even than in the original database. The high p-values concerning the questions about clause constituents indicate that erasing 10 questions randomly from each such set has not altered the averages of these questions significantly, suggesting that the normalised database is a good simulation of what would have happened had the students been asked fewer questions about clause constituents. On the other hand, the difference between the two databases concerning the overall results is statistically significant for the years 2011-2013, corroborating the assumption that knowing clause constituents plays a major role in knowing descriptive grammar. The consistently higher overall averages in the normalised database fit in well with the expectation that the overall average increases when the dominance of clause constituents is reduced since the average efficacy in clause constituents is lower than the average overall efficacy in all the years investigated (see Table 2).

3.2. Identifying the best indicators of success

In order to find the skills that are the best indicators of overall success when learning descriptive grammar, multiple linear regression analysis has been used to search for those triplets of skills that best correlate with the exam result (Elbro and Scarborough 2003; Hatch and Farhady 1982; Urdan 2012). Triplets of skills have been chosen in order to be able to achieve a high coefficient of determination (R^2), preferably around 0.9, while maintaining generalizability by not taking too many variables into account, which would make the results of the regression analysis too specific, valid only for the dataset analysed. Using three variables out of 11 to 14, that is about 25% of all the variables/skills, seems a reasonable compromise in balancing the descriptive and the predictive power of regression analysis (Baayen 2008).

Furthermore, skills that are tested by fewer than five questions have been excluded from the regression analysis. The reason for the exclusion is that the measured efficacy in these skills is very likely to contain too much random variation, which would make the regression analysis unreliable (DeVellis 2011; Dörnyei 2014; McNamara 2000). As a side effect of the regression analysis and the considerations connected with it, I have modified the grammar exam in 2014 so that all the skills in question are tested by at least five questions each in order to improve the reliability of the testing of the overall acquisition of descriptive grammar. Further research is required to investigate whether these measures have indeed improved the reliability of the testing of the knowledge of descriptive grammar.

3.3. Implicational analysis

In order to answer research question 2, i.e. whether there is an implicational relation between the acquisition of skills, those pairs of skills have been searched for that conform to the implicational equation I have developed. This equation assumes the following: For the acquisition of skill B to be dependent upon the acquisition of skill A, the level of efficacy in skill B must generally be below that of skill A; only when a high level of efficacy in skill A has been

attained, can a high level of skill B be present; and when the level of skill A is below a certain threshold, the level of skill B must be close to zero. This relationship is illustrated in Figure 1.

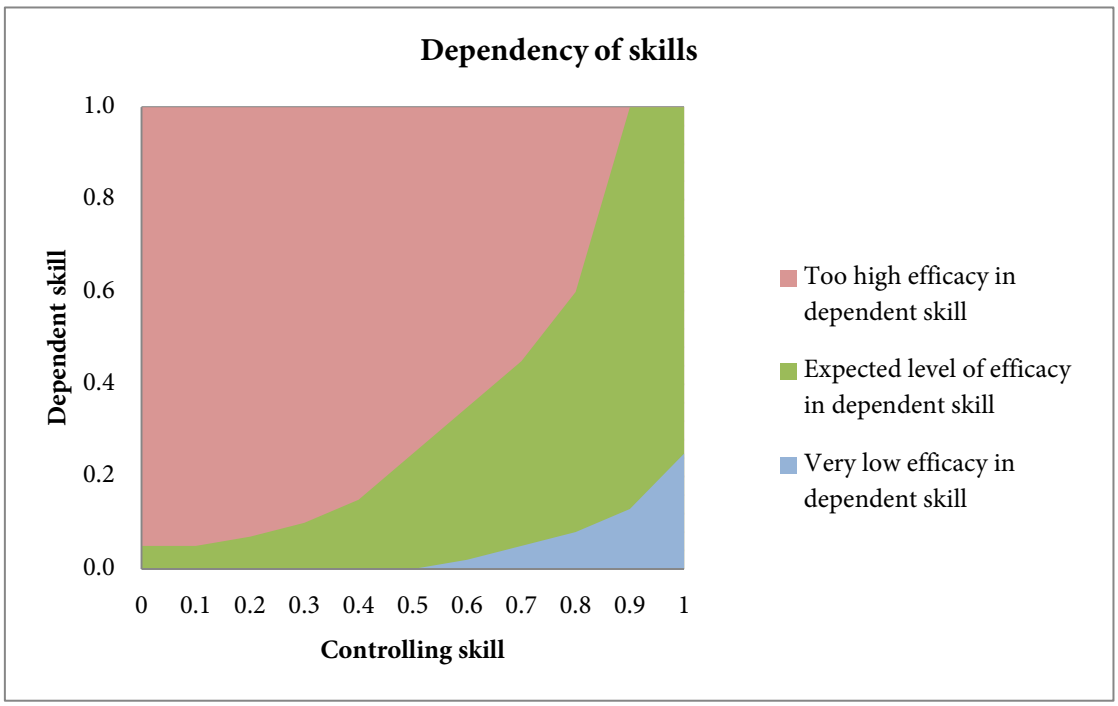

Figure 1. *Dependence of one skill on another one*

The green area consists of the values of efficacy in the dependent skill which conform to the hypothesis of implicationality, i.e. the level of knowledge in the dependent skill is always lower than or equal to the level of knowledge in the controlling skill – within some margin of error. The red area contains the values of efficacy in the supposedly dependent skill which actually contradict the implicational hypothesis as they indicate too high a level of knowledge in this skill compared to the knowledge in the supposedly controlling skill.

The borderline between the red and the green roughly follows the S-curve of learning. Because of the low granularity of the data, it is impossible to draw the S-curve exactly. At the leftmost end of the curve, the dependent skill is allowed to take on higher values than the controlling skill in order to allow for the possibility that students by chance score higher on the dependent skill at the outset of their learning. Up to the threshold level of 0.5 in the controlling skill, the dependent skill must be very low to comply with the implicational hypothesis; only after this point is it supposed to show an increase, tailing the increase in the knowledge of the controlling skill. At the rightmost end of the curve, where knowledge of the controlling skill tops out, the knowledge of the dependent skill is allowed to be as high as that of the controlling skill, or even slightly higher to allow for a chance variation in the scores.

The blue area indicates values of efficacy in the dependent skill which are rather low compared to the level of knowledge in the controlling skill. These values do not falsify the implicational hypothesis. They rather indicate students with unusually biased knowledge or a skill unusually difficult to master.

The threshold for the efficacy in the controlling skill has been set at 0.5 for two reasons. First, the granularity of the data is too low since the average number of questions is only about

6-9 per skill (see Table 3), and because of this low resolution it would be difficult to distinguish between true dependency and random variation below 0.5. Second, the implicational hypothesis states that the learning of the dependent skill only commences measurably when the controlling skill has already been acquired to some degree.

Naturally, it would be best to measure implicationality during the process of learning descriptive grammar with the help of a longitudinal investigation. If such a study showed that the learning of skill B takes off later than the learning of skill A, that would corroborate the assumption of implicationality convincingly. Unfortunately, I have no such data at my disposal. True, the students have to deliver three home assignments during the grammar course, and the assignments resemble the final exam. In fact, they are all composed of questions from previous grammar exams. These home assignments could in principle be used as dynamic data for a longitudinal study of the learning of descriptive grammar.

However, the home assignments have been deemed unreliable data for the purpose of studying the supposed implicational nature of the acquisition of skills in descriptive grammar. The reason for this is that the students have a whole week to do the home assignments, and they have access to every possible aid for their work. Hence it is questionable whether the home assignments accurately reflect the students' knowledge of descriptive grammar, knowledge defined as what one can do on one's own without any outside help, i.e. expert knowledge in Vygotsky's terminology (1978).

This belief is corroborated by the fact that students generally produce inferior results at the exam compared to their home assignments. Therefore, my analysis is based on the end results, the results of the grammar exam. These data may be static, yet they are a more accurate measure of the students' knowledge. And since most students do not attain a high level of knowledge of descriptive grammar by the exam (very few have an overall efficacy above 0.9), it is safe to assume that the results of the exam are a true reflection of the implicational relations among skills if there are any such relations.

4. Analysis

The results of the regression analysis and the results of the implicational analysis are first presented separately in their own subsections and then compared in the last subsection of the analysis.

4.1. Regression analysis

Tables 4a and 4b show the results of the regression analysis based both on the original database and the normalised database, in which the relative dominance of knowing clause constituents has been reduced. Table 4a gives an overview of the analysis, and Table 4b the details. As can be seen in Table 4a, it is indeed sufficient to have triplets of skills to attain high values of the coefficient of determination (R^2), often over 0.9, although the normalised database provides consistently lower values than the unmodified database. The p-values are in each case below 0.001, for which reason they are not listed separately.

Table 4a. *Regression analysis overall*

		2009	2010	2011	2012	2013
Original database	standard error	0.0325	0.0405	0.0377	0.0445	0.0421
	R^2	0.902	0.922	0.939	0.931	0.917
Normalised database	standard error	0.0341	0.0466	0.0420	0.0515	0.0504
	R^2	0.883	0.889	0.924	0.900	0.872

$p < 0.001$

Table 4b shows the equations that the regression analysis has provided for predicting the end result of learning descriptive grammar. It also tabulates the standardised coefficients, which show the individual variables' weight within the triplets in explaining the variability of the end result (the anonymous reviewer; Urdan 2012). The skills are ordered according to their weight in Table 4b. The p-values for all the coefficients and intercepts are far below 0.001 except for the intercept in the normalised database in 2011, for which the p-value is 0.113.

Table 4b. *Regression analysis in detail*[7]

Year	Database	Intercept	Skill 1	Skill 2	Skill 3
2009	Orig.	0.262	phrase types	subclause types	clause constituents
		b	0.320	0.179	0.194
		β	0.593	0.302	0.298
		r	0.431	0.544	0.527
	Norm.	0.248	phrase types	subclause types	phrase vs subclause
		b	0.256	0.244	0.189
		β	0.494	0.429	0.343
		r	0.431	0.232	0.266
2010	Orig.	0.163	clause constituents	NP constituents	parts of speech
		b	0.362	0.285	0.151
		β	0.550	0.401	0.188
		r	0.506	0.436	0.669
	Norm.	0.162	pronouns	NP constituents	parts of speech
		b	0.416	0.196	0.187
		β	0.797	0.285	0.242
		r	0.357	0.436	0.478
2011	Orig.	0.200	subclause types	pronouns	clause constituents
		b	0.436	0.190	0.160
		β	0.659	0.326	0.214
		r	0.486	0.609	0.409
	Norm.	0.047	NP constituents	subclause finiteness	phrase vs subclause
		b	0.404	0.263	0.225
		β	0.572	0.414	0.302
		r	0.206	0.210	0.520
2012	Orig.	0.173	clause constituents	phrase constituents	parts of speech
		b	0.352	0.244	0.177
		β	0.514	0.345	0.209
		r	0.706	0.729	0.793

[7] *b* stands for the coefficients (slopes) in the regression equations, and *β* for the standardised coefficients; *r* stands for the correlation coefficient between a skill and its neighbour to the right with wrap-around, that is, *r* under skill 3 shows the correlation between skill 3 and skill 1.

	Norm.	0.192	pronouns	clause constituents	parts of speech
		b	0.356	0.207	0.157
		β	0.677	0.334	0.192
		r	0.626	0.726	0.664
2013	Orig.	0.181	pronouns	clause constituents	phrase vs subclause
		b	0.378	0.227	0.148
		β	0.677	0.350	0.189
		r	0.543	0.430	0.163
	Norm.	0.201	clause constituents	pronouns	phrase vs subclause
		b	0.274	0.253	0.192
		β	0.479	0.470	0.254
		r	0.507	0.163	0.373

The results vary from year to year; however, it is hardly a surprise in light of the changes that the grammar exam has undergone in the period investigated (outlined in Section 3). The occasionally high *r*-values indicate a high degree of multicollinearity, which suggests that in these cases, even merely two variables would have produced a reasonable regression. This means that these two variables would suffice to make a reasonable prediction as to the oveall outcome of the final exam. It does not mean that it would be preferable to test only these two skills (or, for that matter, only the three skills that come out as the best predictors in the regression analysis) at the exam. For not having a great predictive power does not render a skill useless or unnecessary with respect to knowing descriptive grammar. Nor does weak predictive power mean that such a skill is learnt automatically through the learning of skills with greater predictive power or more easily than skills with greater predictive power. In fact, it is safe to assume that if the skills with less predictive power were not tested, the students would not bother with learning them. Nevertheless, there seem to be clear patterns. In order to make the interpretation easier, Table 4c summarises the ranking of the skills that appear in the regression analysis. Whenever a skill appears as skill 1, it is given three points; when it appears as skill 2, it is given two points, and when it appears as skill 3, it is given one point.

Table 4c. *Ranking of skills according to their predictive power*

Skill	Frequency	Points
Clause constituents	7	15
Pronouns	5	13
NP constituents	3	7
Subclause types	3	7
Phrase types	2	6
Phrase vs subclause	4	4
Parts of speech	4	4
Subclause finiteness	2	2
Phrase constituents	2	2

The skill of knowing clause constituents clearly emerges as the most powerful predictor of success in descriptive grammar. Not only does it appear in every year in the analysis based on the unmodified database and even in the analysis based on the modified database for 2012 and 2013, but it also has the strongest weight overall. This demonstrates beyond reasonable doubt that it is absolutely essential to know clause constituents if one wishes to master descriptive grammar.

It is closely followed by the skill of knowing pronouns. Together with the third most frequent, yet not particularly weighty skill, namely that of knowing parts of speech in general, it underlines the assumption that being able to classify words is a fundamental part of knowing descriptive grammar. Pronouns and clause constituents are far ahead of the rest of the skills.

Although there is a tie between NP constituents and subclause types, the knowledge of NP constituents is considered to be the third most dominant skill because it is supported by the closely related skill of phrase constituents. Nonetheless, it appears to be important to have the ability to distinguish between types of higher order structures such as phrases and subclauses since these two skills have a similar prevalence, closely followed by the skill of being able to distinguish between phrases and clauses. It makes good sense because the distinction between phrases and clauses is one of the most fundamental ones, at least in the theory (Hjulmand and Schwarz 2009) that the students studied here are presented with.

4.2. Implicational analysis

For all possible pairs of skills that are tested with at least 5 questions in the grammar exam, calculations have been made for both the degree of conformity and the degree of nonconformity to the hypothesis of implicationality, as outlined in section 3.3. The degree of conformity is the ratio of students that fall in the green area of Figure 1; that is, the number of students falling in the green area divided by the total number of students in the year concerned. The degree of nonconformity is accordingly the ratio of students who fall in the red area of Figure 1. The three pairs of skills that score highest on conformity in either the original database or the normalised database are tabulated in Table 5.

Table 5. *Implicational analysis; A: Controlling skill and degree of conformity; B: Dependent skill and degree of non-conformity*

		2009 original	2009 normalised	2010 orig. and norm.	2011 orig.	2011 norm.	2012 orig. and norm.	2013 orig. and norm.
1	A	phrase vs subclause	phrase vs subclause	number of matrix clauses	NP constituents	NP constituents	subclause finiteness	no pairs reached at least 0.500 con-formity
		0.732	0.714	0.714	0.759	0.722	0.828	
	B	clause constituents	clause constituents	parts of speech	clause constituents	phrase types	phrase vs subclause	
		0.250	0.268	0.270	0.222	0.278	0.138	
2	A	finding the subclause		subclause finiteness	NP constituents	number of matrix clauses	subclause finiteness	
		0.696		0.698	0.722	0.722	0.776	
	B	parts of speech		pronouns	phrase types	parts of speech	parts of speech	
		0.286		0.222	0.278	0.278	0.155	
3	A	finding the subclause		number of matrix clauses	number of matrix clauses	number of matrix clauses	phrase vs subclause	
		0.696		0.698	0.722	0.704	0.759	
	B	subclause types		phrase types	parts of speech	NP constituents	parts of speech	
		0.286		0.270	0.278	0.259	0.241	

The reason why conformity and nonconformity do not always add up to 1.000 is that the remainder is the ratio of students who demonstrate a very low level of skill B, the blue area in Figure 1. These students do not weaken the hypothesis, thus they could have simply been added to the students who show conformity. Nonetheless, I wanted to see if there were any students who have an unusually low score in a given skill B as compared to a given skill A. As it turns out, such students are very few in number, their proportion never exceeding 0.069.

The results of the implicational analysis are rather varied. In some cases, the pairings of skills make good sense. For example, as shown by pair 1 from the year 2009, it seems logical that the ability to discern clause constituents would depend on the ability to recognise what is a clause. Also, as shown by pair 3 from 2009, it makes good sense that the ability to determine subclause types would rely on the ability to find subclauses.

However, in other cases, the pairings seem rather random. For example, with respect to pair 2 from 2010, it is not at all clear why the ability to recognise pronoun types should in any way depend on the ability to determine the finiteness of a subclause. There are even pairs that go against common sense. For example, with respect to pair 1 in 2011, it would seem more plausible that the recognition of the constituents of an NP would depend on the recognition of phrase types, which includes the recognition of what is an NP, rather than the other way around as shown by the said pair of skills. This pair of skills also contradicts pair 1 and pair 3 from 2009, which seem to corroborate the assumption that recognising a higher order element, such as a subclause, is a prerequisite for being able to discern its constituting elements or subtypes.

I have as yet no explanation as to why the year 2013 has yielded such radically different results compared to the other years, there being not a single pair of skills showing clear implicationality. There are of course also from 2013 pairs of skills whose degree of conformity is higher than their degree of nonconformity. However, the degree of conformity of all these pairs is below 0.500, which entails that their degree of nonconformity is almost as high as their degree of conformity. Therefore they cannot be considered to show implicationality. The proportion of students who demonstrate an unexpectedly low level of skill B (the blue area of Figure 1) is miniscule also in 2013, thus negligible.

It is certainly worth further investigation to see whether the unexpected and quite surprising results of the implicational analysis have any validity or are the product of false assumptions or flawed calculations. Unfortunately, as mentioned above, I have limited access to reliable data on the longitudinal development of the skills, especially of the skills which are discussed late in the grammar course, e.g. subclause types and a number of matrix clauses.

4.3. Comparison of the regression and implication analyses

The results of the regression and implication analyses only partially overlap as shown by Tables 4b, 4c and 5. The skill most frequently appearing in the results of the implicational analysis is the knowledge of parts of speech, which is also among the skills most frequently appearing in the regression analysis. The fact that it is the dependent skill in all cases in the implicational analysis goes hand in hand with the low predictive power of knowing parts of

speech despite its fairly high frequency (Table 4c). However, these results contradict the expectation that knowing parts of speech is a basic skill. This, of course, does not diminish the importance of knowing parts of speech as an important element in the knowledge of descriptive grammar.

The skill which is the most dominant one in the regression analysis, namely clause constituents, is not quite as prevalent in the implication analysis. Nevertheless, it is the dependent skill in all the pairs it appears in. Considering that this skill is so significant in predicting the students' overall achievement in descriptive grammar, and that the students consistently score lower in this skill than the overall average (see Table 2), it is probably worth paying more attention to teaching the skills that this skill is presumably dependent on, most notably the ability to distinguish between phrases and clauses.

Interestingly, the skill that is the next most prevalent one in the implication analysis, namely the recognition of matrix clauses, does not emerge in the regression analysis at all. On the other hand, skills such as distinguishing between phrases and clauses, phrase types, and NP constituents, have a similar prevalence in both analyses. Their emergence from both analyses suggests that it is worth paying more attention to them in the teaching of descriptive grammar because they are likely to play an important role in grasping descriptive grammar.

5. Conclusion

Somewhat surprisingly, the results of the regression analysis vary from year to year. However, clause constituents, pronoun types, phrase constituents, the distinction between clause and phrase, and parts of speech surface as recurring elements. Knowing clause constituents is often significant even when their bias in the grammar exam has been reduced.

The analysis of implicational relationships has yielded an even more varied result. Many of the pairings of skills seem rather random without any logical relationship between the skills in the pairs. However, there does seem to be a weak pattern in the way that knowing (being able to recognise) higher level (i.e. more complex) constructions, such as clauses and phrases, is a prerequisite for mastering the ability to distinguish between lower level elements, such as the constituents and types of subclauses.

All in all, the analyses do seem to confirm the assumption that knowing clause constituents and the distinction between phrases and clauses are the most significant items in learning descriptive grammar whereas knowing parts of speech does not seem to play as large a role as expected. Besides these, the analyses revealed that it may be beneficial in the teaching of descriptive grammar to pay more attention to certain skills, such as distinguishing between phrase types, and phrase constituents. The possible importance of these skills was not part of the original hypothesis, but discovered as a result of this study.

References

Baayen, R. Harald. 2008. *Analyzing Linguistic Data: A Practical Introduction to Statistics Using R.* Cambridge: Cambridge University Press.

Clahsen, Harald, Jürgen M. Meisel, and Manfred Pienemann. 1983. *Deutsch als Zweitsprache. Der Spracherwerb Ausländischer Arbeiter* [German as a second language. Language acquisition of guest workers]. Tübingen: Gunter Narr.

Corder, S. Pit. 1987. *Error Analysis and Interlanguage.* Oxford: Oxford University Press.

DeVellis, Robert F. 2011. *Scale Development*: *Theory and Applications*. SAGE Publications. Kindle Edition.

Dörnyei, Zoltán. 2014. *Questionnaires in Second Language Research*: *Construction, Administration, and Processing*. Taylor and Francis. Kindle Edition.

Elbro, Carsten, and Hollis Scarborough. 2003. Early identification. In *Handbook of Children's Literacy*, edited by Terezinha Nunes, and Peter Bryans, 339-359. Dordrecht: Kluwer Academic Publishers

Ellis, Rod. 2009. *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. Buffalo: Multilingual Matters.

Ellis, Rod. 2012. *Second Language Acquisition*. Oxford: Oxford University Press.

Jarvis, Scott, and Aneta Pavlenko. 2008. *Crosslinguistic Influence in Language and Cognition*. London and New York: Routledge.

Jarvis, Scott. 2011. Conceptual transfer: Crosslinguistic effects in categorization and construal. *Bilingualism: Language and Cognition* 14 (1), 1-8.

Hatch, Evelyn, and Hossein Farhady. 1982. *Research Design and Statistics for Applied Linguistics*. Rowley, Massachusetts: Newbury House Publishers.

Hjulmand, Lise-Lotte, and Helge Schwarz. 2009. *A Concise Contrastive Grammar of English*. Frederiksberg: Samfundslitteratur.

Krashen, Stephen D. 1981. *Second Language Acquisition and Second Language Learning*. Los Angeles: University of Southern California.

Lund, Karen. 1997. *Lærer Alle Dansk på Samme Måde? En Længdeundersøgelse af Voksnes Tilegnelse af Dansk som Andetsprog* [Does everyone acquire Danish in the same way? A longitudinal study of adults' acquisition of Danish as a second language]. Herning: Special-pædagogisk forlag.

Lyon, G. Reid, and Louisa C. Moats. 1997. Critical conceptual and methodological considerations in reading intervention research. *Journal of Learning Disabilities* 30 (6), 578-588.

Madsen, Richard. 2014. Correlation between theoretical knowledge of grammar and performance in the production of written texts. In *Multidisciplinary Perspectives on Language Competence*, edited by Lotte Dam, and Rita Cancino. Aalborg: Aalborg University Press.

Madsen, Richard. (*unpublished*). *Initiatio Grammaticæ Anglorum-Saxonum et Principiis Methodisque Analysis Grammaticalis* [Introduction to English grammar and the principles and methods of grammatical analysis].

McNamara, Tim. 2000. *Language Testing*. Oxford: Oxford University Press.

Pienemann, Manfred. 1998. Developmental dynamics in L1 and L2 acquisition: Processability theory and generative entrenchment. *Bilingualism: Language and Cognition* 1 (1), 1-20.

Urdan, Timothy C. 2012. *Statistics in Plain English*. Taylor and Francis. Kindle Edition.

Vygotsky, Lev. 1978. *Mind in Society*: *The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.