Learner Language and Language Learning

CLAUS FÆRCH, KIRSTEN HAASTRUP
ROBERT PHILLIPSON

In collaboration with
Esther Glahn and *Knud Bæk Kristiansen*

Gyldendals Sprogbibliotek

For an up-to-date summary of Kellerman's work on transfer see Kellerman (in press). Our discussion of semantic issues is based on Lyons 1977, which is a demanding standard work. For a more popular introduction to semantics see Palmer 1976. On semantics and language teaching see van Buren 1975. For Danish learning problems see Davidsen-Nielsen/Færch/Harder 1982. Lexical inferencing is based on Carton 1971 and Færch 1983a.

Chapter six: Grammar

Grammar comprises syntax and morphology. Syntax has received considerable attention in recent decades, whereas morphology has been relegated to a less prominent position. The main reason for this is that linguists have been concerned to discover the nature of linguistic creativity, the human ability to comprehend and produce an infinite number of sentences not yet encountered, and this has been assumed to be intimately related to the productivity of syntactic rules.

True though this undoubtedly is, holding up syntax as the most productive area of language should not be taken to imply that syntax is in any sense more important for communicative purposes than, say, morphology or vocabulary. The importance of syntax is of a different sort.

Both Danish and English are highly ANALYTIC languages. Word order helps express specific types of meaning, for instance agentivity (compare "the child scolded the teacher" with "the teacher scolded the child"). This means that Danish learners of English are used to word order being strictly fixed. Learners with L1s which are more SYNTHETIC, ie languages which like Latin express many types of meaning by means of affixes or declension, and which have a freer word order, have difficulty in learning an analytic language because they are not accustomed to paying sufficient attention to word order. The reverse situation, speaking an analytic language and learning a synthetic one, is equally difficult, as when Danes attempt in vain to identify subjects and objects in a Latin text by means of word order only.

Inexperienced learners are probably much less conscious of word order characteristics than of vocabulary. This implies that they may be more inclined to transfer L1 syntactic patterns to L2 than to transfer vocabulary. The fact that the similarities between Danish and English syntactic structures outnumber whatever differences there are between the two languages makes it possible for Danish

learners to make extensive use of L1 structures in IL speech, and to gradually modify L1 structures as learning progresses.

This book is not the place for a systematic contrastive description of Danish and English grammatical structures.[1] What we shall do is to present results from various analyses of learner language which serve the dual purpose of exemplifying interlanguage grammar analysis and of describing a number of characteristic aspects of the learning of English grammatical structures by Danish learners. Finally we shall consider the nature of grammatical rules.

6.1 Analysis of errors in written texts

In this section we present some results from the PIF analysis of errors in essays and film descriptions written by learners at different levels. The main focus will be on grammatical errors made by Gymnasium learners, but in order to put these in a wider perspective, we also consider texts written by learners from grade 8, as well as errors in orthography, punctuation and vocabulary. Table 5 provides details of the three sets af data.

Table 5: Text sample used for error analysis

Grade	8	1gs	3gs
Numbers of informants	12	12	11
Total number of words	4846	7859	8892
Average length of individual essays in words	404	655	808

An analysis of all errors identified in the texts (for the procedure of error identification, see the description in chapter 19) resulted in the following figures, divided into four broad categories.

On the basis of these figures alone, it would be possible to answer questions such as "are lexical errors more frequent than grammatical errors in texts written by learners at grade 3g?". However, the

Table 6: Errors, absolute figures

Grade	8	1gs	3gs
Orthography	524	313	228
Punctuation	111	123	120
Vocabulary	230	287	241
Grammar	505	438	300
Total	1370	1161	889

figures are *absolute* figures. They pay no attention to the frequency of letters, punctuation marks, words and grammatical constructions in the texts.

The simplest way of obtaining a (very crude) picture of the overall frequency of errors is to calculate the ratio between the figure for each error type and the total number of words. This is shown in table 7.

Table 7: Errors as percentages of total number of words

Grade	8	1gs	3gs
Orthography	11 %	4 %	3 %
Punctuation	3 %	2 %	1 %
Vocabulary	5 %	4 %	3 %
Grammar	10 %	6 %	3 %

Put differently, these figures for the grammatical errors imply that a learner at grade 8 has one grammatical error per 10 words written, a learner at grade 1gs has one grammatical error per 16 words, and a learner at level 3gs has one grammatical error per 33 words. There is therefore a clear development across the three

levels towards more correct English, measured in purely quantitative terms.

The figures in table 7 give us no information about which features of learners' IL change and become correct and which resist change. Do learners at different levels make different types of error? To investigate these questions, all grammatical errors in the PIF written corpus were classified into error types, and figures calculated for the frequency of each type. For reasons of space we shall from now on ignore the 8th grade, and concentrate on the Gymnasium learners. Table 8 lists results for the 1gs and 3gs texts, giving frequencies for the 10 most frequent types of grammatical error.[2]

Table 8: Frequent error types – "top 10 list"

	1 gs		3gs	
	% of errors	rank order	% of errors	rank order
(a) Concord	12,4	1	6,6	6
(b) Determiners	11,7	2	11,1	2
(c) Predicate	11,0	3	14,2	1
(d) Verb phrase	9,8	4	10,4	3
(e) Tense	7,7	5	2,4	14
(f) Noun phrase	7,4	6	9,3	4
(g) Word order	7,2	7	7,3	5
(h) Adverbial phrase	5,5	8–9	5,9	9
(i) Aspect	5,5	8–9	6,2	7–8
(j) Number	4,8	10	6,2	7–8
Total	83,0		79,6	

In order to interpret these figures, it is necessary to scrutinise closely the error types which actually occur. We shall discuss each of these.

(a) Concord. These are errors in subject-verb concord, following either a pronominal or a non-pronominal subject. Concord errors at the Gymnasium levels do not usually occur when the subject is a countable noun and the verb follows immediately. They primarily occur in the following situations:

1. the sentence is initiated with *there* ("There *are* violence as well in the school")
2. the subject is not a countable noun with singular and plural form ("... the police *has* never found his mordereres")
3. the subject is an indefinite pronoun ("Some *maintains* that ...")
4. the verb is divided from the head noun in the subject ("The fights in Ireland *is* not because of political problems ...")

(b) Determiners. Most of these errors involve using *the* instead of no determiner in noun phrases which would contain a determiner in Danish, typically when marking generic reference:

"*The* violence has been developed ..."/ "*The* schools are also a place where violence ..."

or with reference to institutions, following a preposition:

"... in *the* school you do not call it violence but mobing".

Overuse of determiners when marking generic reference is by far the most frequent type of determiner error at 3g level; whereas the "in the school" type is common at level 1gs but hardly ever occurs in 3g.

(c) Predicate. These errors are a very heterogeneous group. Approximately half the errors are borderline cases between grammar and use of idioms, as illustrated by the following example:

"... it's of course also good to see movies about people who *have it worse* than you".

Such errors are often caused by transfer from Danish. A less frequent group of predicate errors are due to wrong verbal complementation, eg:

"... when you walk in New-York in the night you can *risk to be dagged*".

(d) Verb phrase. Errors of tense, aspect, auxiliary and modal verbs, and concord have been analysed separately. This category covers all other verb phrase errors. Most frequent are morphological errors ("... the man ran after the woman he loved and *beated* her"), whereas errors in the use of *do* as a prop word are extremely rare at these levels.

(e) Tense. These errors, which are much more frequent in 1gs than in 3gs, mostly involve a shift to the present tense in a context which requires past reference:

"It wasn't that kind of violence we have now, but *it's* surely a start to it".

Interestingly, there is only one occurrence of a learner using a perfect tense instead of a past tense form ("The picture which I've seen some minutes ago") and very few examples of wrong reference to the future ("They say that if their demand is not being followed they *are going to* kill the people ...").

(f) Noun phrase. This category does not contain errors of number or determiner usage, but typically errors of nominalization of adjectives or numerals:

"You also see some reel pictures like *two* shaking hands",

wrong compounding:

"To get to a quite other *violence-problem* it is the terrorism in Germany"

or lack of pronominalization:

"It's very confused to see the film the first time but after having seen it two or three times you begin to understand a bit of *the film*".

(g) Word order. By far the most frequent type of word order error is adverbial placement, both in 1g and 3g texts:

"All kind of violence *dayly* take place in all kinds of ways at all kinds of places"
"Violence is *today* a very big problem...".

Wrong inversion does not seem to present many problems at these levels (see below for the situation at grade 8).

(h) Adverbial phrase. These errors are all caused by the learner choosing a wrong structure for expressing an adverbial phrase:

"For some weeks ago, I read ..."
"... they were looking all over *for to find guns*".

(i) Aspect. This type of error is less frequent in the PIF data than we had expected. When the number of aspect errors is related to the number of finite verbs in the corpus (1133 in the 1gs texts and 1179 in the 3g texts), it is found that only 2 % of the verb phrases at 1g level and 1.5 % at 3g level are wrongly marked for aspect. It is more usual for learners, especially at the lower level, to overuse the progressive aspect where none is required than vice versa: At level 1g, the ratio between these two types of aspect error is 10:1, at level 3g it is 3:1. The relatively low figures for aspect errors must in part be due to the nature of the two writing tasks.

(j) Number. A few of these errors are due to learners not marking for 'distributive plural': "...they [people] are blaming them [the terrorists] that innocent people become the victims of *their act* [ie acts]".

Other errors, especially at level 1gs, are of the type *"informations", "peoples"* (for *people).*

These results suggest that when learners leave the Gymnasium, they generally know how to handle English morphology; that some aspects of English syntax (determiner usage and adverbial placement) remain a problem; that certain error types which occur at the Folkeskole level are very infrequent at the Gymnasium level (use of *do;* inversion; aspect); and finally that *many* errors, broadly categorized

as 'predicate', 'noun phrase' or 'adverbial phrase', are borderline errors between grammar and idiom usage, and seem to be modelled on Danish.

Certain categories of function words, like prepositions and conjunctions, have not been covered so far in the analysis of grammatical errors. Like idioms they belong to the border area between grammar and vocabulary. We shall complete our description of grammar errors in written texts at the 1gs and 3gs levels by considering the most frequent type of errors in the selection of function words, namely errors in prepositions. At both 1gs and 3gs levels, 7.3 % of all prepositions used are erroneous. One might speculate that the reason for this apparent lack of progress is a greater inventory of prepositions at grade 3gs. However, both levels have exactly the same type/token ratio (0.47). If we assume that this PIF result also holds for other Danish learners, we might conclude that Gymnasium learners do not seem to improve their knowledge of when to use which prepositions. This is an area which needs much more systematic treatment in teaching – which presupposes the availability of both good descriptions of prepositional usage in English[3] and contrastive descriptions of prepositions in Danish and English.

6.2 Analysis of errors in speech and writing

The results presented above give a very simplified picture of grammatical errors in learner language. Among the most obvious limitations are that:

– within each group of learners (cf table 5), individual variation has been levelled out,
– the analyses focussed on errors exclusively and did not consider to what extent learners knew how to use grammatical structures correctly,
– the analyses were based on written language texts only, and variation in the individual learner's spoken and written interlanguage was not considered.

The following example (based on Færch 1983c) illustrates how each of these considerations can be tackled in learner language analyses.

To do so it is necessary to narrow the focus to specific types of grammatical error.

The area investigated was subject-verb inversion following an initial adverbial element. This is obligatory in Danish in all cases, but restricted in English to certain types of adverbial phrases (cf Quirk/Greenbaum/Leech/Svartvik 1972, chapter 8, Davidsen-Nielsen/Færch/Haarder 1982:44ff.):

Danish "I går *mødte jeg* en flot ung fyr"
English "Yesterday *I met* a handsome young guy".

The analysis was based on PIF texts produced by 12 grade 8 learners: an oral film description, a written film description, and an essay (cf chapter 18). In each text, clauses were identified which would have inverted word order in Danish because of an initial adverbial element. The result of this phase of the analysis is presented in table 9, line A. The next phase involved calculating how many of these English clauses actually did contain inverted word order, and whether inversion was correct or erroneous. These results are listed in lines B–D, table 9. As can be seen, the essays contained the largest number of inversion errors (16.7 %), followed by the written film descriptions (9.2 %). It is striking that the oral film descriptions contain no inversion errors whatsoever, and that none of the initial adverbial elements was such that inversion was required in English (cf line C).

Table 9 takes up two of the three problems mentioned above (errors – non-errors and speech – writing), but it still levels out individual variation. In order to investigate this, a calculation was made, see table 10, of how many learners made inversion errors

– consistently, line x,
– variably, line y,
– or not at all, line z.

The results show that in the essays, more than half the learners make inversion errors (one consistently, the rest variably). In the film descriptions, only one third of the learners make inversion

Table 9: Errors/non-errors in inversion in three tasks

		Oral film description	Written film descr.	Essays
(A)	Number of clauses in which inversion would be obligatory in Danish	34	65	84
(B)	Number of clauses with inversion, erroneous in English	0	6 = 9.2 %	14 = 16.7 %
(C)	Number of clauses with inversion, correct in English	0	0	0
(D)	Number of clauses with non-inverted syntactically determined word order	34 = 100 %	59 = 90.8 %	70 = 83.3 %

Table 10: Number of learners making inversion errors

(x)	Number of students who use inversion only	0	0	1
(y)	Number of students who use inversion and non-inversion variably	0	4	6
(z)	Number of students who use non-inversion exclusively	12	8	5

errors (variably). And as could be seen in table 9, no learner makes inversion errors in the oral film descriptions.

At first sight, the inversion results in the three tasks look surprising. One might reason that with more time available, the better the learner's chances of finding the correct rule. As compared with writing away quietly, speaking in the language laboratory with a cassette recorder switched on might be expected to increase, rather than reduce, the transfer of L1 grammatical structures. Why then

do the learners produce correct English sentences in the most demanding situation, and make consistently more errors the more time they have at their disposal to 'monitor' speech (cf chapter 8)? The answer is probably that the learners at this level have practised English structures containing non-inversion to such a degree that, as long as they perform without thinking too much about it, they can perform correctly. Their less monitored performance is highly 'automatized' (cf chapter 11). In situations which allow them to be more conscious about their language, some of the learners become uncertain about the rule and at times end up writing what they would *not* say spontaneously.

This uncertainty is in fact directly attested in one of the essays. The learner's handwritten text reads:

"... if you run with the ball ~~the defender~~ would the defender kick you".

The result of the learner monitoring his language is in this case a self-correction, leading to an error.

The 'de-automatization' effect of self-monitoring can be compared to the well-known situation of opening a bicycle combination lock. We can perform the operation of pressing and pulling the right buttons a thousand times, without thinking about it. But when lending the bicycle to a friend, we may have difficulty explaining which buttons to pull and which to press, and if we then try to demonstrate it at the same time, we fail. One might say that we switch from a 'skill-based' type of performance to a more cognitively based performance, and that we run into difficulty if we do not have a cognitive representation which matches our skills. To return to the grade 8 learners, it is a fair guess that the written texts would have contained fewer inversion errors if the learners had cognitively *known* the correct rules for inversion/non-inversion in Danish and English.

6.3 Types of grammatical rules

We demonstrated above that there was variation in the inversion rules followed by some of the learners at grade 8 in three different tasks. In doing so, we conflated two distinctly different ways of talking about rules:

(1) LINGUISTIC RULES: rules etablished in order to account for language data, formulated in linguistic terminology and belonging to a specific model af language description.
(2) PSYCHOLINGUISTIC RULES: rules as psychological entities, activated by individuals when they produce language.

In 6.2 we took as our starting point linguistic rules, established on the basis of learner language data, and 'projected' them onto the learners' skill-based or cognitively-based representation of language. This is defensible as long as we make it clear that in referring to learners' rules as psychological entities, we do not imply that the rules are *psychologically* represented in a way which necessarily looks like the *linguistic* formulation of the rules. Linguistic models provide convenient and, hopefully, adequate ways of talking *about* language. The extent to which they can be assumed to have psycholinguistic validity is a very debatable issue (see for instance Gregersen/Hermann 1978 and chapter 11 for a discussion of this).

Another distinction needs to be observed between

(3) NATIVE RULES: used by native speakers or rules which aim at being descriptively adequate for the performance of native speakers
(4) INTERRULES: rules used by learners of a foreign language or formulated for the benefit of foreign language learners.

The four types of rules are brought together in figure 8.

We need to be more specific about the differences between linguistic rules of type C and D. Ideally, native speaker linguistic rules (C) exhaustively describe the language performance of native speakers. By means of C-rules, one can 'generate' all grammatical structures

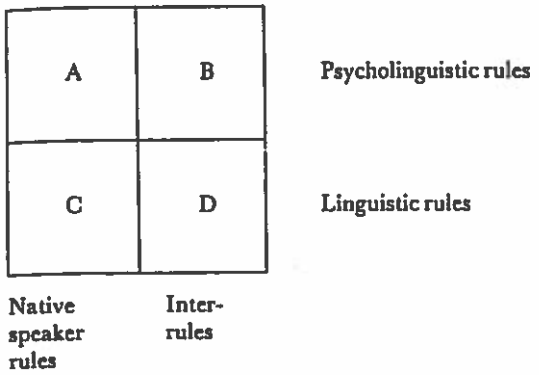

Fig. 8: Four types of grammatical rules

accepted by native speakers as belonging to the language (cf Lyons 1981). Interrules (D), on the other hand, can be one of two types:

(5) rules describing the performance of IL users (ie the exact IL parallel to C-rules)
(6) PEDAGOGICAL INTERRULES constituting af (simplified) subset of native speaker rules (C) and established for didactic purposes.

Learner language studies, of the kind demonstrated in 6.1 and 6.2, are concerned with type (5) rules. We shall now go on to consider pedagogical interrules (6). Use is chiefly made of these in grammar books and teaching materials produced for foreign language learners. Various criteria are relevant when considering the nature of pedagogic grammars, among others the following:

(1) The relationship between a pedagogic grammar and a descriptively adequate native grammer. Differences may reflect *selection principles* (are there rules which have been left out in the pedagogic grammar?) or *degree of detail* (how precisely has a rule been formulated?).
(2) The linguistic theory on the basis of which the pedagogic grammar is written (word-class based? syntax based? eclectic approach?).

(3) Whether the pedagogic grammar is written for learners with a particular mother tongue. Such grammars may select areas in the L2 which are specially problematical because of the contrastive relationship between the two languages. They may also explicitly mention native language structures wherever these can clarify specific problems.
(4) The extent to which the pedagogic grammar incorporates considerations of language learning, for instance by discussing typical cases of learners' grammatical errors.
(5) Whether the pedagogic grammar is intended as primarily a reference grammar or a 'teaching' grammar.

The follow-ups make suggestions for applying these criteria in the analysis of a number of pedagogic grammars currently used in Denmark.

A different use of pedagogical interrules occurs in the classroom, when teachers formulate grammatical rules. Sometimes these are identical with rules found in grammar books. But in addition there appears to be a tradition of using 'rules of thumb', ie rules which are easy to remember and which work in some, but not necessarily all cases. The study of such rules may provide us with valuable information about the types of grammatical knowledge that learners develop.

In the first example the class are translating sentences from Danish into English. For homework they had been asked to prepare the past tense forms of *bend* and *break*. For this reason the teacher has difficulty accepting the learner's plea of ignorance.

Text 12 [3]

T	now
L	naturally he didn't – – jeg ved ikke rigtig hvad det hedder at bøje –
T	– wait a minute – are you seeking for the past tense – remember – now
L	
T	you said – he didn't – and what then – yes you
L	det ved jeg ikke (giggle)

T	know – when you use do and did – which form of the verb is that we use
L	
T	then I know you know – OK er – Thomas –
L	nej det ved jeg ikke – –
T	"naturligvis bøjede han den ikke" she said naturally he didn't – –
L	how –
T	BEND it – because you use the infinitive Susanne – don't you know
L	bend it
T	– yes –
L	

We can extract points from the text as below, identifying the activity types carried out by teacher and learner and the way the rule becomes formulated. In the middle column are key words from the extract:

Activity type	*Keywords from performance*	*Interrule*
Problem formulation	past tense? (T)	
Induction	he didn't – and what then? (T)	
		when you use *do* and *did* ... (T)
	he didn't? (T)	
Solution	bend it (L) BEND it (T)	
		you use the infinitive (T)

Fig. 9: Example of interrule formulation

It is the teacher who formulates the problem. She tries to elicit the correct response from the learner, partly in terms of performance ("he didn't – and what then?") partly by shifting to the linguistic rule level and giving the first half of a rule which she apparently expects the learner to know. She then elicits the correct completion to the sentence from a different learner, repeats this and then gives

the second half of the rule ("you use the infinitive"). Immediately after this, the teacher goes on to the next sentence and never returns to the *do* + infinitive rule.

We end this chapter by quoting another example of the teacher correcting a learner and formulating a grammatical rule. In this instance the rule finally formulated by the teacher is of rather dubious validity, or at the very least an overgeneralisation.

Text 13 [4]

T	i stedet for instead of
L	everyone could see that it would break – instead of
T	what – yes – if you say instead
L	instead of bending – – (laughter in class)
T	of bending – then it is correct – can't you hear it sounds strange – to say
L	
T	that you will do something instead of – – you never end it – Søren
L	every-
T	instead yes – you can't end a
L	one could see that it would break instead
T	sentence with a preposition like that – either you have to add some-
L	nå –
T	thing or you leave out the the – er for instance here the of – you just say
L	
T	instead – yes er –
L	

Footnotes to chapter 6

1. For a description of some areas of grammar, see Davidsen-Nielsen/Færch/Harder 1982.
2. The error categories in table 8 have been conflated from many more error categories in the original analysis. See the discussion of error analysis in chapter 18.
3. See chapter 4, footnote 8.
4. Same as text 12.

Chapter 6. Follow-ups

1. Identify grammatical errors in a set of free compositions (for the procedure of error analysis see chapter 18). Compare results with the "top10 list" in table 8.
2. Identify all instances of the simple and the progressive present tense forms in a set of learner texts. Calculate the proportion of errors relative to the total number of occurrences of each category. Which of the two tense forms seems to be most troublesome to the learners? Check in a number of pedagogical grammars whether the pedagogical interrules formulated for the simple/progressive tenses adequately explain the use of these. Could learners have avoided their errors by correctly following the interrules?
3. Compare the description of a certain grammatical phenomenon in (a) Quirk/Greenbaum/Leech/Svartvik 1972 (b) a number of pedagogic grammars of English, using the criteria listed on p. 115. Are there other relevant criteria?
4. Ask fellow students/teachers to note down examples of "rules of thumb". Find criteria which can be used to characterise them. Discuss the pedagogic value of such rules.

Chapter 6. Sources and further reading

Hatch 1978a describes different approaches to the analysis of grammar in IL data. Further studies of a more theoretical nature are listed in chapter 17 and 18.

The concept of linguistic rule is covered in most introductory textbooks on linguistics. See for instance Lyons 1981, chapter 2. Pedagogical interrules are discussed in Bausch 1979, which is a comprehensive collection of articles on pedagogic grammar. For a brief introduction to pedagogic grammar, see Færch 1977. "Rules of thumb" and other teacher-formulated rules are discussed in Færch 1983d.

The relationship between explicitly formulated grammatical rules and language teaching is discussed in various places in Krashen 1982. For different views, see Seliger 1979, Sharwood Smith 1982, Bialystok 1982.