DOCUMENT RESUME

ED 346 745 FL 020 468

AUTHOR Woodley, Marie-Paule
TITLE Non-Nativeness in Second Language Texts: The Syntax Factor.
PUB DATE 87
NOTE 14p.; In: Written Language: British Studies in Applied Linguistics 2. Papers from the Annual Meeting of the British Association for Applied Linguistics (Reading, England, September 1986); see FL 020 460.
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Contrastive Linguistics; English; Foreign Countries; French; Language Research; Second Language Learning; *Sentence Structure; *Syntax; *Written Language

ABSTRACT

The so-called deviant character of a set of non-native texts is examined by looking closely at how sentence syntax realizes and affects textual functions. Two broad groups of syntactic phenomena are considered: subordination and "marked structures," such as passives and clefts. Emphasis in this paper is on the following four ways in which syntax can be seen as contributing to explicit coherence: linking, or the establishment of explicit links between propositions; foregrounding/backgrounding (within sentence or within discourse); topic selection/continuity; and focus marking. The data for this exploratory contrastive study consist of three sets of texts, including native French texts, native English texts, and non-native French texts. The research task required the subjects to take sides in a debate current at the time of the data collection, backing their argument with elements drawn from simple statistical data provided. Results suggest the following tentative conclusions: (1) lower syntactic complexity may be related to lower linking density, to less topic selection, and to looser topic continuity; (2) the different ways in which the groups used syntax to foreground and background elements within the sentence and within the text as a whole led them to produce quite different types of texts; and (3) there appeared to be little direct transfer of text-building devices from first language to second language. It is suggested that the evidence of considerable differences between text-building devices used by native writers of French and English calls for detailed contrastive research. Contains 8 references. (LB)

NON-NATIVENESS IN SECOND LANGUAGE TEXTS: THE SYNTAX FACTOR

Marie-Paule Woodley
University of Manchester

That non-native texts 'deviate' from the corresponding native norm is obvious to all second and foreign language teachers, most of whom see it as their task precisely to guide learners toward a more native-like performance. One approach to the characterisation of these 'deviations' - itself a fundamental step in the planning of intervention in the language-learning process - has been the systematic analysis of errors, and it is an indication of the 'privileged' status of syntactic problems that most error analyses have concentrated on them. Syntactic error analyses, however, have tended to look at syntax as a sentence phenomenon, purely in terms of grammaticality, and not as one of the means of establishing coherence. This paper seeks to arrive at an understanding of the 'deviant' character of a set of non-native texts by looking closely at how sentence syntax realises - and affects - textual functions. The questions asked therefore are not only: 'Is non-native syntax different?' and 'In what way?', but 'Do writers do different things with syntax in their L1 and in their L2?' and 'Is what they do different from what native speakers of L2 do?'

Two broad groups of syntactic phenomena will be considered, both of which have been the object of much recent research on their links with textual functions: firstly, subordination, and secondly, 'marked structures', such as passives and clefts (see Lautamatti 1978 and Tomlin 1985 on subordination, Prince 1978, Borkin 1984 and Lambrecht 1985 on clefts). They can be looked at as some of the devices that make a text *explicitly* coherent, a highly valued characteristic in the type of academic writing I am concerned with (see Reinhart 1980 on explicit coherence, Lakoff 1984 on the cultural value attached to certain linguistic choices). In this paper I will concentrate on the following ways in which syntax can be seen as contributing to explicit coherence:

1. linking: the establishment of explicit links between propositions;

2. foregrounding/backgrounding (within sentence or within discourse);

3. topic selection/continuity

4. focus marking.

The data for this exploratory contrastive study consist of three sets of texts:

F = 'native' French texts (13 subjects, 5975 words)

EL1 = 'native' English texts (15 subjects, 9165 words)

EF = 'non-native' French texts (same 15 subjects as EL1, 5985 words)

The task required the subjects to take sides in a debate current at the time of the data collection, backing their argument with elements drawn from simple statistical data provided.

A first answer to the question: 'is non-native syntax different?' can be sought by looking at T-unit length in words. A T-unit is a syntactic unit consisting of a main clause plus any subordinate clause or non-clausal structure that is attached to or embedded in it (Hunt 1970). T-unit length works as an - albeit very rough - index of syntactic complexity since the number of words tends to grow with the number of dependent clauses. The mean T-unit lengths for my data are as follows: F = 21.7 words/T-unit, EL1 = 24.4 words/T-unit, EF = 19.5 words/T-unit. As could be expected, the English subjects write considerably longer T-units in their native language than in French. There is also a gap between EF and F, though it is narrower. The figures point - in a somewhat atypically subdued manner compared with other studies by this author - to the existence of a difference in levels of syntactic complexity between the three sets of data. This cannot be equated in any direct way with any textual function as there is rarely a one to one relationship between syntactic form and textual function. For example, in the case of the first of the textual functions listed above, namely **linking**, because propositions within a given text are expected to 'go together',

juxtaposition can suffice to establish the link. Propositions can also be joined by link words, as in:

> (1) En Grande-Bretagne, le gouvernement aide l'éducation supérieure en payant pour les jeunes à aller aux établissements d'éducation supplémentaire. Aux Etats-Unis, par contraste, ce sont les jeunes eux-mêmes et leurs familles qui doivent payer, alors ça laisse le gouvernement avec plus d'argent à dépenser sur l'éducation. (EF)

There can clearly be, therefore, no direct relation between syntactic complexity and 'linking density'. A higher index of syntactic complexity may, however, reflect the amount of **syntactic linking** in a text. And syntactic linking is linking with a difference: with juxtaposition the type of link is not made explicit and all the propositions are on the same plane; with link words the semantic nature of the links is made clear but there is still no hierarchy; subordination on the other hand can combine explicit link and hierarchy, in the sense that one proposition is seen in relation to another, as the following example illustrates:

> (2) Although - as the case of Germany shows - economic performance is not always directly related to the age at which people leave the educational system, the perpetuation of such short-sighted policies can only bring economic disaster upon our country in an age which should be one of optimism. (EL1)

Whereas (1) provides adequate linking with *par contraste* and *alors* but no hierarchical structuring of its three propositions, (2), thanks to a two level dependency structure, makes it clear that the extract is not about Germany, nor economic performance, but about previously described policies concerning higher education.

A way of looking at this creation of depth or relief in text is in terms of the second in our list of textual functions: **foregrounding/ backgrounding**. Following Tomlin (1985:89), we characterise **foreground information** as 'information which is more central or salient or important to the discourse theme', which bears 'a stronger thematic relation of significance to its superordinate discourse theme'. **Background information** is 'information which elaborates or

develops foreground information', which bears 'a weaker thematic relation of significance to its superordinate discourse theme'. It is stressed that the foreground/background distinction is not to be seen as a simple binary concept but as a continuum. In a study of three types of narratives based on a video recording, Tomlin finds that a great majority of dependent clauses (regardless of their absolute frequency, which is determined by stylistic factors) code background information, while main clauses regularly code foreground information.

Beyond the possibility of taking T-unit length as a rough index of 'relief-creation' in text, these findings led me to look at my data from a slightly different angle: is there a difference in the way syntax is used to code foreground and background information in native and non-native texts? The task designed for the data collection specifically requested subjects to combine two textual activities - ie *argument* backed up by *commentary* (cf statistical data provided) - the commentary element being as it were informationally embedded in the argument. Following Tomlin's suggestion that main clauses and dependent clauses represent different types of information, two predictions were formulated:

- main clauses (and simple sentences) could be expected to be used preferentially for:

  (a) 'framing' statistical data;
  (b) stating opinions;

- straight reporting of statistics would be more likely to occur in dependent clauses.

To start with **framing**, I counted as **framing** expressions phrases signalled by *on/nous + noter, remarquer, constater*, etc, or by impersonal constructions such as *il apparait, il convient de noter, it would appear*, etc:

> (3) Il apparait également que c'est aux USA que le pourcentage du PNB consacré aux dépenses publiques dans l'enseignement supérieur public est le plus fort. (F)

(4) En étudiant ce tableau, on s'aperçoit qu'il existe une différence importante entre l'âge moyen à l'entrée... (F)

Opinion expressions were those signalled by *je pense, I think, believe*, etc; *selon moi, à mon avis, in my opinion*, etc; by modals such as *falloir, should*, etc; or by lexical items such as *essential, intolerable*, etc.[1]

(5) ...selon moi la démocratisation de l'enseignement supérieur est un facteur de progrès pour un pays. (F)

(6) I think it would be a great shame to let the standards lapse now. (EL1)

(7) Par contre, il nous faut augmenter nos dépenses afin d'améliorer le système d'enseignement. (EF)

Table 1 shows the results of the analysis of main clauses in terms of **framing** and **opinion**. As hypothesised, these two functions occupy a large percentage of main clauses in all three data sets, but considerably more so in L1 (French and English) than in L2. The table also shows that the distribution of **framing** to **opinion**, similar in F and EF, differs greatly in EL1, with a much higher proportion of **opinion** main clauses. It is noteworthy that the non-native texts, with fewer main clauses involved and a higher proportion of **framing** clauses, express opinion only half as often as the native texts by the same writers.

Table 1: Percentage of 'framing' and 'opinion' main clauses

			% framing/opinion	
F	framing: 17.4% opinion: 27.3%	44.7%	38.9%	61.1%
EL1	framing: 10.5% opinion: 37.0%	47.5%	22.1%	77.9%
EF	framing: 13.9% opinion: 19.2%	33.1%	42.0%	58.0%

Table 2 throws a different light on these figures: it shows the occurrence of straight - ie not interpreted or evaluated - accounts of elements from the statistical data:

Table 2: Straight account of statistics

	% in main clauses	% in dependent clauses	% of all T-units
F	45.6%	54.5%	31.8%
EL1	43.8%	56.2%	16.2%
EF	57.5%	42.5%	25.4%

The smaller proportion of EF main clauses expressing opinion may partly be accounted for by the fact that a good deal more are just reporting figures. It is also interesting to see, although this may be more closely related to the fascinating field of contrastive rhetoric than to the immediate concern of this paper, that the English subjects make considerably less use of the statistical data in their native texts than in their French texts, and that the French writers cling the most to the data provided (cf right hand column of table 2). As far as the study of non-nativeness is concerned, the figures in tables 1 and 2 converge to suggest that the three data sets exhibit such a different balance of fact vs opinion that they may legitimately be seen as representing different text types, a very important consideration indeed in the pedagogy of written communication.

I shall now return to syntactic dependency, and illustrate its link with the third textual function under consideration - **topic selection and continuity** - with four examples of accounts of statistics:

> (8) Il se passe le phénomène inverse en Grande-Bretagne, pays où le taux de scolarisation de la population âgée de 19 à 24 ans est le plus faible, où l'âge moyen à la sortie du système d'enseignement est le plus bas, où le taux d'accroissement des

dépenses publiques relatives à l'enseignement superieur est le plus faible mais où au contraire le pourcentage du PNB consacré aux dépenses publiques dans l'enseignement supérieur public est assez élevé en comparaison avec celui des autres pays. (F)

(9) The figures clearly reveal that the British education system suffers the most in relation to the annual rate of increase of public expenditure, the British figure being only 8.12% compared with Sweden and France, for example, whose governments spend an annual increase of 28.9 and 24.7% on higher education respectively. (EL1)

(10) Britain has the lowest percentage of 19-24 year olds in full time education. This means that the majority of people in Britain do not go on to full time further education, and this is backed up by the figures in the second table, since the average leaving age in British schools is 17.7 years. (EL1)

(11) En Grande-Bretagne moins de la population âgée de 19 à 24 ans font l'enseignement supérieur qu'en France (9.3% en Grande-Bretagne, 12.3% en France). Mais le pourcentage du P.N.B. consacré aux dépenses publiques dans l'enseignement supérieur est plus grand en Grande-Bretagne (0.89%) qu'en France (0.43%). Au Japon 14.7% de la population âgée de 19 à 24 ans ont été admis au système d'enseignement supérieur, mais le pourcentage du P.N.B. n'est que 0.38%. (EF)

Whereas in (8) and (9) the figures are integrated, motivated by the point the writer is making, and backgrounded in relation to this point, in (10) the figures precede the point ("This means..."), and the 'data' sentence is given the same weight as the interpretive one. In (11), which is a complete paragraph, the succession of elements from the statistical tables in simple sentences leaves the reader with a great deal of inferring to do in order to guess the writer's point. More specifically, the reader may wonder whether the paragraph is about Britain or Japan, about numbers in higher education or percentage of GNP: the succession of simple sentences has an impact on a fundamental aspect of textual functioning: **topic selection and continuity**. The link between syntactic

complexity and topic selection and continuity will not be further investigated here, but it seems reasonable to conjecture that lower syntactic complexity may lead to greater proliferation of topics and therefore to more difficult reader-identification of main discourse topics[2].

The notion of topic continuity is basic to an understanding of the first of the **marked structures** under consideration in this study: the **passive**, which is seen as a way of matching syntactic structure and information structure, by allowing the topical element to occur in the grammatical function most normally associated with sentence topic:

(12) The United States' figures represent the longest time for pupils to stay in the education system: that is 16.7 years on average are spent by a pupil at school. (EL1)

(13) This is reflected in the second set of statistics showing that the British school leaving age is on average much lower than elsewhere. (EL1)

(14) In the majority of the countries in the tables, an increased investment in education can be seen. (EL1)

(15) ...the system of higher education ...can be of great benefit to the individual. On entering ..., he or she is being afforded the chance of self-fulfilment... (EL1)

(12) is particularly striking. Its oddness seems to indicate that ensuring topic continuity was very much the primary concern of the writer (a native speaker). The use of the passive as a topic continuity device is found in the three sets of data, with the following frequencies (calculated per 100 words):

F 0.36/100 words

EL1 1.28/100 words

EF 0.18/100 words

The abundance of passives in EL1 as opposed to F and EF prompts several remarks and questions. In contrastive terms, it suggests that the topic continuity devices used by English and French writers are markedly different, and points to specific directions for research which should be of a contrastive nature, both intra-lingual - different text types - and inter-lingual - eg French/English. For the present study of non-nativeness, the figures indicate that there is no simple transfer of items reflecting textual strategies - even when, as is the case of the passive - it is linguistically possible. It could be hypothesised that these advanced learners are in fact more influenced by the target 'norm', represented by the F texts, than by their L1, to the extent that they overshoot that norm considerably, using passives only half as much as the F group. The question for the characterisation of non-nativeness then is: what topic continuity devices - if any - do English writers use in their French texts to compensate for this strikingly low use of passives (half as many as F, just over one seventh of the EL1 frequency figure)? The beginnings of an answer may perhaps be suggested by the inordinately frequent, and often textually inappropriate, use of **on** in EF texts. **On** seems to be perceived by EF writers as an equivalent for their L1 passive, but while it does share certain features with that construction, it is no good as a topic continuity device. This will be the subject of further research, not to be reported in this paper.

The second and last marked structure to be examined is a group of constructions occurring in both English and French and known as **clefts**. The sub-classification into two main types holds for both languages, with great structural similarities. Occurrences of the type illustrated by example 16, generally called **pseudo-cleft**, though common to all three data sets, are too few to permit any generalisation and will not be analysed:

(16) What remains to be seen is whether the government will change its policy. (EL1)

The other type, which will simply be called **cleft**, is seen in examples 17 to 19 below:

(17) The figures clearly reveal that the British education system suffers the most in relation to the annual rate of increase

of public expenditure... compared with Sweden and France... Again it is the British education system which is contracting the most in view of the amount of money consecrated to it. (EL1)

(18) En 65 et 70, c'est en France que le pourcentage est le plus bas, ...; mais en 74, c'est l'Italie qui détient le plus faible pourcentage, suivie du Japon; on peut noter que c'est le Japon qui connait la plus faible progression... (F)

(19) En ce qui concerne le taux moyen d'accroissement annuel des dépenses publiques relatives à l'enseignement supérieur public de 1965 à 1970, il est à voir que ce sont la Suède et la France qui dépensent le plus d'argent pour l'enseignement supérieur et que c'est la Grande-Bretagne qui dépense seulement 8.12%... (EF)

Most remarkable about clefts is their frequency in F, which is three times greater than in EL1 and twice the figure for EF (table 3). Again there appears to be a clear difference between English and French native texts, which warrants a closer look at what clefts actually do in texts.

Table 3: Clefts

	Frequency per 100 words	% informative presupposition	% stressed focus
F	0.40/100 w	91%	9%
EL1	0.13/100 w	54%	46%
EF	0.20/100 w	81%	19%

A cleft is generally characterised as a focus-marking device, with the most precise definition given in Lambrecht (1984:8-9): 'a syntactic device by which a constituent whose referent plays the pragmatic role of focus in a proposition but whose normal position does not mark it as focal is allowed to appear in the preferred post-verbal position for focal ele-

creation of an additional or "auxiliary" clause containing the copula être, in whose free focus position the constituent can "appear". The corollary of a focal IT/C'-clause, however, is not necessarily an informationally low-status WH/QU-clause. Prince (1978) judiciously distinguishes between **stressed-focus cleft** and **informative presupposition cleft**: in the first, the WH-clause represents information which the speaker assumes the hearer knows or can deduce, but is not presumably thinking about; in the second, the WH-clause represents information which the speaker takes to be a known fact, though definitely not known to the hearer. This distinction proved essential in the contrastive analysis of my three sets of data, as it transpired (table 3) that not only do F subjects use three times more clefts than EL1 subjects, they also use them differently: whereas most F clefts are of the informative presupposition type (91%), almost half of EL1 clefts (46%) are of the stressed focus type (cf example 17). EF clefts mostly resemble F clefts, with a large number used to pick out a referent - usually a country - in the IT/C'-clause, and indicating its extreme position in a pre-mentioned scale in the WH/QU-clause (see examples 18 and 19). Contrasting with these 'local' clefts, the EL1 stressed focus clefts seem to perform the function identified by Borkin (1984) of foregrounding the whole clause within the surrounding discourse.

As regards clefts, it would therefore appear that it is frequency rather than use that distinguishes EF from F, with again no evidence of direct transfer from L1 to L2 in the English subjects. But many aspects of the textual function of clefts remain to be clarified, and a larger number of instances would be needed to substantiate these initial findings.

Conclusion

This exploratory study of syntax and textual function in native and non-native texts allows the following tentative conclusions: lower syntactic complexity may be related to lower linking density, to less topic selection and looser topic continuity, three important factors in the establishment of explicit coherence; more importantly perhaps, the different ways in which the groups used syntax to foreground and background elements within the sentence and within the text as a whole led them to produce quite distinct types of texts, 'doing different things'. Finally, there

appeared to be little direct transfer of text-building devices from L1 to L2, and as far as passives and clefts were concerned, non-native texts were closer to F than to E1.1 in terms of use, but differed considerably in terms of frequency. The evidence of considerable differences between text-building devices used by native writers of French and English calls for detailed contrastive research.

The exploratory nature of this study must again be stressed, and the suppression of individual differences must be seen as the result not of a research principle but of lack of space. The pedagogical message emerging from these tentative conclusions is that any correction of a learner's written syntax, or any form of intervention in it, which does not take into account the textual function of particular syntactic choices is bound to be insufficient, and may lead to the production of textual infelicities.

Footnotes

1. The unavoidably impressionistic nature of the list of markers for framing and opinion expressions points to the danger of circularity ever present in the field of functional syntax (cf Tomlin 1985). All I can say is that relevant instances seemed easily identifiable.

2. cf Lautamatti 1978 on how texts which have had their syntax simplified for pedagogical reasons can prove to be more difficult to read.

Bibliography

BORKIN A. Problems in form and function. London: Ablex (1984)

HUNT KW. Syntactic maturity in schoolchildren and adults. Monograph of the Society for Research in Child Development, Serial No 35, Chicago (1970)

LAMBRECHT K. On the status of SVO sentences in French discourse. Paper presented at the Symposium on Discourse Relations and Cognitive Units, University of Oregon, Eugene, June 1984. To appear in R Tomlin (ed) Coherence and grounding in discourse. Amsterdam: J Benjamins (page numbers correspond to manuscript)

LAUTAMATTI L. Some observations on cohesion and coherence in simplified texts, in JO Ostman (ed) Cohesion and semantics. Reports on Text Linguistics. Abo (Finland): Abo Akademi Foundation (1978)

LAKOFF R. The pragmatics of subordination. Berkeley Linguistics Society, 1984, vol 10, pp 481-92

PRINCE E. A comparison of WH-clefts and IT-clefts in discourse. Language, 1978, vol 54, no 4, pp 883-906

REINHART T. Conditions on text coherence. Poetics Today, 1980, vol 1, no 3

TOMLIN R. Foreground-background information and the syntax of subordination. Text, 1985, vol 5, no 1-2, pp 85-122