Theme-Rheme Structure and the Article in Sorbian

Gunter Schaarschmidt, Department of Slavonic Studies
University of Victoria

Sorbian, as spoken today in the southeast corner of the German Democratic Republic, is the remnant of a once territorially much larger group of early Slavic dialects which spread between the Elbe and Neisse rivers from the sixth to the twelfth centuries. After the twelfth century, rapid German colonization progressively decimated the Sorbian language area with resulting bilingualism, first among the educated, later as a massive phenomenon. This long period of bilingualism has left unmistakable traces in all domains of the Sorbian linguistic system, one of them being the acquisition of an article, a feature otherwise alien to Slavic grammatical systems. There is virtually no specialized literature on the syntax of the article in Sorbian, perhaps because its usage is restricted to the dialectal and colloquial spheres as well as to an abundant church literature since the sixteenth century. There is no question, however, that the degree and kind of integration of the article in Sorbian is of considerable interest to the general linguist and the sociolinguist alike. From a general linguistic point of view, the intersection of variables resulting from the addition of an article to other syntactic means of expressing definiteness can shed more light on the general theme-rheme structure of languages. For the sociolinguist, the use of the article by different generations of Sorbian speakers can provide valuable data about the nature and evolution of bilingualism.

The use of the article in the Indo-European language family appears to be a syntactic innovation. As Lakoff (1972:174) states:

In the earliest stages of most languages of this family, there were no articles (though demonstratives did exist, and the morphological shape of the articles developed from them). In some of the languages only the definite article developed; in others, both. In one or two, like Russian, neither has developed. But the trend is to a system containing both a definite and an indefinite article.
Thus, the demonstrative pronoun *ille 'that' in Latin developed ultimately into the definite article *le in French; the numeral unus 'one' resulted in the indefinite article un 'a'. In English, a weakened form of that eventually yielded the definite article the, while one developed into a(n).

Among the Slavic languages, only Bulgarian and Macedonian have developed a fully integrated grammatical article, although Proto-Slavic had a way of marking definiteness morphologically by the postponed element -i, originally an anaphoric pronoun (Vaillant 1942). The use of an article has also been noted in some styles and dialects of other Slavic languages (see, especially, Krámský 1972). The use of the definite article and, to a lesser extent, an indefinite article, is found in Sorbian in the earliest written sources from the sixteenth century. The definite article takes its form from the original demonstrative pronoun *t'ob 'this', as in tón (Lower Sorbian ten), ta, to (cf. Russian tot 'that'). The indefinite article, if used at all, is homophonous with the numeral jedyn (Lower Sorbian jaden) 'one'. Perhaps to compensate for the decrease in deixis of tón, Sorbian has developed new lengthened versions for the demonstrative pronoun, such as tutón or tóne 'this'. As distinct from the language of the written documents from the sixteenth well into the twentieth century, the use of an article is not recommended in the contemporary Sorbian literary languages. The article is widely used, however, in colloquial Sorbian and in regional dialects (see, for example, Šewc-Schuster 1968:124 and Janaš 1976:203).

There can be little doubt that the article arose in Sorbian under the influence of German since language contact and bilingualism of some form or another must be assumed for a very early period in an area roughly east of the Saale and west of the Neisse rivers. It is less clear, however, to what extent the use of an article in Sorbian mirrors that of the article in German, and how the acquisition of an article was integrated with other means of expressing definiteness, specifically with the theme-rheme structure of sentences in Slavic languages. Trying to answer this question provided the incentive for the present investigation which continues previous research.
into the degree of integration of the article in the grammatical system of Sorbian (Schaarschmidt 1983).

The method consisted of contrasting sentence by sentence Sorbian translations of German texts, and German translations of Sorbian texts, as well as bilingual versions of stories. In order to have some meaningful way of evaluating the contrastive results, the 'Markedness Differential Hypothesis' (Eckman 1977) was applied to determine the relative ease/difficulty of the integration of the definite article in the speech of Sorbian bilinguals.

It was found, as suspected, that the Sorbian definite article was not used idiosyncratically or in free variation, except possibly in some small subpart of the data analyzed, and here variables, such as style and emotive-expressive functions, may play a role. It was also found that there was no one-to-one correspondence between the use of the article in Sorbian and that of an article in German. This is not surprising in light of the fact that the use and functions of the article vary considerably in those languages that possess this overt category (see, especially, Naylor 1983). The deviations in article usage between Sorbian and German were furthermore not random, but followed from a hierarchy of markedness (in the typological sense) with respect to the integration of the article in Sorbian. Specifically, the more marked a given function of the article in German, the more often the article was found to be absent in the Sorbian text, and vice versa.

At the lower end of the spectrum, Sorbian generally lacked the article when its function in German was that of a purely grammatical marker, a relatively marked function in terms of typological markedness. This was the case especially with prepositions, for example (SI 1:165,169):

(1) (a) Die Kinder fahren mit dem Bus in die Schule. 'The children go to school by bus.'

(b) Te dźeči jëdëjøa z busom do šule.²

At the other end of the hierarchy, the article was found to be virtually
obligatory in Sorbian when its meaning was close to that of the demonstrative pronoun, a relatively unmarked function (SI 2: 197,204):

(2) (a) Wir sahen uns zuerst die Räume, die grossen Häuser mit den Bildern und mit der künstlerischen Arbeit an.
'At first we viewed the rooms, the large houses with the pictures and the art work.'

(b) My wobhladachmy sebi najprjedy te rumnosće, te wulke chêže z tymi wobrazami a z tym wumôskim dzêwom.

In between these two points in the markedness hierarchy, there are several other variables determining the use or non-use of the definite article in Sorbian. At one point, there is indeed a somewhat grey zone where the use or non-use of the article appear to be in free variation (SI 2: 175,180):

(3) (a) Na ja, was nun? Gehst du in die LPG?
'Well, what now? Will you join the co-op?'

(b) Nó haj, što nêk? Póndžeš do LPG?

(4) (a) Ich habe mich entschlossen und bin in die LPG gegangen.
'I made up my mind and joined the co-op.'

(b) Sym so rozsadžwa a sym šwa do toh' LPG.

Even in this example, the context and stylistic factors may be responsible for the choice of toh' (< toho) in (4 b).

Other variables include generational differences: the article is used more frequently by speakers of the older generation than by those belonging to the younger generation, a difference which is sufficiently wide to be statistically significant. Thus, in some dialect texts, the article in all those cases where it is possible shows twice the frequency for an older-generation speaker born in 1886 (the data were recorded in the late sixties and early seventies), as compared with two younger-generation speakers born in 1929 and 1938, respectively (see, especially, Protze 1974:409-10). Thus, descriptions of the use of the article may vary considerably depending upon the age group of the dialect speakers interviewed.
Last, but not least, the article in German is multifunctional, and it is unlikely that all of its functions can be subsumed under one common denominator or invariant (see especially, Behaghel 1923:31-134). The multifunctional nature of the German article is reflected in Sorbian usage as well, although not all of the functions in German find a reflection in Sorbian and, vice versa, Sorbian shows a few functions which German lacks.

To cover all of this territory in one paper would be an exercise in futility. The present paper will therefore limit itself to the question of the extent to which the article in Sorbian has taken over the functions of other syntactic means of expressing definiteness, especially word order, and which are typical of an article-less language where definiteness is said to be expressed covertly (Birkenmayer 1979:149). Methodologically, the best way to investigate this question seemed to be to contrast colloquial or dialectal usage of the article with literary Sorbian since the latter nowadays prohibits the use of an article.

Although it seems clear that the article alternation in a text cannot be fully equated with the theme-rheme progression (Nikolaeva 1979:170), one of the basic functions of the definite article is to denote known, old information, while the indefinite article denotes unknown, new information. The same information in a text can be conveyed by word order, for example, in Slavic languages where known, old information precedes unknown, new information.

(5) Mal'čik 'prišel.
 'The boy came.'

(6) Prišel 'mal'čik.
 'A boy came.'

The preceding readings will only work given an identical intonational configuration in both sentences, as indicated here by the apostrophe before the second element in each Russian sentence.

The generalization made above can be expressed in the form of two ordering
rules applying to linearly unordered structures.³

(7) \([\text{DEFINITE, MASCUILINE, HUMAN \{BOY\}}, \text{CAME}]\)
\[
\begin{array}{c}
1 \\
2 \\
\end{array}
\rightarrow 1 \& 2

(8) \([-\text{DEFINITE, MASCUILINE, HUMAN \{BOY\}}, \text{CAME}]\)
\[
\begin{array}{c}
1 \\
2 \\
\end{array}
\rightarrow 2 \& 1

The feature \([\text{DEFINITE}]\) in these rules is perhaps best understood as a symbol which may stand for a number of variables including the traditional notion of definiteness. To define this symbol here would go beyond the scope of the present paper (but see Chvany 1983). In any case, the basic functioning of rules of the type (7) and (8) has been described for Russian (Pospelov 1970 and Fursenko 1970) and seems to be operative in other Slavic languages as well. Thus, in Polish, a discourse consisting of two sentences (9) can become ungrammatical if the order of elements is changed in one of the two sentences (10):

(9) W pokoju siedziała dziewczyna. Wszedł chłopiec.
"There was a girl sitting in the room. A boy entered."


The definiteness of \(\text{chłopiec} \) 'the boy' in sentence (10) is acceptable here only if some other discourse segment with prior mentioning of the boy is understood (see also Szwedek 1973:206).

In Sorbian, the rules for ordering elements work in essentially the same way as in Russian and Polish, if we ignore the tendency to place the verb at the end of the sentence, not only in subordinate clauses, as in German, but, as distinct from German, in main clauses as well. Thus, in Sorbian, an indefinite noun phrase will generally follow a definite one (see, in this report, Sewc-Schuster 1976:107-19):

(11) W zahrođce róža kćeje.
"A rose is blooming in the garden."
Since one cannot really talk about definiteness or the definite article without also saying something about indefiniteness or the indefinite article, a few words will be in order here about the indefinite article in Sorbian.

As Civ'jan (1976:182-92) has shown in his study of Modern Greek, the indefinite article is a marked category in the sense that it carries advance information. Therefore, its occurrence is limited (1:3 in the texts studied as compared with the occurrence of the definite article). The indefinite article marks the appearance of every new object in a discourse and is thus a kind of boundary signal for text segments which are more or less semantically complete units. The term 'new object' is defined here as a singular object not known contextually in the progression of a text. Plural objects, proper names, and objects known consitutionally, although new in a given text, do not require an indefinite article.

Perhaps this can be best illustrated by the following Sorbian popular tale (SV 9-10):

(13) Na jenym twarjenju do kamjenja wurubany jedyn mnich.
    On a building into masonry carved a monk

Hay tón šćen tam džens widžec.
Yes he still there today to-see

Te twarjenjo je slušato mnicham. Něk cyrkej sluša
The building is belonged to-monks Now church belongs
ewangelskim, ale mjeneje so pšec "Mniša cyrkej".
to-Protestants but is-called always Monk's church

Te su tam wjac moli zamazali a z pucom
It are there several times smeared and with plaster
zamjetali. To šitko zaso wotpada a tón mnich
covered This all again fell-off and the monk
tam zaso jo.
there again is
There is a monk carved into the masonry of the building. Yes, one can still see him there today. The building belonged to the monks. Now the church belongs to the Protestants, but it is still called "Monk's Church". It was bedaubed several times and covered with plaster. All this fell off again, and the monk is there again.

In general, Sorbian avoids the use of an indefinite article, and where it is used, its function cannot always be clearly differentiated from that of the numeral jedyn 'one'. However, in the above text, the function of jedyn is clearly that of an indefinite article introducing new, unknown information (jene twarjenjo 'a building' and jedyn mnich 'a monk').

Where a new noun phrase in a given text refers to an object which is consituationally known, the use of the indefinite article is not required. This is most commonly the case in folk tales where known mythical figures act like proper names, such as wódnį muğ 'water sprite' in the beginning sentences of two Sorbian folk tales, with (15) being Lower Sorbian and (16) Upper Sorbian (SV 61 and 25, respectively):

(15) Wo Borkowach bydlii wodny muğ ...  
     'In Burg there lived a water sprite ...'  
(16) W Komorowje boł wódny muğ ...'  
     'In Commerau there was a water sprite ...'

After the introduction of the two new objects twarjenjo 'a building' and mnich 'monk' in (13), they are then referred to by the definite article (te and tón, respectively), or by the anaphoric pronoun which in dialectal speech has the same form as the definite article. In this pronominal function, there is little fluctuation in the texts studied, and Sorbian dialectal usage shows complete agreement with the German usage of the article der (die, das) as an anaphoric pronoun, while literary Sorbian prescribes wón, wona, wono (wone). The assumption made in a number of linguistic studies that articles and pronouns are merely superficially distinct manifestations of a single underlying category works particularly well for Sorbian where, dialectally, articles and pronouns are not even superficially distinct (see also Postal 1966 and Sommerstein 1972). A rule, such as (14a), which deletes...
the included, specific subset of a generic nominal under conditions of referential identity in some context, expresses this generalization for article/pronoun creation. A lexical rule (14a) will then spell the generic features as either \( \text{ ton } \), for dialectal speech, or \( \text{ won } \), for the literary language.

\[
(14) \begin{array}{l}
\text{(a)} \quad \left[ \left[ \text{MASC, HUMAN [BOY X]} \right] \left[ \text{SAME, MASC, HUMAN [BOY Y]} \right] \right] \\
\text{(b)} \quad \left[ \text{MASC, HUMAN } \right] = \begin{cases} \\
\text{ton} & / \text{Dialect} \end{cases}
\end{array}
\]

In languages which have both variable order of the type discussed in (5) and (6) and a system of articles, the alternation definite article/indefinite article would seem to be redundant. Thus, in German, the sentences (17) and (18) read with a normal, non-contrastive intonation, mark the known/unknown alternation by means of both word order variation and article alternation.

(17) In der Stube sass ein Junge.  
'A boy was sitting in the living room.'

(18) Der Junge sass in der Stube.  
'the boy was sitting in the living room.'

Since Sorbian, like other Slavic languages, can use word-order variation much more freely than German for expressing the known/unknown distinction, one wonders whether the use of an article is not simply a redundant concomitant feature of such variation in Sorbian dialects, or whether word-order variation has become more restricted. Our preliminary analysis of the dialect data shows that rather the opposite is the case, i.e., the addition of a definite article would seem to allow a broader range of word order variants than is otherwise possible in literary Sorbian. It must be emphasized, of course, that the present investigation is just the beginning of a more comprehensive examination of this problem. It is very important, for example, to include the variable of intonation in a more comprehensive study. Nevertheless, in the two examples below, a comparison between dialect data and the literary standards of Upper and Lower Sorbian would seem to indicate that, ceteris
paribus, the choice of a definite article allows an additional word order variant not usually found in the given contexts in the literary language.

Consider first the following passage from Upper Sorbian dialect data (19b) contrasted with the same passage (19c) in the literary language (S1 1:139,144):

(19) (a) In alter Zeit sind die Leute zu Fuss auf die Wiesen gegangen ... dann kamen die Räder auf, dann sind wir mit den Rädern dorthin gefahren ... und jetzt fährt auch der Bus in die Stadt, vier oder fünf Mal jeden Tag. Da kann man immer mit dem Bus fahren.

'In the old days people went to the meadows on foot ... then bikes appeared, and we went there by bike ... and now there is also a bus going into the city, four or five times a day, so one can always go by bus.'

(b) Wstarych časach su ludži cufus chodžili na řuki ... pon su pribišli te kolesa, pon smy z kolesami jėzdžili tam ... a nět dže tež bus do města štyri o pjač moli koždy džen. Da mōžeš pľecz z busom jėzdžić.

(c) Wstarych časach su ludžo pěši chodžili na řuki ... potom su kolesa pribišli, potom smy z kolesami jėzdžili tam ... a nět dže tež bus do města štyri abo pječ kroc koždy džen. Tuž mōžeš pľecz z busom jėzdžić.

It should be noted that auxiliary verb forms, such as su 'are' in (19b) and ordujo 'will be' in (20b), are not counted in determining the theme/rheme positioning in Sorbian. Thus, in (19b) the noun phrase ludži 'people' is considered to be preverbal, as is ta gus 'the goose' in (20b).

The phrase in question in (19b) is su pribišli te kolesa 'the bikes appeared' which corresponds to literary Sorbian su kolesa pribišli. The noun kolesa 'bikes' is in this context understood to be known by contrast since the preceding sentence made reference to pěši 'on foot'. The postverbal positioning of te kolesa appears to be a direct consequence of the use of a definite article here, thus nullifying the normal theme/rheme positioning. That positioning is also possible with the article, i.e., su te kolesa pribišli, without any apparent difference in meaning.
The same situation exists in the dialect passage (20b), taken from Lower Sorbian (SD 3:12-13):

(20) (a) Und vor der Kirmes muss noch der Sand gefahren werden ... und dann wird auch die Gans geschlachtet ... und ein Korn wird auch gekauft ... und am Sonnabend dann werden die Kuchen gebacken.

'And before the kermis the sand must still be hauled ... and then the goose will be slaughtered ... and a corn schnapps will also be bought ... and then on Saturday the cakes will be baked.'

(b) A prjoz teje kjarmuše yšći musy byš ten pjesk wezony ... a pen teke yšći ordujo ta gus zašlachtowana ... a jaden korn ordujo teke kupjony ... a sobetu pen ordujo pjacone te mazance.

Here the extra word order variant ordujo pjacone te mazance 'the cakes were baked' stands out even better due to its contrast with preverbal constructions, such as musy byš ten pjesk wezony 'the sand must be hauled' and ordujo ta gus zašlachtowana 'the goose will be slaughtered'. The existence of this typical rheme order in the same context with theme orders shows that the use of the definite article is clearly not a redundant feature in Sorbian, but one which in fact cancels out the usual theme-rheme sequencing.

Conversely, the definite article can apparently be omitted where preposing already indicates that a noun phrase is the theme, for example, ludži in (19b). The data are not sufficient to advance this as a general hypothesis, but it seems that a rheme may not necessarily have to be postposed when the overt indefinite article is present, e.g., jaden korn ordujo teke kupjony 'a corn schnapps will also be bought' instead of teke ordujo kupjony korn, provided that jaden is indeed the indefinite article here, and not the numeral 'one'.

The phrase nět dźe teš bus do města 'now there is also a bus going into the city' in (19b) is somewhat problematic because bus seems to be indefinite here, yet the German version has a definite article. Replacing der Bus 'the bus' with ein Bus 'a bus' does not seem to make any difference here, at least
for this writer, so perhaps the English translation is a more accurate rendering of the Sorbian text. In any case, a phrase with ten bus 'the bus' preposed, as in nět ten bus teř do města dě 'now this bus also goes to the city' has quite a different meaning because the implication here seems to be that a given route was extended to include the city, while in (19b) an entirely new route could be involved.

In summary, the use of an article in Sorbian does not appear to be a mere redundant addition to the covert means of expressing definiteness, but actually allows more flexibility in the choice of variants for the sequencing of syntactic elements in terms of the theme-rheme progression in discourse. It may well be the case that these dialectal variants, which are not normally found in the literary language, carry additional meanings and functions which, when combined with intonation, make the syntax of dialectal and colloquial Sorbian very different from that of literary Sorbian. The answer to this question must be left for a future study.

It will be appropriate to conclude this paper on a sociolinguistic note concerning the relationship between the use of the article in Sorbian and the degree and kind of Sorbian-German bilingualism. In a study of the Lower Sorbian language, Juro Krygar (1956:57) notes in 1675 that 'năša řec znajo pši wěćownikach artikl' (our language uses an article with nouns). In a footnote, the editors state that the educated of that time acquired Sorbian on the basis of German, which explains their desire to place an article before nouns. What seems to be implied in this note is that article usage is possible only when Sorbian has ceased to be a first language for bilingual speakers and is therefore learnt via German. That statement also seems to imply that those Sorbs who were uneducated learned Sorbian first and German second, or not at all, and thus did not use any article. It is of course impossible to provide any direct verification of this claim about language use in the seventeenth century, except perhaps through a study of tales, but there appears to be some indirect evidence in favour of this statement.

The widespread use of an article in colloquial Sorbian and the dialects
among the older generation of speakers interviewed in the 1960's would seem to be in line with the view that, among the literate Sorbs (and most of them were), Sorbian was learned through German. In fact, the entire Sorbian language area had been shrinking considerably since the late nineteenth century, with some dialects becoming completely extinct by the 1930's, thus making many Sorbs in effect monolingual, i.e., German-speaking.

After 1945, when Sorbian had begun to be taught intensively in schools, with some schools using only Sorbian as a means of instruction, the language again acquired a status close to that of a first language, if not that of a first language, for most of the speakers born in the late twenties and the thirties. As a result, less mixing occurs, and the younger generation uses the article with considerably less frequency than the older generation. The prohibition against the use of an article in the literary language must thus be considered not a purist trend in its entirety, but a reflection of the changing usage by the new 'native' speakers of Sorbian.
FOOTNOTES

1 Actually, "Sorbian" is the generic term for what some linguists consider two distinct languages, viz. Upper Sorbian with appr. 65,000 speakers and Lower Sorbian with appr. 25,000 speakers. For the purpose of the present study, this distinction can be safely ignored since the usage of the article shows identical patterns in both Upper and Lower Sorbian in the texts investigated.

2 The transcription of dialect data is that used in the primary sources. Specifically, \(w\) is used for either \(w\) or \(t\) in those dialects where \(t\) has a bilabial pronunciation.

3 The format of the rules given in this paper follows Sanders' equational grammar notation, with square brackets indicating sets of syntactic-semantic features and the ampersand (\&) denoting a linear order relation. See, in this respect, Sanders 1972 and Sanders 1975.
REFERENCES

I. PRIMARY SOURCES


II. SECONDARY SOURCES


Naylor, K.E. 1983. On Expressing 'Definiteness' in the Slavic Languages
and English. In M. Flier (ed.), American Contributions to the Ninth
International Congress of Slavists. Vol. 1: Linguistics. Columbus, Ohio:
Slavica, 203-20.

Nikolaeva, T.M. 1970. Akcentno-prosodičeskie sredstva vyraženiJa kategorii
opredelennosti-neopredelennosti. In T.M. Nikolaeva (ed.), Kategorija
opredelennosti-neopredelennosti v slavjanskix i balkanskix jazykax.

Pospelov, N.S. 1970. O sintaksičeskom vyraženii kategorii opredelennosti­
neopredelennosti v sovremennom russkom jazyke. In T.P. Lomtev and A.A.
Kamygina (eds.), Issledovanija po sovremennomu russkomu jazyku. Moskva:
Izd. Moskovskogo un-ta, 182-89.

Postal, P. 1966. On so-called 'pronouns' in English. In F.P. Dinneen (ed.),
Monograph Series on Language and Linguistics. Washington, D.C., 19,
177-206.

Protze, H. 1974. Typische Erscheinungen deutsch-sorbischer Interferenz unter
Berücksichtigung ihrer sozialen Bedingtheit. In Aktuelle Probleme der
sprachlichen Kommunikation. Soziolinguistische Studien zur sprachlichen
Situation in der DDR. Berlin: Akademie-Verlag, 389-416.


in Sorbian and German. Canadian Slavonic Papers, 25, 147-62.

i morfologija. Budyšin: Domowina.

Sommerstein, A. 1972. On the so-called definite article in English. Linguistic
Inquiry, 3, 197-209.

Szwedek, A. 1973. A Note on the Relation Between the Article in English
and Word Order in Polish. Papers and Studies in Contrastive Linguistics,
2, 212-225.

Vaillant, A. 1942. L'article en vieux slave. Revue des études slaves, 20,
5-12.