

Terje MATHIASSEN

ON CLITICS IN BALTIC

Over the last decades linguists all over the world have taken great interest in problems related to clitics. Among the Indo-European languages which have attracted their interest from this point of view are above all Romance, Slavic, and Germanic. Little attention seems to have been paid to clitics in Baltic. Thus, the contributions listed in the comprehensive 1994 bibliography on clitics by Joel A. Nevis (in cooperation with Brian D. Joseph, Dieter Wanner and Arnold M. Zwicky) are few and – for the most part – marginal and general. In our opinion there is, however, some interesting evidence of clitics candidates in Baltic that deserves our attention, but tracing clitics is not an easy task due to the fact that there can hardly be said to exist any generally accepted definition of this concept. Nevertheless many linguists would subscribe to the following clitic features:

- a) clitics lack stress;
- b) clitics cannot be used in isolation;
- c) the internal order of clitics co-occurring in a string is not free.

Deviations from the first feature can be encountered. Thus, „established“ pronominal clitics in South Slavic may be stressable, cf., for example, Bulgarian *az ne go znam* ‘I don’t know him’ (which is due to the fact that the Bulgarian negational particle has an extrapositional stress which automatically places the stress on the immediately following element irrespective of its status as non-clitic or clitic).

Two basic types of clitics can be discerned. The first is constituted by clitics which have non-clitic counterparts. An example would be the accusative and dative forms *cię* and *ci* ‘you’ of the personal pronoun vs. the longer forms *ciebie* and *tobie* of Polish. The second type is made up of elements which have no free alternative forms like, for example, the interrogative particle *-kö/-ko* in Finnish or the focus marker *-kin/-kaan* in the same language.

Dependent on their position *before* or *after* the *host word* clitics are referred to as *proclitics* or *enclitics* respectively. Clitics which can have both roles are labeled *endoclitics*.

Endeavours have been made to separate clitics from *affixes*. The distinction, however, is not always clear, and this question will not be discussed here. For practical reasons, however, the term *affixal clitics* will be used.

The aim of this article is to investigate certain clitic candidates in Baltic with special emphasis on element order and the motivations for that particular order. The investigation will have both a diachronic and synchronic dimension.

Our analysis will be limited to Latvian and Lithuanian, i. e. East Baltic, since Old Prussian material seems too scanty and precarious with respect to clitics.

Among the clitic candidates of Latvian and Lithuanian are:

- (1) the interrogative particle Latvian *vai*, Lithuanian *ar* (proclitic);
- (2) the deictic (focus) particle *gi* of Lithuanian (endoclitic)¹;
- (3) the negational particle Latvian and Lithuanian *ne* (proclitic);
- (4) the clitic affixes *te-*, *be-*, *tebe-*, *nebe-* of Lithuanian (proclitic);
- (5) the reflexive particle *-s(i)* in Latvian and Lithuanian (originally endoclitic);
- (6) the pronominal element of the pronominal (long) adjective form (originally endoclitic, cf. Old Lithuanian);

(7) the atonic personal pronouns of East-Baltic (*mi*, *ti*) which were endoclitics.

We will concentrate on affixal and pronominal clitics, thus leaving aside (1) and (2).

A Lithuanian verb can have a cumulation of prefixes, e. g. *tebepasiližka* 'still remain(s)', *nebepasiližka* 'no longer remain(s)' which are subject to strict hierarchic rules in terms of relative order in the string. On the basis of examples of this kind we will propose the following two hypotheses, namely

a) that the closer to the root (host word), the greater is the degree of the *grammaticalization* of the prefix,

b) that the position of a prefix in the string reflects the relative chronology of its attachment to the host element.

These hypotheses look so simple that it is almost embarrassing to present them, but as far as we can see, nobody has stated this explicitly with respect to Baltic material before. Furthermore, they seem to be in harmony with the principles of the so-called *natural morphology* (cf., for example, B y b e e, 1985, 4–5) which is a recognized strategy and approach.

First we will try to verify hypothesis a) bearing in mind that grammaticalization „highlights the tension between relatively unconstrained lexical structure and more constrained syntactic, morphosyntactic, and morphological structure“ (H o p p e r, T r a u g o t t, 1993, 1–2). A near to ideal way of proving hypothesis a) would have been a situation where the elements in question could occur both in proclitic and enclitic position in the form of a mirror-image-like structure. Such a structure, however, is found only in the case of the reflexive particle.

¹ Deictic particles are also found in Latvian, see E n d z e l i n, 1922, 541 ff.

(1) With unprefixed verbs the *reflexive particle* has developed more or less into an *ending* whereby it is (optionally) reduced to *-s*, e. g. *liėkas* 'seem(s)'. This process testifies to and underscores its high degree of grammaticalization.

(2) The next element in our example (*tebepasiliėka/nebepasiliėka*) is *pa-* which can be said to function on the borderline between the *lexical* category of *Aktionsarten* (procedurals) and the *grammatical* category of *aspect*. Although the category of *aspect* has not reached the same level of grammaticalization as in many Slavic languages, nevertheless it has – in our view (see Mathiasen, 1996, 119 f.) – been grammaticalized to a considerable degree also in Lithuanian. Thus, it is not surprising that this element occupies the position next to the reflexive.

(3) We are then left with the elements *be*, *ne* and *te*.

Although Fraenkel, 1965, 41, advocates another etymology, the *be-* affix can in our opinion be connected with the verb *būti* and reflect a form of the type *bě* (< **bhwē?*) found in Slavic. If this assumption is correct, the *be-*element of Lithuanian could be interpreted as an original verb auxiliary with some degree of autonomy which later developed into a particle with a certain affinity to aspectuality ('still being'). The short vowel in Baltic (Lithuanian) *be* could be explained according to Leskien's Law.

In Lithuanian the *negational* particle appears before the element or word which is negated. Most frequently, this is a verb form. Thus, the sequence *nebe-* rather than **bene-* is expected. It always occupies the leftmost position in relation to the host word (element). If our hypothesis of *be* as an original auxiliary verb is correct, the synchronically simple form *nebepasiliėka* would reflect the same element order as the compound *nėra* (< **ne + yra*) / *nebuvo pasilikeš (pasilikusi)* where the negation appears before the overt auxiliary of the compound verb form. A further comparison with compound verb forms reveals that whereas constructions like *neėra / neėbuvo pasilikeš (pasilikusi)* are equally possible, the alternative **yra nebepasilikeš (-likusi)* seems unacceptable. This shows that the negational particle occupies a special place in the system. Its leftmost position is a good indication that it has been attached (agglutinated) to the verb later than the other affixes. Another hint in the same direction is its different behaviour in terms of accentuation. Thus, one will say, for example, *pàguli*, but *negùli*.

The particle *be* can co-occur both with *ne* and *te*, but *te* cannot combine with *ne*. The relative order of elements is *tebe*, not **bete*. Can this order be accounted for?

The *te-*particle has an emphatic / rhythmical or delimitative function. Except for delimitativeness, none of these epithets are closely linked with grammaticalization. To be true, the *te-*element shows a very considerable degree of grammaticalization in connection with the *permissive*, but originally it was only a concomitant

(accessory) marker (since the verb was characterized by non-indicative endings). Furthermore, it may be argued that the *te*-marker of the permissive represents a homonym of *te*- in a case like *tebepasiliēka*. From a synchronic point of view *tebe* can probably be looked upon as an indivisible whole which only in a diachronic perspective can be analyzed in two components *te* + *be*. In the latter perspective *te* is likely to occupy the position before *be* since there is evidence of *te* only in preverbal position, cf. our interpretation of *be* as a verb. An alternative would have been a combination like **tebetepasiliēka* with deletion of the second *te*. A sequence of the type **betepasiliēka* is not very likely to have existed, but if it had, *te* could from a diachronic point of view have been interpreted not as an *enclitic* to the (originally more autonomous?) *be*, but as a *proclitic* (candidate) to *(pa-si)liēka*. Finally, it should be stated that the *linear* order *eo ipso* excludes the *logical* possibility, namely that two elements can occur *simultaneously*; one element necessarily has to appear either *before* or *after* another. The order *tebe* can be the result of a generalization which meant the elimination of a competing **bete*. Etymologically *te* seems to have been a deictic particle, derived from the same stem as the demonstrative pronoun / adverb in *t-* (Frankel, 1965, 1071).

According to the second hypothesis, i. e. b) above, the relative order of affixes observed in cases like *tebepasiliēka* / *nebepasiliēka* also reflects the *chronology* of the attachment of the affixes to the verb. That it really is so, seems likely from the analysis carried through under point a). This order is further confirmed through an observation made by Kazlauskas, 1968, 72: „Yra pagrindo manyti, kad dalelytės (*ne-*, *be-*, *te-*, *tebe-*) su veiksmazodžiu į vieningą fonetinį kompleksą suaugo vėliau, negu prieveiksmiai. Šiaurės vakarų aukštaičių ir daugelyje žemaičių tarmių, sangražinės morfemos *si* padėtis priešdėliniuose veiksmazodžiuose priklauso nuo priešdėlio tipo. Po priešdėlių, kilusių iš prieveiksmių, sangražinė morfema *si* šiose tarmėse būva prieš veiksmazodžio šaknį, plg. *nu-si-neš(a)*, o po priešdėlių, kilusių iš dalelyčių (*be-*, *te-*, *tebe-*), ši morfema paprastai atsiduria veiksmazodžio gale (kaip ir nepriešdėliniuose veiksmazodžiuose), plg. *bė-šneka-s(i)*, *tė-neša-s(i)*, *tebe-spārdo-s(i)*, *nė-neša-s(i)* ir t.t.“ Thus, we can read diachronic data in a way parallel to that of counting the yearly rings of a tree or peeling an onion². This simple interpretation seems far more convincing than a theoretical „leap“ of affixes into specific positions according to complicated and completely *ad hoc* rules which derive

² Professor Stundžia has kindly pointed out to me that an experiment with children for the purpose of testing their ordering of the elements in question would have been interesting and instructive. I quite agree, but for the lack of an appropriate test group I have not been able – at least not so far – to conduct such an experiment.

from a special original pattern which can hardly be reconstructed. Nevertheless, the question about the linear structure (or structures?) prior to the cliticization (or tendency towards cliticization) of the above mentioned affixes is a legitimate and important one. One must in our opinion assume a greater degree of autonomy of the elements in question, but we are not able to develop these ideas further. Nor do we have evidence to propose a specific „original“ neutral word / element order which could give an optimal interpretation. Baltic (Lithuanian) seems in the case of prefixes to have developed a kind of *agglutinating* structure which is not characteristic of Indo-European languages; we hold it to be an innovation, rather than an archaism.

Finally, it should be stated that also the conjunctival *b(i)-* (cf. further Old Slavonic *bi-*) in cases like Old Lithuanian *jeib* and *atmintumbime* (see Stang, 1966, 429, and Zinkevičius, 1981, 127) is a good candidate for a copula clitic. Above *be* was commented upon. If even other auxiliaries from *būt(i)* can have been clitics in Baltic like, for instance, in Bulgarian (i. e. *sām, si, e*) and very typically in Polish, I am not able to say, but the enigmatic forms Lithuanian *yra*, Latvian *ir* could have replaced a former clitic **es* (?) = Bulgarian *e*, cf. also *esti* which can have been the non-clitic variant of **es*.

We will now leave the affixal proclitics of the type which has been dealt with above and proceed to the *pronominal* clitics (types 5 through 7 in the above table). As is well known, the reflexive verb particle derives from atonic forms of the reflexive 3rd p. personal pronoun. A parallel development is observed in Slavic and Germanic (Scandinavian). For Proto-East-Baltic Rosinaš 1995, 34 posits **si* for the genitive, dative and accusative as opposed to the tonic (non-clitic) forms in **sev-*. The situation was parallel for the 1st and 2nd p. sg. where the atonic forms *mi, ti* (gen., dat., acc.) had free alternative (non-clitic) counterparts in **men-* and **tev-* respectively. In the dual as well as the 1st and 2nd plural alternative forms do not seem to have existed. As mentioned in the table the atonic forms were endoclitics. In the 3rd p. there was still another endoclititic pronoun which is best illustrated within the framework of the nominal phrase, namely the pronominal element of the so-called pronominal (long or definite) form of the adjective/participle which is cognate to the allegedly free forms *jis* ‘he; it’, *ji* ‘she; it’ which have explicit orthotonic parallels of the type *jisal, jinal*. In the definite adjective forms of Latvian the original situation has become obscured. The original endoclititic character of the reflexive 3rd p. pronoun is still reflected in Lithuanian, cf., for example, *pa-si-lieka* (proclitic) vs. *lieka-s(i)* (enclitic). In Latvian only relics of this situation are preserved, cf. a case like *sa-s-tapt* ‘meet’. The normal situation in contemporary Latvian is encliticization of the reflexive particle both with unprefixated and prefixed verbs, e. g. *liktie-s* ‘seem’,

sa-runātie-s 'chat'. A possible development towards encliticization in Latvian is described by Stolz, 1989, 17 f. The endoclititic position of a non-reflexive 3rd p. pronoun is demonstrated through cases like Old Lithuanian *pa-io-prasta*, *nu-ghie-wargie* (procliticization in participle forms with a prefix) vs., for example, *didimėjime* (encliticization in unprefixated forms), cf. Zinkevičius, 1981, 29 ff. In contemporary Lithuanian this kind of „infixation“ has been abandoned in favour of encliticization (postposition). The common principle both for the reflexive element and the (j)i-clitic is that they could not occur in word initial position. Exactly the same situation is found in the case of the reflexive element in Slavic, cf. Bulgarian. A prerequisite is here that the „word“ is conceived as „phonetic word“ whereby the spelling (separately or together) is of no importance, since it is based on orthographic conventions alone which differ from one language to another.

In contemporary Lithuanian the pronominal character of the reflexive verb particle and the (j)i-element of the long form adjective/participle doesn't seem to be felt any longer. In this sense one can say that these originally atonic pronouns have been lost. Also the atonic forms of the 1st and 2nd sg. of the personal pronoun have disappeared. In Old Lithuanian they were still intact, cf., for example, the proclitic *pa-mimokink* (vs. today's *pamokink mane*) vs. enclitic in unprefixated forms, e. g. *gėlbėkim* (instead of *gelbėk mane* in the contemporary language). The examples are taken from Zinkevičius, 1981, 49.

In Bulgarian one can say, for example, *az sām mu go dal* 'I gave it (e. g. *moliv* 'pencil') to him' with a string of three clitics, first a copula clitic, then two pronominal ones of which the dative clitic obligatorily precedes the accusative. *A priori* there seems to be no good reason for this particular order of the two pronominal objects. They are likely to occur simultaneously, but this, of course, is impossible in a linear order. A careful study of old Lithuanian texts with respect to co-occurrence of two pronominal objects would perhaps be fruitful and rewarding. Care should, however, be taken for possible foreign (Polish) influence. Unfortunately we are not able to prove that the anaphoric *jis/ji* pronoun of Lithuanian represents a former clitic in East-Baltic with a function parallel to that of *go/ja* in Bulgarian, but it can more or less be proved that it has existed in Latvian (through direct evidence in dialects, indirect in the long forms of the adjectives/participles), but it has been replaced by the non-clitic *viņš/viņa* (and *tas/tā*) in the contemporary Standard language. Further its phonological structure is that characteristic of a clitic. An argument in the discussion would also be the extended forms *jisai* and *jinaī* of Lithuanian which could lead one to the assumption that Lithuanian is in a stage of transition towards the same development as Latvian, i. e. a system without anaphoric pronominal clitics.

As already pointed out we have in our analysis not postulated a specific word order in the deep structure from which movements of clitics could have taken place according to a set of rules. However attractive such an approach may be, Baltic (Lithuanian) offers a big variety of frequent, near to neutral word order patterns which in our view makes it *ad hoc* to postulate a specific word order as a point of departure. We have also abstained from discussing the validity in Baltic of Wackernagel's Law which states that clitics appear in second position (i.e. after the first stressed word) in the sentence in IE.

In conclusion, the above presentation has revealed a comprehensive system of clitics in Baltic which can almost cope with those of Slavic and Romance. Many questions, however, remain to be solved. Thus, the study of clitics in Baltic deserves more attention both from Baltologists and from linguists in general.

LITERATURE

- Bybee J. L.*, 1985, *Morphology. A study of the relation between meaning and form*, Amsterdam, Philadelphia.
- Endzelin J.*, 1922, *Lettische Grammatik*, Riga.
- Fraenkel E.*, 1962-65, *Litauisches etymologisches Wörterbuch*, Heidelberg.
- Hopper P. J., Traugott E. C.*, 1993, *Grammaticalization*, Cambridge.
- Kazlauskas J.*, 1968, *Lietuvių kalbos gramatika*, Vilnius.
- Mathiassen T.*, 1996, *A short grammar of Lithuanian*, Columbus.
- Nevis J. A. et al.*, 1994, *Clitics. A comprehensive bibliography 1892-1991*, Amsterdam, Philadelphia.
- Rosinas A.*, 1995, *Baltų kalbų įvardžiai: morfologijos raida*, Vilnius.
- Stang Chr. S.*, 1966, *Vergleichende Grammatik der baltischen Sprachen*, Oslo-Bergen-Tromsø.
- Stolz Th.*, 1989, *Zum Wandel der morphotaktischen Positionsregeln des baltischen Reflexivzeichens*, - *Folia Linguistica Historica*, IX, 13-27.
- Zinkevičius Z.*, 1981, *Lietuvių kalbos istorinė gramatika*, II, Vilnius.