BALTISTICA XXXIII(1) 1998, 5–13

William R. SCHMALSTIEG
The Pennsylvania State University

OLD PRUSSIAN ORTHOGRAPHY: STATISTICS VS. COMMON SENSE

In recent years several scholars (Levin, 1982; Inoue, 1982; 1984; 1992) have attempted to establish more scientific approaches for the study of Old Prussian documentation. Although here I hope to dispute the statistical approach, I must admit that the perspicacity shown by Levin and the sheer labor of the statistical calculations performed in Inoue's work are a tribute to those two scholars' intelligence, industry and perseverance.

Levin (1982, 204) has written that if we know the total number of occurrences of a spelling '...we can calculate the probability that the spellings were distributed randomly. The greater the total number of items, the greater the probability that a difference in the distribution of the spellings is consistent. This of course is elementary methodology for studying any *population*, including a population of items in a text. This is now the methodology that must be applied to the study of the Prussian corpus. The significance of spelling variants is not determinable by *a priori* declarations, impressions, or gut feelings, but by a careful calculation of the actual distributional probabilities.'

Inoue (1982, 3) introduces the notion of 'stability' and 'frequency' of the word-forms encountered in the Enchiridion. He selected from Trautmann (1910) 972 items appearing in the Enchiridion. Of these items 425 had allographically alternating forms in the stem and the ending. The remaining 547 items had no allographic alternations. These latter items were classified according to their number of occurrences, those with the maximum number of occurrences being considered the most 'stable'. Items occurring one, two or three times are marked with the numbers (1), (2) and (3) respectively, whereas items which occur four or more times are marked with the symbol (X). Thus, for example (Inoue, 1982, 6):

kērmens (2) masc. nom. sg.
kermenes (X) gen. sg.
kērmenen (1) – *kermenen* (1) – *kērmnen* (1) – *kermnen* (1) – *kērmenan* (1) acc. sg.
The stem alternation is then:
kermen- (X) – *kērmen-* (X) – *kērmn-* (1) – *kermn-* (1).
The alternation patterns (AP) are as follows:
AP 1: e (X) – ē (X) (i.e., presumably, the alternation encountered in the root)
AP 2: e (X) – Ø (2) (i.e., presumably, the alternation encountered in the suffix)

The alternations encountered in the endings are:
masc. acc. sg. *en* (X) – *an* (1)
The following endings are stable:
masc. nom. sg. *s*
masc. gen. sg. *es* (X)

For each word alternative spellings are grouped together and the different vowel and consonant alternation patterns are extracted. Thus, for example (Inoue, 1982, 11, 16), the noun stem *aucktimmisk-* (X) vs. *auktimmisk-* (2) establishes the consonantal alternating pattern *ck* (X) – *k* (2).

We find then in the Enchiridion an impressive grand total of 426 alternating patterns (Inoue, 1982, 15–54), beginning with (001) *j* (1) – *i* (X) as illustrated by *pickull-* (2), *pikull-* (1), *pjckul-* (1), *pickūl-* (1) and ending with (426) *u* (X) – *ū* (1) illustrated by *teinu* (X), *teinū* (1).

From this grand total of 426 alternating patterns one can find varying degrees of significance of the alternations. Inoue (1982, 55–56) considers as 'significant' only those alternations which have (X) : (X), i.e., four or more attestations of each allograph. Thus, for example, *ī* vs. *ij*, represented in *grīk-* (X), *grijk-* (X), *grik-* (2); *mijl-* (X), *mīl-* (X); *stawijd-* (X), *stawīd-* (X), *stawid-* (2) is a 'significant alternation' and allows Inoue to draw the conclusion (1982, 57) that the allograph *ij* is a special sign to express the long *ī*.

Inoue writes also (1982, 56–57): 'The macron placed over a single graph means either the lengthening of the vowel or the intonation... The macron placed over the diphthong is, presumably, the sign of intonation...' Significant alternations 2. *e* – *ē*, 4. *ei* – *ēi*, 6. *a* – *ā*, 7. *ia* – *iā*, 8. *ai* – *āi* and 9. *ou* – *oū* seem to consist of only the presence or the absence of a macron.

Among the diphthongs Inoue includes the allographs *ia* and *iā* which are illustrated by the word *crixtiānisk-* (X) – *crixtianisk-* (X) – *cristiānisk-* (2) – *krixtiānisk-* (1) – *crxtiānisk-* (1). However, the graphemic sequence *-iā-* is encountered also in *waitiāt* 'to speak' and *biātwei* 'to fear'. I would phonemicize the first word either as /vaitijāt/ or /vait'āt/ (the latter on the basis of Old Church Slavic *věštati* 'to inform') and the second word as / bi[j]ātvei/ (on the basis of the Lithuanian counterpart of *bijóti* 'id'). It would seem likely to me that the Old Prussian graphemic sequence *-iā-* does not necessarily denote a diphthong, but two syllables */i[j]ā/, or monosyllabic */jā/, or */ā/ preceded by a palatalized consonant. I would point out, that Mažiulis (1966, 50) writes that in the sequence **-ij-* the sound **-j-* is not rendered by a separate letter in *kalab-i-an* (EV 424), *crixt-i-a* III, *biātwei* III. Furthermore, Mažiulis (1966, 55) suggests that in the word *cris-ti-āniskan* III the letter *-i-* denotes the softening of the preceding **t*. The first three words were later phonemicized by Mažiulis (1981, 277, 281, 261) as [kalabījan, krikstija, bijātvei] respectively. Although the accusative singular *cris-ti-āniskan* is not mentioned in his vocabulary, Mažiulis phonemicizes the nomina-

tive singular *cristiāniskas* as [krist'āniskas] (1981, 281). Mažiulis, however, apparently relying on his 'gut feeling' and general knowledge of Baltic philology, presents an analysis which in my view seems more likely, although it might be hard to support statistically.

Among the consonants the significant alternations are 11. *b – bb – p*, 12. *d – dd*, 14. *r – rr*, 15. *s – ss*, which leads Inoue to the conclusion (1982, 57): 'The alternation between a single consonant and a double consonant occurs with no apparent restriction, i.e., the circumstances of their appearances are far from definition.'

Inoue's view, however, would be in sharp contrast to views such as that of Trautmann (1910, 196) who thought that a doubled consonant denoted a preceding short stressed syllable or Kortlandt (1974, 300) who has proposed that a doubled consonant denotes that the following vowel was stressed. Another view is that of Smoczyński (1989a, 128–132) who writes that a comparison of the words marked with the macron in the Enchiridion with the corresponding words in the I and II catechisms shows that the macron in Abel Will's text corresponds to certain letters in I and II which, not being segmental phonemes, may have some accentual significance. Such writing he calls 'letter accentography' (*literowa akcentografia*). In the Enchiridion he finds doublets which illustrate the principal of either writing an extra letter or the macron, e.g., /ú:/: *b-o-uton* = *būton* 'to be' (cf. Lith. *bū́ti*); *p-o-utwei* = *pūton* 'to drink'; I (acc.) *mut-t-in* = Ench. *mūtien* 'mother'. My own view is closest to that of Trautmann, since in Middle Low German orthography doubled consonants commonly (but not always!) denoted that the preceding vowel was short (see Lasch, 1914, 56–57, 134–135). On the other hand I am readily willing to admit that there were numerous deviations, cf., e.g., *mut-t-in* listed above and *naseilliwingiskan* 'geistlich, spiritual' in which the second syllable, being a diphthong, probably has two morae.

Concerning the possible neutralization of the phonemic contrast /e/ and /a/ Inoue (1984, 25) writes: '...I can not imagine that the vocalic system of Old Prussian should have shown more radical a development than that of MODERN STANDARD Lithuanian'. But the chronological attestation of a language says little about the stage of phonological development. Ancient Sanskrit had already merged Indo-European **ē* and **ō*, cf. *dhā-* 'to put' as opposed to *dā-* 'to give' whereas contemporary Lithuanian keeps them apart in *dė́ti* 'to put' and *dúoti* 'to give'.

Now in my view the Lithuanian near neutralization of the Baltic */e/ vs. */a/ contrast is in two steps. The first step is a result of the merger of */tje/ and */te/, (cf. Lith. voc. sg. *svetè* from *svẽčias* 'guest' vs. *teñ* 'there'), such that only */tja/ or */t'a/ (Lith. *čià* 'here') becomes possible, i.e., there can be no contrast between /e/ and /a/ after a palatalized consonant, */tje/ having been lost. This first step was probably common to all the Baltic languages, since it is this merger which led to the unification of the thematic vowel in the Baltic verb (see Schmalstieg, 1963). Presumably at this time there could still be a contrast of */e/ vs. */a/ after an unpalatalized consonant,

thus */te/ vs. */ta/ would be possible, as they are indeed in contemporary Latvian, e.g., *te* 'here' vs. *tad* 'then'.

The second step occurs when front vowels begin to palatalize preceding consonants, i.e., when Lithuanian */te/ passes to */t'e/, the contrast between /e/ and /a/ is lost completely in post-consonantal position, but is just barely maintained in word initial position, see Girdenis (1995, 62).

For Old Prussian the first step is generally recognized. Thus, e.g., Klusis (1995, 91) writes 'the opposition /a/:/e/ was neutralized after the palatals in Prussian similarly to modern Lithuanian'. If front vowels palatalized preceding consonants in Old Prussian, then the second step would be realized and the situation would be parallel to that of Lithuanian. Possible evidence for this might be supplied by such spellings as *pyienkts*, *piēncts* 'fifth' beside Lith. *peñktas* 'id'. In any case Klusis (1995, 91) in discussing the form of *ast* 'is' and *asse* '(you) are', suggests the possibility that the *e* : *a* opposition was weak in Old Prussian as in modern Lithuanian. I would agree with him in this respect.

One cannot, of course, exclude the possibility that the palatalization of consonants by a following front vowel was common to all the Baltic languages and that Latvian later lost this feature. In this case the post-consonantal neutralization of /e/ vs. /a/ would have been completed in a single step.

I have proposed in (1964 and 1974, 18–19) that in Old Prussian there may have been two co-existent vocalic systems, a primary system denoted by the number [1], and a secondary system denoted by the number [2]. Thus [1] /ē/ > [2] /ī/ (Lith. *turéti* 'to have' beside Old Prussian *turīt* 'id.'), [1] /ō/ > [2] /ū/ (*nosēilis* 'spirit' beside *nuseilin*), [1] /ī/ > [2] /ei/ (*gīwan* 'life' beside *geīwan*), [1] /ū/ > [2] /ou/ (*būton* 'to be' beside *boūton* and *baūton*). Now the chain shift that I propose is fully in accord with contemporary principles of historical linguistics (Labov, 1994, 123–124). A parallel can be found in the English Great Vowel Shift, where according to Labov's principle I (raising of long vowels) /ǣ/ > /ē/ > /ī > iy/ and /ɔ/ > /ō/ > /ū > uw/ and according to his Principle IIa (lowering of diphthongal nuclei) /iy > ey > ay/ and /uw > ow > aw/.

Inoue (1984, 26), however, relies on Marchand (1970, 111) to reject the notion of co-existent phonemic systems in Old Prussian. But I should like to quote in full what Marchand (1970, 111–112) wrote: 'In his work on the phonemes of the Old Prussian Enchiridion, Schmalstieg has made use of the criterion of coexistent phonemic systems to explain certain features of that text... I am not prepared to say that Schmalstieg is wrong; in fact, I am convinced that he must be to a great extent right, and it is obvious that any language will have such systems and subsystems. It is precisely this generality that makes the use of such a method dangerous; an invulnerable statement (i.e., one which cannot conceivably be contradicted) is scarcely a statement... This remark must apply to all statements which begin with an opener such as „It seems quite likely that..." (to be read: „It may well be true that..." or „The follow-

ing assumption allows me to make a number of other assumptions"). This does not mean that I wish to delete all enabling or operational statements; they must be used with extreme care and labeled for what they are.'

With respect to this Marchand (1970, 116, fn. 10) refers to such logical positivist philosophers as Carl G. Hempel, R. Carnap and H. Reichenbach, so I assume Marchand has in mind the 'falsifiability theory of meaning'. But a meaningful statement is not only one which can be tested, but also one for which one could imagine a test. One could imagine the phonemic analysis of the speech of an Old Prussian peasant.

In addition I would object that logical positivism itself is just one possible theoretical stance and is certainly not universally accepted. For example, Professor Kurt Huebner of the University of Kiel, as quoted by Feyerabend (1978, 145), maintains that 'the source of scientific progress lies neither in abstract rules of falsification, nor in inductive inferences and the like, but in the entire mental and historical situation in which a scientist finds himself... The decisive weakness of contemporary philosophy of science seems to me to lie in this: despite the great variety of schools and thinkers it still proceeds unhistorically. It tries to solve its basic problems – the character of the methods to be applied and the justification of the statements obtained with their help – by mere reflection, where thinking apparently is only left to itself and to its sophistication...' Even a casual reading of Feyerabend's book reveals the weaknesses of the so-called 'scientific method'.

Likewise the notion of co-existent phonemic systems is not a strange and unusual notion, but one which was developed by the eminent American linguists, Charles C. Fries and Kenneth L. Pike (1949, 29) who wrote: 'The speech of monolingual natives of some languages is comprised of more than one phonemic system; the simultaneously existing systems operate partly in harmony and partly in conflict. No rigidly descriptive statement of the facts about such a language accounts for all the pertinent structural data without leading to apparent contradictions.' Fries and Pike write further that there are several types of problems connected with the analysis of coexistent phonemic systems. These include '...a conflict in the system of sounds of a single speaker during a transition stage wherein a phonemic contrast is being introduced or lost or replaced by linguistic change of some type over a period of time within a single dialect'. The notion of coexistent phonemic systems can also prevent (1949, 30) '...internally inconsistent and self-contradictory analyses which result if one treats on a single descriptive level (a) those particular differences of sound which occur in one uniform style, and (b) those which are due to a qualitative or stylistic change, in whispering, song, extra-fast utterance, extra-precise pronunciation, or the like, in the speech of one individual'. Labov (1995, 158) notes that the data of spontaneous speech provides '...wide dispersion of vowel tokens, ranging along the path of the change from more advanced to less advanced tokens'. The more advanced tokens are used in '...intense social exchanges with peers'.

Inoue (1992, 16–17) writes that one can easily imagine that frequently hearing and then frequently copying the same word form, a person could make mistakes, i.e., that sometimes he will hear a form incorrectly and then copy it down exactly as he heard it. One might suppose that the greater the frequency the correspondingly fewer mistakes, i.e., thanks to the frequent repetition of the correct form, the hearer naturally corrects his possible mistake, as a result of which the number of incorrectly copied forms decreases. In reality, however, according to Inoue, the matter isn't so simple, i.e., within a single short text there are very few words or word forms that are repeated sufficiently often, several dozen times or even more than 100 times. Thus, for example, in the Enchiridion the word *bhe* 'and' appears about 400 times, whereas its orthographically incorrect variants, *bbe* and *bke* each appear three times and have no importance for the establishment of the conjunction 'and'. Practically, however, according to Inoue, the great majority of words or word forms occur more or less than five times with their orthographic variant(s). From this arises the difficulty in the establishment of the correct form. One example is the infinitive 'to be': *būton* (1x), *baūton* (2x), *bouton* (2x), *boūt* (6x), *bout* (2x). For this word the root has the following alternation: *bū* (1x), *baū* (2x), *bou* (4x), *boū* (6x) and the infinitive suffix as *-ton* (5x), *-t* (8x).

Indeed it is just the notion that a more frequently repeated form would somehow be more likely to be correct that I dispute. It is my experience from dealing with many non-native speakers of English that an 'incorrect' form can be repeated over and over again and thereby become firmly established in the speech and/or writing of the non-native. For example, a certain native speaker of Russian in one of my classes consistently spelled the English word *quite* as *quiet* in spite of my numerous corrections. Also typically the hearing of the older non-native does not improve appreciably over a period of time and once he has acquired certain phrases and accent, these do not appreciably change, frequently because the native speaker does not feel comfortable about correcting the non-native. Although I felt quite comfortable in trying (unsuccessfully!) to correct my non-native student, one might wonder how easy it was for an illiterate Prussian peasant to correct a Protestant clergyman.

Again relying on the evidence of Lithuanian I would suggest that in the word *bhe* the orthographic *-h-* did not reflect anything in the pronunciation, so that repeated hearing of the word had no bearing on whether the *-h-* was written or not. The Middle Low German habit of writing *-h-* after consonants frequently had no phonological significance (see Lasch, 1914, 136–137). The *-h-* was commonly used in short words such as *ghân*, *shê*, *vhê*, *nhâ* in order to give them more weight.

It is well known that typically in the Germanic languages the stop consonants have delayed voicing (see Martinet, 1986, 87). One could therefore even argue that for the Old Prussians the writing *bbe* would have better expressed the pronunciation, because it would have emphasized the voicing of the initial consonant, whereas

a phonological sequence of /b/ plus /h/ would seem to have been uncanonical in both German and Old Prussian. If one were to believe that the *-h-* had phonological significance, one would have to propose, I assume, a phonological system for Old Prussian with aspirated voiced stops, like that of Sanskrit, a scenario which would seem highly unlikely to me, and I believe, to most Balticists, even though it would seem to have statistical support.

In regard to statistical approaches I should like to reiterate here my thought (1990, 360–361) that although, of course, no two situations are ever exactly the same, it appears to me likely that the Old Prussian situation parallels the Old Latvian situation since many of the Old Latvian texts were written down by German pastors. Thus Ozols (1965, 76) writes that the letter *e* is used for all of the Latvian endings with a short vowel. Following the advice proffered by Levin (1982, 204) one would then conclude that the ending *-e* is always the correct ending for all Latvian final short vowels. Now there are always auxiliary hypotheses which can justify any theory. One could suppose, for example, that in early Latvian all of the short vowel endings had indeed merged and that only later did they appear in different form to somehow match their Lithuanian phonological counterparts. Or maybe these particular texts reflected some dialects which disappeared and that the dialects which survived maintained the distinctions in question. My 'gut feeling' and common sense tells me that some German pastor merely failed to distinguish Latvian short final vowels.

But let us look briefly at the first Latvian grammar, Johann Georg Rehehusen's *Manuductio ad Linguam Lettonicam* (1644). With regard to the orthography Fennell (1982, 113–114) writes: 'In fact, Rehehusen really has no system at all: cf. his treatment of open and closed *e*. From his remarks, it would appear that these are to be distinguished, and represented by *ä* and *ö* respectively, but in practice, the distinction is often neglected: *es* is most often written *öß* (pp. 14, 15, 16, etc.), but the form *äß* also occurs (p. 22); the contrast between *esmu* (with *ä*) and *esi* (with *ö*) is nowhere observed – the only forms given are *ößmu* and *ößi* (pp. 26, 35, etc.). Similarly the expected *äßit* (*esat*) occurs only once (p. 20) against numerous examples of *ößit* (pp. 15, 18, 22, etc.).'

Notice the correct spelling *öß* 'I' (in the sense of reflecting the pronunciation, not in the sense of a consistent orthography) as opposed to the incorrect *äß*. On the other hand the expected **äßmu* (modern Latvian *ęsmu*) doesn't occur at all. A statistical preponderance of *ößit* 'you are' would certainly outweigh the correct (by contemporary standards) *äßit* (modern Latvian *ęsat*) which occurs only once. Again all kinds of auxiliary assumptions would be possible, e.g., that Rehehusen and interlocutors did indeed use the form *esat* instead of contemporary *ęsat*, etc.

It seems to me that the Latvian linguist R. Grabis (as quoted and translated by Fennell, 1982, 113) is correct in expressing his 'gut feeling' in the following way:

'Word endings are for the most part incorrect and attached at random, since the author was unable to distinguish whether the vowel of the ending was *-a*, *-e*, *-u* or *-o*...'

I would contrast the statistical approach to that of Smoczyński (1989b) who, for example, reminds us of the fact that the macron over a vowel in an Old Prussian word does not necessarily mean length, but may be a 'suspension', i.e., it may denote the vowel plus a following nasal, usually *-n*, but occasionally *-m*. Smoczyński gives such examples from Duke Albrecht's and Abel Will's forewords to the Enchiridion as *gnedigē* for *gnedigen* 'gracious', *getragē* for *getragen* 'carried', etc. This can be supported by such orthographic variation as the following: *potaukīsnan* 'Verheissung, promise' beside the apparent gen. sg. *potaukin-snas* and the past act. part. *potaukinnons* 'verheisst, promised'; *powartīsnan* 'Busse, penance' (but with an original meaning of 'Bekehrung, reversal, conversion, turning') beside the verb *wartint* 'kehren, to turn'. This allows Smoczyński (1989b, 180) to suggest for *sīdans* 'sitzend, sitting' a reading *sindans* and compare it with *sindats* I and *syndens* II. In my view Smoczyński without the aid of any statistical theory has made an important contribution to our understanding of Old Prussian. Since within the stem Inoue separates vowel alternating patterns from consonant alternating patterns the alternating pattern *-in-* vs. *-ī-* doesn't occur and the words *potaukīsnan* and *potaukinsnas* (perhaps as the result of a minor misprint) are included in the alternating pattern 005, viz., *-i-* vs. *-ī-* (see Inoue, 1982, 13–14). In this manner the statistical approach fails to capture the possibility of an interesting generalization.

Therefore I conclude that Old Prussian texts offer too limited a corpus for meaningful statistical interpretation. In the final analysis, the interpretation of any phenomena depends more on the taste of the researcher, than on observed facts. Those who are impressed by more or less automatic 'discovery procedures' supported by logical positivist philosophers of science will probably never accept the findings of those who rely on their general knowledge of Baltic philology and 'gut feelings' based on a comparison with parallel situations.

REFERENCES

Fennell T. G., 1982, The first Latvian grammar: J. G. Rehehusen's 'Manuductio ad linguam lettonicam...', Melbourne.

Feyerabend P., 1978, Against method, London.

Fries Ch. and K. L. Pike, 1949, Coexistent phonemic systems, – Language, XXV, 29–50.

Girdenis A., 1995, Teoriniai fonologijos pagrindai, Vilnius.

Inoue T., 1982, A graphology of Old Prussian Enchiridion = Annals of Foreign Studies XIII, Kobe City University of Foreign Studies.

Inoue T., 1984, A graphonological and phonological study of Old Prussian Enchiridion with the critical text = Annals of Foreign Studies XV, Kobe City University of Foreign Studies.

Inoue T., 1992, Normalizacija prusskogo jazyka Enxiridiona, Kobe, Monograph XXII, Institut inostrannyx jazykov g. Kobe.

Klusis M. and B. Stundžia (eds.), 1995, Pirmoji prūsų knyga, Vilnius.

Kortlandt F. H. H., 1974, Old Prussian accentuation, – Zeitschrift für vergleichende Sprachforschung, LXXXVIII, 299–306.

Labov W., 1994, Principles of linguistic change, Oxford and Cambridge (USA).

Lasch A., 1914, Mittelniederdeutsche Grammatik, Tübingen.

Levin J., 1982, Graphology and sound change in Old Prussian, – Anders Ahlqvist (ed.), Papers from the 5th international conference on historical linguistics (Current issues in linguistic theory, XXI), Amsterdam and Philadelphia, 201–210.

Marchand J., 1970, Some remarks on the German side of the Elbing vocabulary, – Magner T. and W. R. Schmalstieg (eds.), Baltic linguistics, University Park and London, 109–117.

Martinet A., 1986, Des steppes aux océans, Paris.

Mažiulis V., 1966, Prūsų kalbos paminklai, Vilnius.

Mažiulis V., 1981, Prūsų kalbos paminklai, II, Vilnius.

Ozols A., 1965, Veclatviešu rakstu valoda, Rīga.

Schmalstieg W. R., 1963, The Baltic thematic vowel reconsidered, – Lingua, XII, 217–219.

Schmalstieg W. R., 1974, An Old Prussian grammar: The phonology and morphology of the three catechisms, University Park and London.

Schmalstieg W. R., 1990, A few issues of contemporary Indo-European linguistics, – Ph. Baldi (ed.), Linguistic Change and Reconstruction Methodology, Berlin, New York, 359–374.

Smoczyński W., 1989a, Studia bałto-słowiańskie, I, Kraków.

Smoczyński W., 1989b, Zur Schreibung im Apr. Enchiridion (I), – Baltistica, III 1 Priedas, 179–195.

Trautmann R., 1910, Die altpreussischen Sprachdenkmäler, Göttingen.