

The Development of PIE *ē in Palaic*

Rex E. Wallace

1. The Anatolian languages show considerable diversity in their treatment of Proto-Anatolian *ē < PIE *ē. In Hittite it is true for the most part that the mid-palatal vowels remain unaltered.¹ In Luwian and Hieroglyphic Luwian *ē has generally lost its palatal color and merged with ǣ² while its long counterpart *ē has become ī under accent and ı in unaccented position. A depalatalization process similar to that attested in Luwian has been proposed for *ē in Palaic on the basis of such words as a-aš and a-aš-du < PIE *h₁es and *h₁estu, the second and third singular imperatives of the verb 'be' (see e.g. Kammenhuber 1959:30; Puhvel 1966:239; Friedrich 1960:183; Carruba 1970:39; and Oettinger 1979:558³). Similarly, the treatment of *ē in Palaic is said to parallel that of Luwian, e.g. *ē > ı̄/ı̄ in ú-i-te/ti-si < uēdhēsi 'you build' (see Oettinger 1979:130). There are, however, a number of reasons to question the claim that Palaic and Luwian have treated the Proto-Anatolian mid-palatal vowels in an identical fashion.

The evidence for a sound change *ē > ǣ in Palaic is quite tenuous. In fact there appear to be only two legitimate examples of the change: a-aš and a-aš-du. Other presumed examples of *ē > ǣ are of questionable value and in most cases, if not all, more appropriate derivations, which involve no such hypothesized change, can be offered (see below pages 3ff and 6). Further, the plene writing of these forms (a-aš not **aš, a-aš-du not **aš-du) indicates that the root vocalism here is phonetically long and not short as would be expected were the vocalism the direct reflex of *ē via sound change.⁴ Thus, if a sound change *ē > ǣ is to be proposed for Palaic, perhaps on the basis of a-aš and a-aš-du alone, then the lengthening process, by whatever means, must be satisfactorily explained in relation to the change *ē > ǣ. These facts, in addition to the fact that possible counterexamples to a change *ē > ǣ exist (see below page 2), prompted Calvert Watkins in his article 'A Palaic Carmen' (Watkins 1978:309) to suggest that *ē may indeed have remained in Palaic.

The sound change *ē > ı̄/ı̄ is based upon the etymologies of two forms: Palaic ni-i 'not' < PIE *nē (Oettinger 1979:535) and ú-i-te/ti-si < *uēdhēsi 'you build' (Oettinger 1979:130). Both words are of questionable value as evidence for *ē > ı̄/ı̄ because the etymologies which have been suggested can be disputed. For the form ni-i Heiner Eichner (MSS 29:40 footnote 33) has proposed a derivation from PIE *nei (compare Latin (archaic) nei, (classical) nī 'not', Oscan nei 'not', Lithuanian niē-kas 'no one', and Avestan naēčis (Pokorny IEW 757)). This etymology has the advantage of utilizing a sound change which must be independently reconstructed for the language on the basis of a form with an unquestioned

etymology: ki-i-ta-ar 'lies' < *keitore (see N. Oettinger MSS 34:113 and 1979:536; H. Eichner MSS 31:78 and 80). Oettinger (1979:130 and footnote 84) has argued that the pre-form of Palaic ú-i-te/ti-ši is likely to be *uēdhēsi, with a long vowel in the first syllable, based on the evidence of the Lydian form vici- < *uēdhē- 'build'. But the derivation of Palaic ú-i-te/ti-ši from a pre-form *uēdhēsi involves at least two problems. First, one must assume that the [ĩ] vowel in the second syllable (after the proposed sound change *ē > ĩ in unaccented position) can be represented by a Ce sign (note that this word is written ú-i-te-ši (2x), ú-i-ti-ši (1x)). But there is no reason to expect Ce to represent [ĩ] phonetically since the appropriate Ci sign would exist in every case (e-valued signs are lacking in a considerable number of cases, see H. Eichner 1980:133 and E. Sturtevant 1933:43-46 and 50-52). Thus, while fluctuations in representation (e-i) may indicate that the vowel in this syllable is indeed short (so it is argued by Oettinger 1979:130), it is more likely that such fluctuations are indicative of [ě] rather than [ĩ] (see Oettinger 1979:130 and 533 ff.). As a result the second syllable of ú-i-te/ti-ši is a problem for the claim that the development of *ē is parallel in Luwian and Palaic. Second, the Hittite cognate (compare, for example, ú-e-te-iz-zi 'builds'; see also Oettinger 1979:129-130) appears to agree with Palaic forms with respect to quantity of the initial syllable, short not long (and the quality of the second, e not i). Further, Oettinger's claim (1979:130) that the Hittite forms reflect a vowel shortening process in the first syllable 'im Nebenton' is completely ad hoc. No evidence is adduced for the shift of (primary?) stress from the pre-verb to the verbal root in verbs of the pehutē-Class (Oettinger's Class I 2 g; see Oettinger 1979:36 ff. and 125 ff.)⁵ In fact it is suggested (Oettinger 1979:36 and 107 footnote 43) that stress on the pre-verb shortened long vowels in subsequent syllables.⁶ Moreover, even if we grant such a shortening process in the first syllable of this Hittite verb, it is still possible to claim that the vowel quantity in the first syllable of the Palaic word is short and hence that Palaic too must have had such a process. The writing ú-i- may indicate not only a long [i] but also a short [ě] or [ĩ] (see Oettinger 1979:533 and 1979a:201). In light of all of these problems with the etymology of ú-i-te/ti-ši there is no overriding necessity to assume that Oettinger is correct in his claim that the pre-form in Palaic contained a long vowel in the first syllable. The possibility exists that the Hittite and Palaic forms, though perhaps ultimately from Proto-Anatolian *uēdhēsi, are to be derived from an intermediate pre-form with a short vowel in the initial syllable: *uēdhēsi. As a result the claim that Palaic has evidence for a sound change *ē > ĩ/ĩ is questionable.

Over and above the shortcomings of these proposed sound changes, there is a more serious problem. The major developments of the mid-palatal vowels offered thus far fail to account satisfactorily for all of the morphemes which have e-vocalism and are of PIE origin: e.g. ú-e-ir-ti 'says' < *uer (IEW 1162); -Ci/e-eš 'nominative plural' < *-ēs consonant stem ending (Kammenhuber 1959:33 footnotes 1 and 2) or *-ēs < *-eies i-stem ending⁷; (-)eš-hu-ur/(-)e-eš-ha/(-)e-eš-ha-na 'blood' < *h₁esh₂r (IEW 343 and Tischler 1977:112-115); te-e-ka-an-za

te-ta-a-an-za 'flowing' < *tek (IEW 1059 and Watkins 1978: 311⁸). It can be shown that neither phonological development proposed for the palatal vowels is compatible with the other when these morphemes are considered.

The sound changes * \check{e} > \check{a} and * \bar{e} > \bar{i}/\bar{y} were reconstructed on the basis of the following items: a-aš/a-aš-du 'be' < *h₁es- (IEW 340); -ua-ni 'first person plural' < *ueni; -uar- 'particle of direct speech' < *uer- (IEW 1162); ua-at-ta-na 'in water (?)' < Pre-Palaic *ueteno (?); ma-li-ta-an-na-as 'honied' < *melit- (IEW 723) for * \check{e} > \check{a} ; and ni-i 'not' < *nē (IEW 756) and ú-i-te/ti-si 'you build' for * \bar{e} > \bar{i}/\bar{y} . If we agree that * \check{e} > \check{a} then it is impossible to fit the remaining lexical items with e-vocalism into the system of phonological developments. Neither ú-e-ir-ti nor -Ci/e-eš can be allowed to have short e-vocalism because these items would then form counterexamples to those very words used to argue for * \check{e} > \check{a} , specifically -ua-ni, -uar-, ua-at-ta-na, and a-aš/a-aš-du (note that the phonetic environments for the opposing sets of words are identical: after u and before s). But if it is assumed that the vocalism in question in these words is long, a suggestion which is perfectly plausible, then it is impossible to maintain * \bar{e} > \bar{i}/\bar{y} in Palaic. Since the environment in which * \bar{e} is found is identical in both words, after u, it is impossible to plead for a special phonological development. Rather one is compelled to concede that * \bar{e} > \bar{i}/\bar{y} is an impossible development, especially in light of the remaining words (-)eš-hu-ur/(-)e-eš-ha/(-)e-eš-ha-na and te-e-ka-zn-za/te-ta-a-an-za which could be offered as corroborative evidence for * \bar{e} > \bar{e} . The same incompatibility can easily be shown if one uses the suggested sound change * \bar{e} > \bar{i}/\bar{y} as a starting point. In this case the evidence speaks for itself: * \check{e} > \check{a} and * \bar{e} > \bar{i}/\bar{y} are mutually exclusive.

Thus, whatever position is taken with respect to the changes * \check{e} > \check{a} and * \bar{e} > \bar{i}/\bar{y} it is clear that they are mutually exclusive and that the evidence for the development of the mid-palatal vowels in Palaic needs to be reexamined and a hypothesis offered in which the phonological developments are accounted for in a way which is consistent with the data, which allows for natural phonological developments, and in which the end product is a reasonably well organized synchronic system. This paper is an attempt to suggest such a hypothesis.

2. Any discussion of the treatment of PIE mid-palatal vowels in Palaic is troubled from the outset. Essentially this is the result of a combination of facts. First, the existing Palaic texts were written by scribes who were native Hittite speakers. As a result the texts contain errors in representation of words as well as failures to make word divisions. Second, there are only a small number of morphemes inherited from PIE which can be used as evidence for phonological developments. Moreover, the writing system does not possess the means to distinguish e and i in all phonetic contexts. And finally, the indication of the quantity of medial vowels is in some cases ambiguous (see Oettinger 1979: 533 ff.). As a result considerable diversity in the treatment of possible phonological developments is conceivable.

The data which are relevant to a discussion of the development of the mid-palatal vowels in Palaic can be most conveniently divided into categories on the basis of a rather disparate and sometimes overlapping set of criteria.

A. The vocalism of the item in question is ambiguous due to the fact that there is no Ce sign available to represent e-vocalism. Hence Ci may be used to represent [Ce] or [Ci] phonetically. To this category belong: az-zi-ki-i 'gobble up' < *ats-e/i-kī < *h₁d-ské⁹; gi-nu-kat 'flesh or part of the body' < *genu-g(h)od¹⁰; ki-iš-ta-a-am-mu 'expired, dull, faded' + the dative of the first person singular enclitic pronoun < *gesdhont-mu.¹¹

B. The vocalism of the item in question is ambiguous due to the manner in which it is represented by the writing system (for the representation of e in the Hittite writing system see Oettinger 1979: 533 ff.). Included in this category are: 1) ē-ě u-e-ir-ti 'says'; -Ci/e-es 'nominative plural'. 2) ī-ī-ě in the first syllable of ū-i-te/ti-ši 'you build'. 3) ē-ě-ī in the second syllable of ū-i-te/ti-ši. 4) ě-ī ša-a-ú-i-ti-ra-an(-) 'horn' < *sauetran < *souh₁-e-tro-m (Oettinger 1979a: 201-202)¹².

C. The vocalism of the item is questionable because alternant pre-forms can be suggested. ma-li-ta-an-na-aš 'honied', -ua-ni 'first person plural', ua-a-šu 'well-', -uar- 'particle of direct speech', belong in this category¹³.

ma-li-ta-an-na-aš. If this form is to be derived from the stem *melit- via the addition of the possessive suffix *h₃on- (see Eichner 1980: 147 footnote 69) then a basic form with zero grade vocalism of the root syllable, *mlithon-, is to be preferred to full grade. The orthography supports such a hypothesis because a (ma-) is the graphic representative of an empty vowel.

ua-a-šu. The root vocalism of ua-a-šu is difficult to ascertain because the original ablaut alternations of u-suffixed (o in strong cases: e in weak cases) substantives was generally leveled in favor of one of the alternants (compare Latin genu : Hittite gēnu but Greek γόνυ : Sanskrit jānu 'knee' (IEW 380). However, since o-vocalism is original in the nominative-accusative form, and since we have no reason to believe this situation was otherwise for this word in Palaic, it is reasonable to maintain that the root vocalism in this form was originally *o, *h₁uosu.¹⁴ The length of the vowel in this form is undoubtedly the result of the lengthening process described in Oettinger 1979: 447 ff. and Eichner 1980: 144 footnote 67.

-ua-ni. Oettinger (1979: 566 footnote 12) claims that Palaic and Luwian -ua-ni 'first person plural' is to be compared with Hittite -ue-ni and as a result reflects the sound change *ě > ǎ.¹⁵ It should be pointed out however that a variant inflectional ending with a-vocalism occurs in Hittite, particularly in the older period (see Oettinger 1979: 9). As a result one could argue that two variant endings existed in Proto-Anatolian, -ue-ni and -ua-ni, and that Palaic has simply generalized the use of the -ua-ni variant at the expense of -ue-ni. It is thus difficult to use -ua-ni as evidence for a sound change *ě > ǎ.

-uar-. The standard etymology of -ua(r)- (Palaic has simply -uar- but Hittite shows two phonologically conditioned variants -ua- and -uar-) connects it with the root *uer- (IEW 1162; see, for example, Eichner 1975: 84; Oettinger 1979a: 201), compare Hittite ueriya- Palaic ú-e-ir-ti, Latin uerbum, and Greek ερέω. The disadvantage of such an etymology, the fact that it fails to account for the -ua-/-uar- alternations in Hittite, has been pointed out by B. Joseph (1982) in his article 'Hittite iwar, wa(r), and Sanskrit iva'. Joseph, developing an etymology first suggested by J. Przyluski in 1934, persuasively argues that -uar- is composed of a particle *ue/o- with an adverbial suffix -r.¹⁶ If Joseph's etymology is correct then the Palaic form -uar- is unacceptable as evidence for *ě > ǎ because the Hittite forms -ua-/-uar- show that the vocalism was originally *o not *e.

D. The lexical item is of questionable value as evidence because alternate writings make it difficult to decide on a pre-form. lu-ki-it (lx)/lu-u-ki-it (lx)/lu-ki-i-it (7x) 'break into pieces' or 'ignite' belong to this category.

Eichner (MSS 31: 81), following the suggestion of A. Kammenhuber BSL 1959: 29, has proposed that the Palaic forms cited above be derived from *léuket 'ignite'. However the fact that plene writing occurs only one time in the first syllable makes such a proposal unlikely. Oettinger (1979: 276 footnote 35) is probably correct in assuming that this word is a secondary formation in *-ié- and that this verb provides evidence for a special development of *ě, *ě > ĭ / ĭ̃. Carruba (1970: 62) has suggested that the meaning of this word is actually 'break into pieces', a perfectly acceptable meaning in the context of a bread ritual. The PIE root *leuǵ- 'break up' can be offered as an etymology (compare Sanskrit rujāti 'break into pieces', Lithuanian laužiu/laužti 'break up', and possibly Latin lūgeō 'mourn'¹⁷, IEW 686). The Palaic verb is then likely to be a deradical -ié- formation similar to Hittite forms discussed by Oettinger (1979: 343 ff.). Since both full and zero grade forms of the root are attested in Hittite, it is impossible to decide whether the root in this case had full grade vocalism with a subsequent sound change *eu > ũ in unaccented position or whether the vocalism of the root was zero grade to begin with, *leuǵ-ié-t or *luǵ-ié-t > Palaic lu-(u)ki-(i-)it.

E. The lexical item is of questionable value due to a scribal error or due to an error of identification on our part.¹⁸ te-e-ka-an-za/te-ta-a-an-za belong to this category because of the ka-/ta- alternation; (-)eš-hu-ur/(-)e-eš-ha/(-)e-eš-ha-na belong here because in every case the form is written together with hapari- (ha-pa-ri-i-ši-e-eš-ha-na/ha-pa-ri-ua-ni-e-es-ha) or hinapi- (hi-na-pi-eš-hu-ur).

F. The lexical item is of questionable value as evidence because the forms attested in the texts present special problems in phonological development. The various forms of the word 'blood', ua-at-ta-na 'in water (?)', ša-a-ú-i-ti-ra-an(-) 'horn', and ša-pa-ú-i-na 'purifies', belong to this category.

The various forms of 'blood' are of questionable value not only because of the 'Zusammenschreibung' with hinapi- and hapari- but also because it is uncertain how the Palaic forms match up with corresponding

forms in other Anatolian languages. Carruba (1970: 53) suggests that (-)eš-hu-ur and (-)e-eš-ha are both nominative-accusative singular forms. (-)e-eš-ha matches up quite well with Hittite ešhar in terms of the initial vowel and with Luwian asha(r) in terms of the loss of final -r but exactly how (-)eš-hu-ur fits into this scheme is not clear. If (-)es-hu-ur is indeed a nominative-accusative form then the -ur must be explained as a special auslaut development since (-)e-eš-ha and other forms such as karsandu 'cut' < *krsentu (IEW 938) indicate that the regular development of *r̄ is -ar (with -ar > a in the context ____ #C). The texts in which the forms (-)eš-hu-ur and (-)e-eš-ha occur show that these forms do occur in different phonetic environments: (-)es-hu-ur an-na-as 5 A KUB XXXV 163 21', (-)e-eš-ha ti-ua-ni 5 A KUB XXXV 163 13. Hence we might tentatively suggest that *r̄ > -ar (with subsequent loss of r) in the context ____ #C and r̄ > -ur in the context ____ #V, though phonetic motivation for such a development is difficult to discern.¹⁹

ua-at-ta-na. Carruba (1970:79) has proposed that this Palaic word is to be compared with Luwian uattanei from an unattested *uatar. Oettinger, following this suggestion, has proposed that these forms are to be compared with Hittite [uedeni] 'in water' (Oettinger 1979: 535). Such a relationship must be considered speculative because the Palaic and Luwian forms have fortis consonants medially while Hittite attests a lenis stop.²⁰

ša-a-ú-i-ti-ra-an(-). This form is not problematic because the Hittite scribe has failed to make what in our opinion must be a word division following the sign -an but rather because of the -i- vocalism after the t. Oettinger (1979a: 202) notes this problem but leaves it unresolved.

ša-pa-ú-i-na-i. Oettinger (1979: 535) relates this Palaic word to the Hittite form šippāi- 'scrape off' (Friedrich 1952: 193). There are a number of problems with such a correspondence. First, the medial consonants do not match up. Hittite has a fortis stop, Palaic a lenis. Second, no attempt is made to explain the additional suffixal material -ú-i-na- (-i is the -hi conjugation third singular present ending). Finally, it should be noted that Hittite does possess a verb šap- 'scrape off' (see Friedrich 1952: 183) which is probably to be connected with šipāi- in some way. Thus it is conceivable that Palaic ša-pa-ú-i-na-i corresponds to šap- rather than to šipāi-.²¹

G. The lexical items are problematic because the vowels in question are unexpectedly written plene. In this category belong: a-aš/a-aš-du and ka-a-ar-ti 'in the heart'. But the a-vocalism of ka-a-ar-ti is not at issue here since it is the result of a Proto-Anatolian change whereby *ē > ā in the environment ____ R (esonant) C (see (Oettinger 1979: 534).

H. The lexical item is of questionable value as evidence because of possible Hittite origin. The forms az-zi-ki-i and e-eš-ta 'was' belong to this category.

az-zi-ki-i. Watkins (1969: 73) assumes that this form is a genuine Palaic word. Carruba (1970: 52) notes that this form might be a Hittite loan, or at least a Hittitized form. There are two reasons why Carruba might be right. First, if the form is Palaic then the epenthetic vowel

-i- (-zi-) must be explained. Oettinger (1979: 318) has argued for a vowel epenthesis process in Hittite in which a vowel with the quality [e] is inserted into -Csk- clusters (*-Csk- > -Cs-e-k-). If we argue that Palaic shared such a process with Hittite then we are forced to admit another example of *ě which does not become ǎ. Second, only Palaic outside of Hittite provides evidence for an iterative-durative formation in *-ske/o-, and this only in the verb az-zi-ki-i. The rest of the Anatolian languages, Palaic included, show evidence for an iterative-durative in -ša (see Watkins 1969: 73). As a result, it is at least possible that the Palaic form az-zi-ki-i is actually a Hittite loan.²²

e-eš-ta. The questionable status of this word is undoubtedly due to the fact that Palaic is assumed not to have continued PIE e-vocalism. A final decision with respect to the status of the form, Palaic or Hittite, cannot legitimately be made until the development of PIE *ē has been determined. If warranted, it could be argued that this form is a legitimate Palaic word.

3. The relevant lexical items left as residue are a-ni-it-ti 'performs, accomplishes' < *h₃n-ié-ti (Oettinger 1979: 535 and 559) and par/pa-ar-ku-i-ti < *bhr̥ghu-ié-ti 'cleanses' (see Oettinger 1979: 330 ff.). These verbs are secondary derivatives in *-ié- and clearly show a special development of *ě > ǎ after the palatal glide ǰ (see Oettinger 1979: 535).

4. The preceding division of Palaic lexical items makes it very clear that a considerable portion of the discussion of the development of *ē in Palaic is necessarily speculative. While it is possible to develop a consistent picture of the development of these vowels, the particular picture developed will depend on one's evaluation of the forms in the various categories. If it is argued, for whatever reason, that a-aš and a-aš-du are excellent examples of *ě > ǎ (with subsequent lengthening) then possible interpretations of other categories will be delimited in some ways. For example, Class A forms will be seen as special developments of *ě after velars; Class B 1 forms will be derived from proto-forms containing *ē. If one claims on the other hand that Class A and B forms are best treated as continuing PIE *ě, then Class C items will be given non-a derivations, and the forms in Class G (a-aš and a-aš-du are the forms at issue here) will be considered to have a-vocalism but of a secondary and non-phonological origin. The main difficulty then is to find criteria which will allow one to make a decision as to whether the starting point should be *ě > ǎ or *ě > ě. At this point we doubt whether there is any truly principled way to decide.

5. N. Oettinger, in his 'Exkurse zur Lautlehre und den anat. Schwestersprachen' (Stammbildung 1979: 530 ff.), has argued "dass Palaische und das Urluwische auf eine gemeinsame, das Hethitische nicht mehr einschliessende Vorstufe zuruckgehen." If we accept Oettinger's proposal let us say on the basis of morphological innovations, then we would be inclined to adopt a solution which enables us to assume the greatest number of common phonological innovations.

In light of such reasoning it is undoubtedly best to select the change *ě > ǎ as the basic phonological development with the forms a-aš/a-aš-du as representatives of such a change. The forms in Class C and Class F (ua-at-ta-na and ša-pa-ú-i-na-i are the items at issue here)

may be seen as evidence in support of such a change, but only at the expense of excluding ú-i-te/ti-si and sa-a-ú-i-ti-ra-an(-) from consideration. Note that since the phonetic environments in both sets of words are the same, after u, it is necessary to dismiss one set from the discussion. Though there are problems with the forms ú-i-te/ti-si and ša-a-ú-i-ti-r-an(-) these two items are as a whole less problematic than the relevant forms in Classes C and F. In fact the syllable in ša-a-ú-i-ti-ra-an(-) which is relevant to our discussion is not terribly problematic at all. As a result it is probably best to argue that the Class C and F forms (save ša-pa-ú-i-na-i) should be dismissed from the analysis altogether and that ú-i-te/ti-si, if it is to be derived from *uēdhēsi, and ša-a-ú-i-ti-ra-an(-) provide evidence for a special development of *ē to ī after u.

Class A forms gi-nu-kat and ki-iš-ta-a-am-mu in conjunction with the Class D form lu-(u-)ki-(i-)it and the residue lexical items a-ni-it-ti and par-/pa-ar-ku-i-ti may be seen as evidence for the special development of *ē to ī after the palatal consonant i and after velar consonants (which were presumably palatalized, hence the development *Ce > *C'e > C'i).

In sum the contexts for which we must assume special developments of *ē must be extended from 'after i and consonants with the features [+ back and + stop]' (proposed by Oettinger 1979: 535) to 'after i and consonants with the features [+ back]'.²³ In passing we note that Luwian seems to provide evidence for such a development also, compare Luwian ú-i-it-pa-ni- 'old' with Hittite [uetspant-].²⁴

The Class A form az-zi-ki-i, due to its medial vocalism -zi- [tse] (for which see Oettinger 1979: 318), is best considered as a Hittite or 'Hittitized' form.

The phonological developments of PIE *ē (for PIE *eh₁e > Proto-Anatolian *ē and PIE *e₂e > Proto-Anatolian *ē see section 6) in Palaic are:

1. *ē > ī / i _____ : a-ni-it-ti 'performs' < *h₃n-je-ti; par-/pa-ar-ku-i-ti 'cleanses' < *bhr̥ghu-je-ti; lu-(u-)ki-(i-)it 'break into pieces' < *leuǵ-je-t or *luǵ-je-t.
2. *ē > ī / [+ back] _____ : gi-nu-kat 'flesh or part of the body' < *ǵenu-h(h)od; ki-iš-ta-a-am-mu 'expired' + dative of the first person singular enclitic pronoun < *gesdhont-mu; ša-a-ú-i-ti-ra-an- 'horn' < *souh₁etrom; ú-i-te/ti-si 'build' < *uedhēsi.
3. *ē > ā in all other contexts²⁵: a-aš/a-aš-du 'be' < *h₁es/*h₁estu; and possibly ša-pa-ú-i-na-i 'purifies', if this form corresponds to Hittite sipāi- [sepae-] 'scrape off'.

These developments are the same (with the exception of 2) as those suggested by Oettinger (1979: 535) except that he bases development 3 on lexical items from Classes C and F (ma-li-ta-an-na-aš, -ua-ni, ua-at-ta-na, ša-pa-ú-i-na-i) and not specifically (see footnote 3) on a-aš and a-aš-du which are problematic due to the length of the initial vowel.

Two avenues of explanation are open for explanation of length in these forms. Palaic offers some evidence that accented (PIE accent) syllables had their vowels lengthened: the plene writing of such forms as

a-aš/a-aš-du, a-hy-ua-a-an-ti (1x) 'drink', a-ta-a-an-ti (2x)/a-da-a-an-ti (1x) 'eat', mu-ú-ši (if from mūsi, see Oettinger 1979: 560; contra Oettinger, see Eichner 1975: 86 footnote 6) : mu-ša-a-an-ti 'stuff oneself full', ka-a-ar-ti 'in the heart', ua-a-šu 'well', could form the basis for such an argument. However, numerous forms exist without plene writing where such a hypothesis would lead us to expect it, and in some forms (e.g. šu-ú-na-at 'fill' < *su-néh-t, see Oettinger 1979: 159) plene writing is found in syllables which did not bear PIE accent. Moreover, the fact that Luwian attests a long vowel in the same form (a-aš-du third person singular imperative of 'be', Oettinger 1979: 561) indicates that the lengthening process may have been common to both Palaic and Luwian. If so then we should probably abandon the hypothesis suggested above and seek a solution which can be shared by Palaic and Luwian.

As far as it is possible to tell there is no evidence for a common phonological process lengthening vowels in Palaic and Luwian. As a result it is probably best to attempt a morphological solution.

Palaic shows evidence for a phonological process monophthongizing diphthongs: *éi > í in ki-i-ta-ar, *eu > ü in lu-(u-)ki-(i-)it (if from full grade of the root *leug-) *ou > ü (if from *mousei 'stuff oneself full' after Eichner 1975: 86 footnote 6; for possible etymology compare Greek μῶ 'close lips' and αυστί '(drink something) in one pull'). In Luwian there is evidence at least for the development *éi > í (Oettinger 1979: 535-536). Such monophthongization processes may provide a key to the length of the root vowel in Palaic a-as/a-aš-du and Luwian a-aš-du. Such phonological developments in amphikinetic -mi verbs with a root shape TEUT would have resulted in a restructured ablaut pattern: EU : U \implies \bar{U} : U. On the basis of such a pattern it is possible to imagine that ablaut was restored to verbs that had previously had it eliminated due to the sound change *ǵ > ǰ : *es- : *as- \implies *as- : *as-. The suggested developments may thus be sketched: *es- : *as- \implies *as- : *as- \implies āš- : aš- (Palaic a-aš/a-aš-du : a-ša-an-du/a-še-en-du²⁶) on the basis of verbs with \bar{U} : U ablaut. If such a development is considered plausible then one of the major stumbling blocks to the claim of a phonological development *ǵ > ǰ has been removed.

6. Finally the forms in Classes B1. 2. 3., E, and H must be interpreted in light of the developments proposed thus far. Some variation in details may again be possible.

It seems clear that the forms of Class B1. are to be derived from pre-forms with long *ē vocalism: ú-e-ir-ti < *uérti; -Ci/e-eš < *-ēs < *-eies (compare the situation in Latin where the contracted i-stem nominative plural ending was generalized as the ending of C-stem nouns, e.g. ducēs 'leaders'); occasional plene writings in Hittite may be offered as support for such a derivation.²⁷ The Class E form te-e-ka-an-za/te-ta-a-an-za may also be derived from a pre-form with long ē vocalism, provided we assume that the -ta- in te-ta-a-an-za is a scribal error (see Watkins 1978: 310). If we assume that te-e-ka-an-za is the correct representation for this word then long ē vocalism is not an unreasonable assumption in light of the fact that the consonant *k has apparently been lenited to g after ē (for the lenition of consonants after ē see H. Eichner MSS 31: 79 ff.). The remaining Class E forms (-) e-es-ha etc. can be used as evidence to support the development *ē > ē since plene writing in two

of the alternants indicates that the pre-form was probably *h₁ésh₂r. The Class H form e-eš-ta 'was' may actually be a legitimate Palaic form. Since we are arguing that PIE *ē, PIE*eh₁e > Proto-Anatolian *ē, and PIE *eie > Proto-Anatolian *ē remain in Palaic there is no basis upon which to claim that this form is a Hittite loan.

Class B 2. and 3. forms ú-i-te-si/ú-i-ti-si, in light of the developments proposed thus far, must be interpreted as coming from a pre-form *uedhēsi. Thus the second syllable of this form may contain a special development of *ē, if Oettinger is correct in assuming that long vowels were shortened in unaccented syllables in Proto-Anatolian (see Oettinger 1979: 36 and 125; see also H. Eichner 1980: 163). But, in order to maintain this position, it is necessary to claim that the vowel which resulted from this Proto-Anatolian shortening process was phonetically different than original *ē since only *ē undergoes the depalatalization process.

The phonological developments for PIE *ē, PIE *eh₁e > Proto-Anatolian *ē, PIE *eie > Proto-Anatolian *ē in Palaic are:

- 4. *ē, eh₁e > *ē, *eie > *ē, > ē: ú-e-ir-ti 'says' < *uērti; -Ci/e-eš < *-ēs < *-eies 'nominative plural'; (-)e-eš-ha etc. 'blood' < *h₁ésh₂r; te-e-ka-an-za 'flowing' < *tēkonts; and possibly e-eš-ta 'was' < *e-h₁es-t (see H. Eichner 1975: 78), if this is a legitimate Palaic form.²⁸
- 5. *ē > ē̃ (a raised mid-palatal vowel) in unaccented position: possibly ú-i-te/ti-si 'you build' if from < *uédhēsi.

7. Such a series of phonological developments as those suggested above necessitate a revision in the number of innovations shared by Palaic and Luwian. While it is still possible to claim that *ē̃ > ĩ after j and [+ back] consonants (note the addition of u to the context in which this special development takes place) and that *ē̃ > ǎ in all other phonetic contexts it is no longer possible to claim that Palaic and Luwian share a common development for the long mid-palatal vowel. It will now be necessary to argue that *ē̃ remained during the common period of development and that the change of *ē̃ > ĩ/ĩ is actually a Luwian innovation.

Footnotes

*I would like to thank Professor Brian Joseph for taking the time to critique earlier versions of this paper. His assistance was invaluable. I would also like to thank Dr. Heiner Eichner for his perceptive comments on an earlier version of this same paper. Dr. Martin Peters has kindly agreed to allow this paper, which will appear in Die Sprache (1983), to be prepublished in OSU WPL.

1. For the special treatments of *ě and *ē in Hittite see Oettinger 1979: 448 and 533-545 and H. Eichner 1980: 144 footnote 65.

2. For the special treatment of Luwian ě see Oettinger 1979: 535.

3. It is difficult to determine Oettinger's stand on a-aš/a-aš-du because he doesn't use it directly as evidence for *ě > a (for which see page 535). However, on page 558, in his survey of Anatolian verbal classes a-aš/a-aš-du are considered examples of Inflectional Class = Hittite I 1 a, i.e. -mi conjugation without stem ablaut (for which in Hittite see Oettinger 1979: 184 ff.).

4. For a discussion of the phonetic reality behind the plene writing of vowels see Oettinger 1979: 533 ff. (For a discussion concerning the Hittite writing system and its representation of Hittite phonology see Eichner 1980.) The assumption that length is indicated by plene writing forms the foundation for a number of hypotheses concerning the development of the Anatolian phonological system in Eichner 1971, 1973, 1975, 1980 and Oettinger 1976, 1979, and 1979a.

5. It may be plausible to suggest that the accent of verbs of the pehutē Class shifted their accent from the preverb to the verbal root on the basis of the accent of the simplex *dhéh₁- (for simplex forms see Oettinger 1979: 109).

6. For additional examples of shortening of long vowels in unaccented syllables see Eichner MSS 31: 76-79 and 1980: 161-163. Eichner claims this process is Proto-Anatolian.

7. The possibility that Palaic may have generalized *-ēs < *eies at the expense of *-es was suggested to me by H. Eichner (personal communication).

8. For most of the forms in Pokorny (IEW 1059) a labio-velar is required. Watkins argues that a labio-velar may well have been generated in dialectal IE times from adjective forms in -u (see Pokorny *teku- and thematized *tekuo-. For parallel cases see Watkins (1978: 311).

9. For epenthesis in Hittite in -Csk- clusters see Oettinger 1979: 318. Oettinger claims the epenthetic vowel in these cases has the quality [e]. For development of the first laryngeal see Benveniste 1935: 49 and also Eichner 1975: 95.

10. Tischler (1980: 553) discusses the suffix -kat. The fact that the root has e-grade vocalism is the result of leveling. The original paradigm had o-grade in strong cases and e-grade in weak cases; see H. Eichner 1979: 59.

11. For the etymology of this word see Tischler 1980: 592-593 and Oettinger MSS 34: 129-130.

12. For a discussion of this Palaic form and its Hittite counterpart see Oettinger (1979a: 197-204).

13. Puhvel (1965: 240) cites the Palaic root ahu-, a-hu-ua-(a-)an-ti 'they drink', as corresponding to the Hittite root eku- 'drink'. But there is no evidence that the vocalism in the plural in Palaic was *ě since the corresponding Hittite plural form shows a-vocalism also: a-ku-ua-an-zi/a-ku-an-zi 'they drink'. As a result Puhvel's correspondence is not valid. Rather Palaic ahu- is to be compared to Hittite aku-. The same is also true of Palaic at-/ad- 'eat' and Hittite ad- since once again the Palaic forms are only attested in the plural: a-ta-a-an-ti (2x)/a-da-a-an-ti (1x).

14. The e-vocalism in *ĝenu -g(h)od > gi-nu-kat must be explained as generalization of *ě from weak cases.

15. Oettinger (1979a: 201 footnote 25) admits that this correspondence is not sure. See also the brief discussion in Kammenhuber (1959: 38-39 and footnote 3 page 38).

16. For further discussion of the formal and functional advantages of such an etymology the reader is referred to Professor Joseph's article.

17. For a similar extension of the meaning 'break up' compare English 'break up' in the sense 'to lose control of oneself': He was all broken up (i.e. with grief) by the death of his aunt.

18. For a brief discussion of errors made by Hittite scribes see Watkins (1975 and 1978).

19. It should be noted that the form (-)eš-hu-ur may not even be related to the forms (-)e-eš-ha/(-)e-eš-ha-na. The word division for these two forms seems reasonably certain since the preceding form appears to be a verb with first plural and second singular inflectional ending respectively. This is not the case for (-)eš-hu-ur. In addition, it has been pointed out to me by Professor Joseph that this form may actually be a *uer/n stem with *-ur > -ur (for which see Eichner MSS 31: 73-76).

20. Oettinger (1979a: 201) points out the questionable nature of this form.

21. These forms, if related, may ultimately come from the PIE root *sep- 'hold in esteem' (IEW 909 and compare Sanskrit sapati 'woo', Avestan hap- 'support' and Latin sep-eliō 'bury'). The meaning of the Palaic word, 'purifies', is a reasonable extension of the basic meaning suggested by Pokorny.

22. It is doubtful that it could be argued, in defense of Watkins, that the quality of the epenthetic vowel in this case was [i] and hence irrelevant for a discussion of the development of e-vocalism (contra [i] see Oettinger 1979: 318). On the other hand one could suggest that the change *ě > ĩ was actually mirror-image, i.e. around velars.

23. Oettinger suggests such a possibility in (1979a: 201 footnote 25). But he assumes, wrongly we believe, that the spelling of Palaic ú-e-ir-ti indicates [e^h], a close [e], rather than [ē] because -ua(r)- suggests an original root aorist formation. However, as we have pointed out above, -ua-(r)- may not be related to *uer- and hence an acrostatic accented present (lengthened grade) is at least possible for this verb in Palaic.

24. This comparison was suggested to me by H. Eichner (personal communication).

25. One additional development of *ě may be noted here. It is probably true, as Eichner has pointed out (MSS 29: 28, 37, and MSS 31: 77), that final unstressed *ě was lost, at least under some circumstances, e.g. Palaic ki-i-ta-ar < *keitore and Palaic nu-ú-ku, nu-uk-ku < *nū-kue 'and now' (see Carruba 1970: 65-66).

26. Carruba (1970: 39) claims that the e-vocalism in this form may be the result of a nasalization process, though what type is not made clear. Carruba's suggestion must be considered dubious because of the fact that, since low vowels are more susceptible to nasalization, it would be odd to represent a nasalized low vowel with a symbol for a mid-palatal vowel.

27. This was suggested to me by H. Eichner (personal communication).

28. For a different explanation of length in this form see Oettinger (1979: 90).

References

- Benveniste, Emile. 1937. Origines de la Formation des noms en indo-europeen. Paris: Adrian Maisonneuve.
- Carruba, Onofrio. 1970. Das Palaische. Studien zu den Bogazkoy-Texten 10. Wiesbaden: Otto Harrasowitz.
- Eichner, Heiner. 1971. 'Urindogermanisch *k^we 'wenn' im Hethitischen'. MSS 29: 27-46.
- Eichner, Heiner. 1973. 'Die Etymologie von heth. mehur'. MSS 31: 53-108.
- Eichner, Heiner. 1975. 'Die Vorgeschichte des hethitischen Verbalsystems.' In H. Rix ed. Flexion und Wortbildung, pp. 71-103. Wiesbaden: Dr. Ludwig Reichert.
- Eichner, Heiner. 1979. 'Hethitisch gěnuššuš, ginušši, ginuššin'. In E. Neu and W. Meid eds. Hethitisch und Indogermanisch, pp. 41-62. Innsbruck.
- Eichner, Heiner. 1980. 'Phonetik und Lautgesetze des Hethitischen -- ein Weg zu ihrer Entschlüsselung.' In M. Mayrhofer, M. Peters, and O.E. Pfeiffer eds. Lautgeschichte und Etymologie, pp. 120-165. Wiesbaden: Dr. Ludwig Reichert.
- Friedrich, Johannes. 1952. Hethitisches Wörterbuch. Heidelberg: Carl Winter.
- Friedrich, Johannes. 1974. Hethitisches Elementarbuch I. Heidelberg: Carl Winter.
- Joseph, Brian. 1982. 'Hittite ivar, wa(r) and Sanskrit iva'. To appear in Zeitschrift für Vergleichende Sprachforschung.
- Kammenhuber, Annelies. 1959. 'Das Palaische.' BSL 54: 18-45.
- Oettinger, Norbert. 1976. 'Der Indogermanische Stativ'. MSS 34: 109-150.
- Oettinger, Norbert. 1979. Die Stammbildung des hethitischen Verbums. Nurnburg: Hans Carl.
- Oettinger, Norbert. 1979a. 'sautra- 'Horn', eine hethitische *-tro-Bildung.' In E. Neu and W. Meid eds. Hethitisch und Indogermanisch, pp. 197-204. Innsbruck.
- Porkorny, Julius. 1959. Indogermanisches Etymologisches Wörterbuch, Vol. 1. Bern und München: Francke.
- Puhvel, Jaan. 1966. 'Dialectal Aspects of the Anatolian Branch of Indo-European.' In H. Birnbaum and J. Puhvel eds. Ancient Indo-European Dialects. Berkeley and Los Angeles: University of California Press.

- Sturtevant, Edgar H. 1933. A Comparative Grammar of the Hittite Language. Baltimore: The Linguistic Society of America.
- Tischler, Johann. 1977-80. Hethitisches Etymologisches Glossar, Vols. 1-3. Innsbruck.
- Watkins, Calvert. 1969. Indogermanische Grammatik, Vol. 3.1. Heidelberg: Carl Winter.
- Watkins, Calvert. 1975. 'Die Vertretung der Laryngale in gewissen morphologischen Kategorien in den indogermanischen Sprachen Anatoliens.' In H. Rix ed. Flexion und Wortbildung, pp. 358-378. Wiesbaden: Dr. Ludwig Reichert.
- Watkins, Calvert. 1978. 'A Palaic Carmen.' In M.A. Jazayery, E.C. Polome, and W. Winter eds. Linguistic and Literary Studies in Honor of Archibald Hill, Vol. 3. Mouton: The Hague.

Diachronic and Synchronic Tone Rules in the Etsako Verbal System:
Some Theoretical Implications*

Donald G. Churma

There has been a good bit of discussion recently concerning the tonal system of Etsako¹, from both a diachronic (Elimelech 1974) and a synchronic (Elimelech 1978, Leben 1978, Clements 1979) perspective, in an attempt to account for some rather complicated synchronic tonal alternations, especially in the verb phrase. In this paper, I will attempt a diachronic account which is considerably more comprehensive than the previous account, as well as, I will argue, more plausible with respect to the areas which have been treated by Elimelech. I will then go on to examine the synchronic system which has resulted from the changes in question, arguing for an analysis which is different in theoretically interesting respects from all previously proposed analyses. Finally, I will pursue the implications of my diachronic and synchronic accounts for the relationship between synchrony and diachrony in tonology in general, especially with respect to non-segmental theories of tone. In particular, I will argue that there are certain kinds of synchronic tone rules that have no formal analogue in rules that express diachronic changes. Based on this discrepancy, I will argue that tonal "autosegments" (Goldsmith 1976) are actually mapped onto tone-bearing units, as originally proposed by Leben (1973), and not merely "associated with" them, as Goldsmith (1976) and most other current researchers--including Leben (1978)--maintain, and explore the consequences of accepting this proposal with respect to what Goldsmith calls "tonal stability". I will also argue that the synchronic system of Etsako provides good evidence against Leben's (1973, 1978) "Obligatory Contour Principle".

1. The data. I will be concerned here almost exclusively with the verbal system, especially the behavior of verb + noun object sequences, since the rest of the tonology appears to be fairly straightforward, even from a synchronic perspective². The following data, all involving a third person singular pronominal subject, are taken from Elimelech (1978:85-107), with the exception of those followed by a question mark, which have been inferred from his rules and the behavior of similar forms. They represent the forms found in the "tenses" listed in the affirmative (I, II) and the negative (III, IV). The five nouns used for illustration have the forms: útsádè 'put', àkpà 'cup', ògèdè 'banana', àtásà 'plate', and úkpò 'cloth'; the verb stems which are representative of most of the one and two syllable verbs in the language, do not occur in isolation, but are presumably underlyingly dé 'buy' and kélé 'look for' (cf. the infinitive forms údémhì and úkélémhì), with the final vowel being "elided" before the vowel-initial object noun.

Table 1

	Present	Future	Past	Past Perfect	Habitual	Past Hab.
I.	ò dùtsádẹ̀	ò Ọ́á dùtsádẹ̀/ 'dùtsádẹ̀	ọ̀ dùtsádẹ̀	ọ̀ xẹ̀ dùtsádẹ̀	ò dùtsádẹ̀	ò yá dùtsádẹ̀/ 'dùtsádẹ̀
	dàkpá	dàkpá/ 'dàkpá	dàkpà	dàkpà	dàkpà	dàkpá/ 'dàkpá
	dọ̀gẹ̀dẹ̀	'dọ̀gẹ̀dẹ̀	dọ̀gẹ̀dẹ̀ ?	dọ̀gẹ̀dẹ̀ ?	dọ̀gẹ̀dẹ̀ ?	'dọ̀gẹ̀dẹ̀
	dá'tásà	'dá'tásà	dâtásà	dâtásà	dâtásà ³	'dá'tásà
	dúkpò	'dúkpò	dúkpò	dúkpò	dúkpò	'dúkpò
II.	kèlùtsádẹ̀	kèlùtsádẹ̀	kélútsádẹ̀	kélútsádẹ̀	kélútsádẹ̀	kèlùtsádẹ̀
	kèlàkpá	kèlàkpá	kélâkpà/ kélâkpà	kélâkpà/ kélâkpà	kélâkpà/ kélâkpà	kèlàkpá
	kèlọ̀gẹ̀dẹ̀	kèlọ̀gẹ̀dẹ̀	kélọ̀gẹ̀dẹ̀/ kélọ̀gẹ̀dẹ̀ ?	kélọ̀gẹ̀dẹ̀/ kélọ̀gẹ̀dẹ̀ ?	kélọ̀gẹ̀dẹ̀/ kélọ̀gẹ̀dẹ̀ ?	kèlọ̀gẹ̀dẹ̀
	kèlá'tásà	kèlá'tásà	kélâtásà/ kélâtásà	kélâtásà/ kélâtásà	kélâtásà/ kélâtásà	kèlá'tásà
	kélúkpò	kélúkpò	kélúkpò	kélúkpò	kélúkpò	kélúkpò
III.	ò dùtsádẹ̀	ò Ọ́ǎ dùtsádẹ̀	ọ̀wà dùtsádẹ̀	ọ̀wà xě̀ dùtsádẹ̀	ò dùtsádẹ̀	ò yǎ dùtsádẹ̀
	dàkpá					
	dọ̀gẹ̀dẹ̀	Tone patterns in other tenses are same as in present for verb-object combinations				
	dá'tásà					
	dúkpò					
IV.	kèlùtsádẹ̀					
	kèlàkpá					
	kèlọ̀gẹ̀dẹ̀					
	kèlá'tásà					
	kélúkpò					

Clearly, the tone patterns in the case of the present, future, and past habitual in the affirmative, and of all "tenses" in the negative, are not what one would expect on the basis of the infinitival forms of the verbs and the isolation forms of the nouns; but the remaining three tenses show precisely the expected behavior. Furthermore, those tense morphemes which have a surface segmental realization (as opposed to the purely tonal marking of the distinction between past and habitual in the affirmative) invariably show a rising tone in the negative, while the affirmative future and past habitual show high tone, and past perfect has low tone⁴. Note also that verb-object sequences show no tonal differences from tense to tense in the negative, unlike in the affirmative. It is these three sets of facts with which my discussion will be primarily concerned.

2. The diachronic origin of the alternations.

2.1. Elimelech's account.

Elimelech (1974) has proposed a rather sketchy analysis of the development of what he calls there the past, present, and customary constructions, where these correspond to his later past, present progressive, and habitual, respectively. His account posits earlier forms of the verb stems and nouns identical to the synchronic forms mentioned above, and high and low toned vowels as markers of the past and present tenses, respectively, with the customary having no overt marker. The development of the customary is then straightforward: the only change is the loss of the final vowel of the verb stem and the appearance of its tone on the following vowel (as predicted by the non-segmental theories mentioned above). The past is also fairly straightforward: after the loss of the vocalic marker of past tense, the stranded high tone is associated with the pronominal morpheme to its left, resulting in the rising-toned pronoun which is characteristic of the past tense. The stranded tone is associated leftward, rather than rightward, in order to prevent merger of the past tense with the customary. The posited changes are schematized in (1) and (2) where \hat{v} represents the pronoun, \hat{v} the past marker and $c\hat{v}$ and $\hat{v}c\hat{v}$ represent the verb stem and a low-low noun⁵.

$$(1) \text{ Customary: } \hat{v} + \phi + c\hat{v} \div \hat{v}c\hat{v} > \hat{v} + c\phi + \hat{v}c\hat{v}$$

v cvcv

$$(2) \text{ Past: } \hat{v} + \hat{v} + c\hat{v} + \hat{v}c\hat{v} > \hat{v} + \hat{\phi} + c\hat{\phi} + \hat{v}c\hat{v} > \hat{v} c\hat{v}c\hat{v}$$

The present tense is a bit more complicated, and Elimelech treats low-low nouns separately from the others, although the loss of the vocalic present marker and subsequent assignment of the stranded low tone rightward is common to the evolution of all examples. The suggested evolution of this construction with a high-low noun is that given in (3), where the second \hat{v} represents the present tense marker.

$$(3) \hat{v} + \hat{v} + c\hat{v} + \hat{v}c\hat{v} > \hat{v} + \hat{\phi} + c\hat{\phi} + \hat{v}c\hat{v} > \hat{v} c\hat{v}c\hat{v} > \hat{v} c\hat{v}c\hat{v}$$

In the case of low-low nouns, and only in the case of such nouns (and low-low-low nouns), a change of "high tone spread" plays a role in addition to changes similar to those which occurred in the case of the past tense. Elimelech's account is schematized in (4), where the fourth stage is the output of "high tone spread".

$$(4) \quad \hat{v} + \hat{v} + c\acute{v} + \hat{v}c\hat{v} > \hat{v} + \check{\emptyset} + c\acute{v} + \hat{v}c\hat{v} > \hat{v} + c\check{v} \hat{v}c\hat{v} > \\ \hat{v} + c\check{v} \acute{v}c\acute{v} > \hat{v} + c\check{v}c\acute{v} > \hat{v} + c\hat{v}c\acute{v}$$

The changes postulated are for the most part quite plausible ones, and the basic insight about the earlier existence of vocalic tense morphemes is, I believe, correct in a broad sense. However, there are a number of problems with the details of this account. First of all, it is incomplete (although, to be fair, Elimelech did not intend a complete account), in that only three tenses and monosyllabic verb stems are treated; in particular, no negatives are considered, and the differences between the negative and the affirmative are among the most striking facts about the verbal morphology of Etsako. Secondly, as Elimelech (1974: 71) suggests, this account entails that at an earlier stage "the customary is the most basic [i.e., least marked--DGC] of the constructions considered..." Since it is the simple present which is typically unmarked (cf. Tiersma 1982 and the references cited there), an analysis which sets up another category as unmarked should be carefully scrutinized. More serious are the numerous appeals to the functional notion of avoidance of merger, especially as a trigger for a phonological change. While it seems clear that otherwise expected changes can be **blocked** if they would result in merger (cf. for example, Campbell 1974a), a change which occurs solely for the purpose of preventing another change from causing a merger is attested nowhere else, to the best of my knowledge. Such a change appears to be a priori extremely unlikely, and I would like to suggest that this is not a possible mechanism of phonological change. Furthermore, the postulated change of "high tone spread", which is said to have such a motivation, does not appear to be a "natural diachronic tone rule" in the sense of Hyman and Schuh (1974), unlike the other changes suggested (although synchronic rules of this nature are not uncommon). Thus, not only does Elimelech's account require a change which has an extremely questionable functional motivation, but the change required is itself not a natural one. Finally, the putatively functionally motivated direction of association of the tones stranded by vowel elision (leftward in the past tense, rightward in the present) violates in the case of the present tense, an apparent universal (Leben 1978, Clements and Ford 1979), whereby stranded tones are associated with the trigger of the rule which resulted in their being stranded.

2.2. The present account.

It seems clear that the evolution of the Etsako tonal system cannot be exactly as Elimelech has suggested--even within the limited domain which he treats--and I will argue here for an alternative, and fuller, account. My account agrees with that of Elimelech as far as the reconstructed form of nouns and verb stems is concerned, and our accounts are also in agreement with respect to the reconstruction of the past tense morpheme. It also seems quite likely that Elimelech is correct with respect to his reconstruction of the "present progressive" (i.e., simple present--cf. note 6) morpheme, although a not terribly plausible alternative is discussed below

(cf. note 10). There is fairly good evidence, however, that there was in fact an overt habitual marker--as one would expect on the basis of the considerations presented above--namely a rising-toned vowel (probably *a--see below). These reconstructions, together with those concerning morphemes not considered by Elimelech, are summarized in (5), where v indicates a vowel whose quality is not reconstructible:

- (5) present--*ṽ; past--*ṽ; habitual--*ṽ; future--*θâ; past
perfect--*ṽ x̃; past habitual--*yâ; negative--*ṽ (clause
initial)

The evidence in favor of the low-high tone sequence as a mark of the habitual concerns the various kinds of aspectual (?) modifications of past tense. Assuming that such forms were once transparently past tense + some other morpheme, and that modern y- in the past habitual comes from earlier *i--presumably the past tense morpheme, which of course bore high tone--then considerable further reconstruction is possible. The past perfect seems to be readily interpretable in this way, with *xe marking "perfective aspect" and if the vowel of the past marker was *i and the habitual marker was *a, then after simplification of *a to *â when followed by the high-toned verb stem in accordance with the sound change posited in (6b), we get intermediate *f â, which after the devocalization of i alluded to above would yield the form given in (5). Given such a further reconstruction, we would have an explanation for the lack of a rising-toned pronoun in the past habitual, as long as devocalization (with concomitant association of the stranded tone to the following vowel, as required by the universal mentioned above) preceded the elision of tense morphemes required by (6e) below. In addition, of course, it would provide an earlier phonologically isolable past tense morpheme. One might question such a reconstruction on the basis of the behavior of the negative forms, since -wâ, which appears to be the (negative) past tense morpheme--appearing in both the simple past and the past perfect--is not present in the past habitual negative. Thus, it could be argued, the lack of -wâ in the past habitual negative indicates that the past habitual never was--at least insofar as internal reconstruction is able to ascertain--composed of phonologically isolable past tense and habitual morphemes. However, given the other oddities with respect to the negative in general, such as the presence of -wâ in any of the tenses, as well as the lack of tense to tense tonal distinctions and the rising-toned tense markers, the behavior of negative forms cannot be taken as very good evidence about the morphological composition, either diachronic or synchronic, of the corresponding affirmative forms.

The basic idea behind these reconstructions is that the unexpected tone patterns (i.e., those in the present, future, and past habitual, and in the negative) came about as the result of a rightward "spreading" of low tones from tense markers onto the verb stem together with a number of subsequent changes specified below. The reason why the past, past perfect, and habitual fail to show the effect of this spreading is the presence of a high tone--either as a simple high-toned vowel in the case of the past, or as part of a rising tone, as in the other two cases--which has the effect of blocking low-spreading. The habitual and past perfect are reconstructed with a rising tone, rather than with a simple high, in order to explain the differential behavior of these morphemes and the past tense morpheme with respect to the tone of the pronominal element. The rationale behind the

reconstructions should become clearer in the sketches of the diachronic development of various affirmative forms given in (7), where the effects of the sound changes given in chronological order in (6) are illustrated (and where '-' represents a morpheme boundary and tones joined by a ligature are attached to a single tone-bearing element).

- (6) a. *L - H > L - LH
- b. * $\widehat{\alpha H} - \alpha H$ (#) $-\alpha H > \alpha H$ (#) $-\alpha H$ (i.e., *HL L > HL, *LH H > L H)
- c. * $\widehat{LH} > H$ (= 'H after H, due to the application of "downdrift" --cf. Churma 1982)
- d. *L H > L \widehat{LH} (morpheme-internally only)
- e. "elision" of reconstructed v's and final vowels of verb stems

- (7) a. Pres.: * ∂ # \check{v} - $d\check{e}$ # $\acute{u}ts\acute{a}d\grave{e}$ * ∂ # \check{v} - $d\check{e}$ # $\partial g\grave{e}d\grave{e}$
- | | | | | | |
|--|-----------------------------------|----|----------------------------------|--------------|--|
| | 6a | | | | |
| | $d\check{e}$ | 6b | | $d\check{e}$ | |
| | $d\grave{e}$ | 6c | | $d\grave{e}$ | |
| | | 6d | | \grave{e} | |
| | $d\check{u}ts\acute{a}d\grave{e}$ | 6e | $d\partial g\grave{e}d\grave{e}$ | \grave{e} | |
-
- b. Fut.: * ∂ # $\theta\acute{a}$ - $d\check{e}$ # $\acute{u}ts\acute{a}d\grave{e}$ * ∂ # $\theta\acute{a}$ - $d\check{e}$ # $\partial g\grave{e}d\grave{e}$
- | | | | | | |
|--|-----------------------------------|----|-----------------------------------|---------------|--|
| | 6a | | | | |
| | $d\check{e}$ | 6b | \acute{a} | $d\check{e}$ | |
| | \acute{a} $d\check{e}$ | 6c | \acute{a} | $'d\check{e}$ | |
| | | 6d | | \grave{e} | |
| | $d\check{u}ts\acute{a}d\grave{e}$ | 6e | $'d\partial g\grave{e}d\grave{e}$ | \grave{e} | |
-
- c. Hab.: * ∂ # \check{v} - $d\check{e}$ # $\acute{u}ts\acute{a}d\grave{e}$ * ∂ # \check{v} - $d\check{e}$ # $\partial g\grave{e}d\grave{e}$
- | | | | | | |
|--|-----------------------------------|----|----------------------------------|-------------|--|
| | 6a | | | | |
| | \check{v} | 6b | \check{v} | | |
| | | 6c | | | |
| | ∂ | 6d | ∂ | \grave{e} | |
| | $d\check{u}ts\acute{a}d\grave{e}$ | 6e | $d\partial g\grave{e}d\grave{e}$ | \grave{e} | |
-
- d. Pres.: * ∂ # \check{v} - $k\acute{e}l\acute{e}$ # $\acute{u}ts\acute{a}d\grave{e}$ * ∂ # \check{v} - $k\acute{e}l\acute{e}$ # $\partial g\grave{e}d\grave{e}$
- | | | | | | |
|--|---|----|--|-------------|--|
| | 6a | | | | |
| | $k\acute{e}$ | 6b | $k\acute{e}$ | | |
| | $k\grave{e}$ | 6c | $k\grave{e}$ | | |
| | | 6d | | \grave{e} | |
| | $k\acute{e}l\grave{e}$ | 6e | $k\acute{e}l\grave{e}$ | \grave{e} | |
| | $k\acute{e}l\grave{u}ts\acute{a}d\grave{e}$ | | $k\acute{e}l\partial g\grave{e}d\grave{e}$ | \grave{e} | |

e. Hab.:	*ò # ǎ - kélé # útsádè		*ò # ǎ - kélé # ògèdè
	_____	6a	_____
	ǎ	6b	ǎ
	_____	6c	_____
	ò	6d	ògèdè`
	kélútsádè	6e	ò kélògèdè`

The final forms are not those found synchronically, of course. One problem concerns the rising-falling tone on ògèdè produced by the operation of (6d), which should be simply falling. Incorrect rising tones due to (6d) are also found in the útsádè forms in (7a, b, d), where we should have for (7b) variants with either low or (downstepped) high, while for the others we should have only one possibility--low tone. The latter problem carries over to the synchronic analyses considered below, and further discussion will be postponed until then. As for the former, it can be handled by positing a subsequent change such as that given in (8):

$$(8) \quad *L \widehat{LHL} > L \widehat{HL}$$

This is not at all implausible, given the highly marked nature of rising-falling tones. In fact, since such tones do not, to the best of my knowledge, occur at all on short vowels (although they, and even tones with four components--cf. Lovins 1971a--can be found on long vowels), (6d) may have been blocked by the presence of the falling tone. (If so, then of course (8) is unnecessary).

As far as the affirmative is concerned, only one further aspect requires comment.¹⁰ Forms such as àkpà in the present, future, past habitual, and the negatives seem to behave quite anomalously with respect to the account suggested above. The expected development for the present tense is sketched in (9):

(9) Pres.:	*ò # ǎ - dè # àkpà	
	_____	6a
	è	6b
	_____	6c
	é	6d
	ò	6e
	dàkpà	

This is not even close to being correct for the present (although it is precisely correct for the habitual!). If àkpà had had high, rather than low, tones prior to the sequence of events outlined in (6)--but only in the anomalous tenses--then something quite close to the correct output is obtained, as illustrated, again for the present, in (10):

(10) Pres.:	*ò # ǎ - dè # ákpá	
	_____	6a
	è	6b
	_____	6c
	é	6d
	ò	6e
	dǎkpá	

I will assume that changes which had the effect of converting low tones to high on nouns which bore only low tones in the present, future, and past habitual had taken place prior to the events in (6), although exactly what these changes were and why they took place is far from clear. The inappropriate rising tone will be treated, as before, as a synchronic problem. There is a generalization to be found concerning which tenses exhibit these changes: precisely those reconstructed with a low tone (either as a simple low tone or part of a falling tone) immediately preceding the verb stem. It seems almost as if, loosely speaking, this low tone, after having been spread onto the verb stem by (6a), "dislodged" the high tone of the latter, which in turn "dislodged" the low tones of these nouns. Even within such a figurative account, however, it is not clear why only nouns which bear exclusively low tones (and not, e.g., *ògèdè*) participate in the "dislodging". Apart from such cases, however--which, it should be recalled, were not given a convincing account by Elimelech, either--all of the forms in the affirmative have been satisfactorily accounted for.

Let us now consider the negative forms. Given the reconstructed forms in (5), we would expect the verb-noun forms to show precisely the same tonal patterns as the corresponding affirmative forms, since the only difference is the presence of the negative morpheme, which is not adjacent to the verb-noun forms. Only in the case of present tense, however, are such forms found (although the future and the past habitual show one of the affirmative variants in the forms for which there is variation). Since the tone patterns found in the negative have been seen to be the result of a low toned present tense morpheme in the affirmative (note that they are identical to not only the negative present, but also the affirmative present), it is tempting to propose that the negative tone patterns are also due to the former presence of a low toned vowel which was part of a discontinuous negative morpheme and directly preceded the verb stem. The existence of the low toned *-wà* in the past and past perfect, furthermore, appears to support such a proposal, although it would be unclear why it surfaces in just these tenses.

In the case of the past perfect, however, this *-wà* is in the wrong place, since the tense morpheme, and not *-wà*, is adjacent to the verb stem. Furthermore, this proposal incorrectly predicts an initial downstepped high tone in the future and the past habitual, rather than the simple high tone actually found, as illustrated for *àtásà* in (11); where *ɔ...ɔ* is the discontinuous negative morpheme:

- (11) * *ɔ # ò # ɔá # ɔ̃ -dɛ # àtásà*
- | | | |
|--------------------|-----------|----|
| | <i>ɛ̃</i> | 6a |
| | <i>á</i> | 6b |
| | <i>'é</i> | 6c |
| | <i>ǎ</i> | 6d |
| <i>ɔ ɔá dātásà</i> | | 6e |

Assuming that the falling tone on the future marker, like the contour tones found in the verb-noun combination, is simplified by a synchronic rule to high tone, we would get the incorrect *ò ɔá 'dǎ'tásà*.

We must, therefore, find an alternative explanation for the negative forms. The only reasonable alternative appears to be analogical leveling

in the negative "superparadigm": using the morphologically least marked present tense as the basis for the analogy, Etsako speakers leveled out the tense-to-tense differences in verb-noun tonal patterns. (The result, in the case of the habitual, was the loss of the only thing that distinguished it from the present.) For some reason, it appears, the tone patterns on the verb-noun combinations were taken to be (part of) the means of indicating negation. Why this should be so is not clear, but there seem to be at least partial (synchronic) parallels in other languages; Larry Hyman and David Stampe (personal communication) have noted that in Haya and in Sora there is a similar striking reduction in tense-to-tense contrasts in the negative. It could well be that this synchronic impoverishment is due to the kind of diachronic development suggested here.

Only the rising tones on the negative tense morphemes remain unaccounted for. The future and the past habitual can be accounted for by positing the following change, which postdated those given in (6):

$$(12) \quad \widehat{HL} \# H > \widehat{HL} \# \widehat{LH}$$

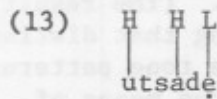
Note that it would not do to describe the change of H to LH as occurring after a L, since this would result in incorrect rising tones on tense morphemes in the affirmative. Nevertheless, the past perfect--where there is no preceding falling tone--has a rising tone. This is apparently another case of analogy. The result of these occurrences of analogy is that in present day Etsako, the negative is indicated in some cases (the future, past perfect, and past habitual) doubly redundantly--by the rising-toned tense marker and the verb-noun tone patterns, in addition to the falling toned pronoun/-wà.

To summarize, the account just described has several advantages over that of Elimelech. First of all, of course, it accounts for the development of a much greater number of forms, including the negative forms, which underwent a rather interesting leveling of tonal contrasts from tense to tense. Even in the case of the tenses treated by Elimelech, some improvement has been made, since no reference to functional factors is required in order to determine the direction of association of stranded tones--in violation of an otherwise valid universal (one which has a quite plausible motivation--cf. section 4). It also allows for a stage in the development of the language which conforms to Greenberg's (statistical) universal concerning the unmarked tense. Furthermore, the one kind of case which was not given a satisfactory treatment--all-low-toned nouns in the problematic tenses--was not satisfactorily accounted for by Elimelech. An explanation for the development of these must apparently await further comparative evidence; at any rate, they have so far successfully resisted both Elimelech's attempted internal reconstruction and my own.

3. The synchronic system. In this section, I will examine the tonology of the Etsako verbal system from a synchronic perspective. After pointing out inadequacies of previous treatments, I will propose a reanalysis which avoids these problems, and requires a rule which is of considerable theoretical interest, since it violates what Leven (1973) calls the "Obligatory Contour Principle".

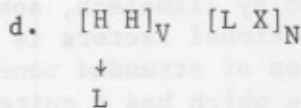
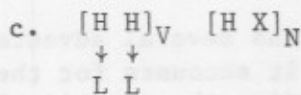
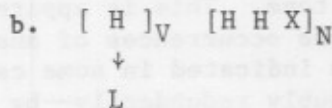
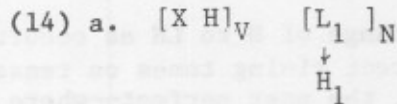
3.1. Elimelech's analysis. Elimelech's account assumes a nonsegmental theory of tone which does not allow for the possibility of mapping a single

tone onto more than one tone-bearing unit. The underlying representation of útsadê, for example, would be:



Tones are mapped onto tone-bearing units by a (presumably universal) set of rules which need not concern us here, since his system is, at least for present purposes, descriptively equivalent to a segmental theory of tone, with the exception that tones remain when the segment that bear them are lost. For reasons of space, I will discuss his analysis using such a quasi-segmental system.

The core of Elimelech's treatment of the verbal system is a set of four tone rules which apply in the present, future, and past habitual tenses, and in the negative (pp. 87-9). I give below the rephrasal of these rules given in Leben (1978:183):



(14a) is intended to have the effect of raising all of the L's in a noun object which consists entirely of low tones (although it is not clear, without some kind of convention for reading such a notation, that this rule would not also affect the initial L's in, say, a (LLH noun); (14b) lowers the H of a monosyllabic verb stem when it is followed by a noun which begins with at least two H's; (14c) lowers both H's of a bisyllabic verb stem when a noun with initial H follows; and (14d) lowers the H of only the first syllable when a noun with initial L follows. He also makes use of a set of independently motivated "optional" rules (cf. note 2 and below) which have the following effects (cf. Elimelech 1978:110 for a summary), where D represents downstepped high tone, and H* non-downstepped high:

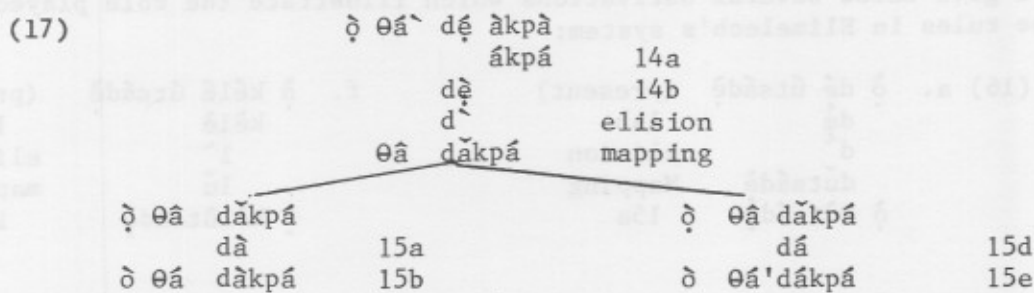
- (15)
- a. $\widehat{LH} \ H^* \rightarrow \ L \ H$
 - b. $\widehat{HL} \ L \rightarrow \ H \ L$
 - c. $H \ \widehat{HL} \rightarrow \ H \ L$
 - d. $\widehat{LH} \rightarrow \ H$
 - e. $\widehat{HL} \ D \rightarrow \ H \ D$ ¹¹

I give below several derivations which illustrate the role played by these rules in Elimelech's system:

- | | |
|---|---|
| <p>(16) a. ọ́ dẹ́ útsádẹ̀ (present)
 dẹ́ 14b
 ḍ elision
 dùtsádẹ̀ Mapping
 ọ́ dùtsádẹ̀ 15a</p> | <p>f. ọ́ kẹ́lé útsádẹ̀ (present)
 kẹ́lé 14c
 ḷ elision
 lù mapping
 ọ́ kẹ̀lùtsádẹ̀ 15a</p> |
| <p>b. ọ́ dẹ́ àkpà (present)
 àkpá 14a
 dẹ 14b
 ḍ elision
 dàkpá mapping
 ọ́ dàkpá 15a</p> | <p>g. ọ̀wà kẹ́lé àtásà (past neg.)
 kẹ́lé 14d
 ḷ elision
 lâ mapping
 ọ̀wà kẹ̀lá'tásà 15e</p> |
| <p>c. ọ́ dẹ́ úkpò (present)
 ḍ elision
 ọ́ dúkpò mapping</p> | |
| <p>d. ọ́ ø dẹ́ àkpà (past)
 elision
 õ dàkpà mapping</p> | |
| <p>e. ọ̀wà dẹ́ àkpà (past neg.)
 àkpá 14a
 dẹ 14b
 ḍ elision
 dàkpá mapping
 ọ̀wà dàkpà 15b</p> | |

Elimelech's analysis succeeds in generating all of the attested surface forms from the input forms he posits. Moreover, the fact that it also generates forms other than those given as surface forms by Elimelech (due to the supposed optionality of the rules in (15)) is common to both Leben's analysis and my own, and I will not consider this matter further; apparently, the only way of preventing overgeneration under **any** analysis is extensive appeal to grammatical conditioning with respect to whether a given rule in (15) is in fact optional, obligatory, or not applicable at all. There are slightly more complicated cases of a similar nature which are unique to Elimelech's analysis (although Leben's account has problems of a different nature in these cases--see below). These problematic cases involve nouns whose first two syllables bear high tone, whether underlyingly or as the result of the operation of (14a), in the future and past

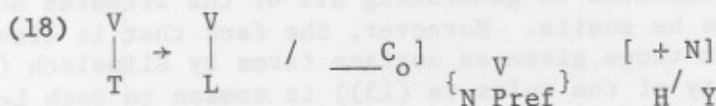
habitual tenses in the affirmative. Elimelech's account in the case of the future with ákpà as object (p. 95) is sketched below:



That is, after the application of all obligatory rules, the "optional" rules are divided into two sets of two rules each. Within each set, the rules apply obligatorily, but either set may be selected to apply; however, one, and only one, of the sets must be selected. In other words, either both (15a) and (15b) must apply, or both (15d) and (15e) must. Clearly, this analysis is highly unusual, and I know of no theory of rule interaction which would permit it; some reanalysis in this respect is obviously in order.

A further, and probably more obvious, problem with this account is the fact that it posits **four** separate rules which refer to an extremely unusual set of non-phonological categories--present, future, and past habitual tenses, and **all** tenses in the negative. Clearly, an account which does not require such extensive reference to an obviously disparate class of grammatical categories such as this is, *ceteris paribus*, to be preferred over one that does.

3.2. Leben's analysis. Leben's account is framed within a theory which is in some respects just the opposite of that implicitly accepted by Elimelech. While in both theories, tone and "ordinary" segmental features are represented on separate tiers (in the sense of Goldsmith 1976), Leben's theory prohibits sequences of like tones in the underlying representation of a morpheme, and contains a "convention on tone melodies" (Leben 1978:181) which simplifies such a sequence to a single tone whenever it arises in the course of a derivation. As a result, he reformulates Elimelech's rules (14b-d) as follows (p. 185):



As in Elimelech's analysis, this rule must be restricted to applying in the present, future, and past habitual tenses, and in the negative. It has the effect of lowering high tones when they occur on a vowel in a verb stem or noun prefix and are followed by a high toned vowel. Together with rules (14a) and the equivalent of (15a, b), most of Elimelech's data can be generated from essentially the same underlying forms as in Elimelech's

account. The following derivations illustrate the effects of these rules in the present tense:

(19) a.	ò dẹ ògèdẹ		b.	ò kélé útsádẹ
	ò d̄ògèdẹ	14a		ò kélútsádẹ
	ò d̄ògèdẹ	elision		ò kèlùtsádẹ
	ò d̄ògèdẹ	15b		ò kèlùtsádẹ
	ò d̄ògèdẹ	18		ò kèlùtsádẹ

This account does in fact handle the present tense forms without difficulty. Furthermore, having replaced (14b-d) with a single rule must undoubtedly count in its favor. There are, however, some forms which are problematic. Some are rather trivial, and can be handled by making use of the remaining rules in (15), but the variant pronunciations in the future and past habitual appear not be amenable to treatment by simply adding another rule, although one of the variants is indeed generated (as long as (15e) is ordered before (18)), as indicated below:

(20)	ò ɛá dẹ àkpà	
	ákpá	14a
	dákpá	elision
	ò ɛá'dákpá	15b
	ò ɛá'dákpá	15e
	ò ɛá dákpa	18

This assumes that (18) applies to downstepped high tones. If it does not, then ò ɛá 'dákpa would be derived--the other variant--but unless it does and is made optional, only one of the variants can be derived. Since there is already an obligatory/optional/inapplicable problem with respect to the rules in (15)--for Leben's analysis, and my own (see below), as well as for that of Elimelech--it might be supposed that making this rule optional would not create any really new problems. However, the rules in (15) are independently motivated as being optional by variant pronunciations in the nominal system, and this clearly cannot be the case for (18), which applies only in the specified tenses. In fact, this unusual morphological restriction is a good indication that this rule **cannot**--because of the semantic task it performs--be optional. It is also worth nothing that this restriction makes it clearly a Stampeian rule (and not a "natural process"), which would entail in Stampe's framework (cf. Stampe 1973, Donegan and Stampe 1979) obligatory application. In addition, making (18) optional would incorrectly entail in the case of bisyllabic verb stems the existence of variants such as present tense *ò kélútsádẹ--an especially striking demonstration of the failure of this suggestion, since this is precisely the (only) form found in the habitual.

One final, though rather minor, problem with this analysis is the further reference made to disjunctive morphological categories in rule (18). This rule thus requires reference not only to the curious set of tenses required by Elimelech's rules, but also to the equally curious set consisting of verbs and noun prefixes. In the next subsection, I will present an analysis which has neither this drawback nor the others pointed out with respect to Elimelech's and Leben's analyses.

3.3. The present analysis. Like both of the previous analyses, mine assumes that tones and segmental features may be represented underlyingly on separate tiers. It also makes crucial use of the possibility of violating Leben's Obligatory Contour Principle (cf. section 3.2).

The replacement for Elimelech's rules (14b-d) and Leben's (18) is suggested by the output of the sound changes proposed in (6). Recall that monosyllabic verb stems ended up with either a high tone or a low tone in the unusual tenses, depending on whether (6b) or (6c) simplified the rising tone that resulted from the occurrence of (6a); bisyllabic verbs were invariably low-rising. Since the synchronic analogues of (6b) and (6c) are independently motivated by alternations in the nominal system (cf. rules (15a, d)), a rule which yields a rising tone on monosyllabic verb stems and on the final syllable of bisyllabic stems (with a low tone on the first syllable) will be able to account for the alternations in question. I suggest the following, which applies in the present, future, and past habitual, and in the negative:

$$(21) \quad \begin{array}{c} X \\ | \\ T \end{array} \rightarrow \begin{array}{c} X \\ L L H \end{array} / \text{--- } C_O]_V N[$$

A few comments about the formulation of this rule are in order. Note first of all that the output contains two consecutive low tones--a clear violation of the Obligatory Contour Principle. Note also that the tones in the output are not linked by "association lines" to any segmental material. This is necessary because the rule is intended to be applicable to both monosyllabic and bisyllabic verb stems; the required association lines will be provided by the universally applicable "Well-Formedness Condition" (WFC) proposed in Goldsmith (1976).¹² It is also worth pointing out that this rule is considerably simpler than Leben's rule (18), since it requires reference neither to noun prefixes nor to a following high toned vowel in a noun, and that it is clearly simpler than the corresponding three rules in Elimelech's analysis.

Before illustrating the application of this rule, I would like to give some attention to the rules in (15), particularly with respect to the future and past habitual forms, which have been seen to be problematic in both Elimelech's and Leben's analyses. In order to avoid the problems with Elimelech's account, it is desirable to collapse (15b) and (15e), if possible. It is indeed possible to do so, since the falling tone is simplified to high in both cases and is required to be followed by the only tones which could possibly follow it (cf. note 11) when the two rules are taken together. They can thus be replaced by the single rule (22):

$$(22) \quad \widehat{HL} T \rightarrow H T$$

(The possibility of collapsing these two rules was apparently overlooked by both Elimelech and Leben, since they both collapsed (15b) with another rule, (15a), by means of alpha variables.) Rules (21) and (22), together with (14a) and (15a, c, d) (ignoring, of course, the optional/obligatory/-inapplicable problem), can account in a plausible way for all of the forms in Table 1, as illustrated below (where the optionality or obligatoriness of rules in a given morphological category is indicated in parentheses):

(23) a.	ò dẹ ògèdè	(present)	b.	ò kélé útsádẹ	(present)
	<u>dẹ</u>	14a		<u>kélé</u>	14a
	dẹgèdè	21		kèlùtsádẹ	21
	ò	elision		ò kèlùtsádẹ	elision
	ò dẹgèdè	22 (oblig.)			22 (oblig.)
		15a (oblig.)			15a (oblig.)
		15d (oblig.)			15d (oblig.)
c.	ò 0á dẹ útsádẹ	(future)	d.	ò 0á kélé útsádẹ	(future)
	<u>dẹ</u>	14a		<u>kélé</u>	14a
	dútsádẹ	21		kèlùtsádẹ	21
	0á	elision		0á	elision
	(ò 0á dútsádẹ)	22 (oblig.)		ò 0á kèlùtsádẹ	22 (oblig.)
	ò 0á'dútsádẹ	15a (opt.)			15a (oblig.)
		15d (oblig.)			15d (oblig.)
e.	ò 0á kélé úkpò	(future)	f.	ò dẹ àkpà	(habitual)
	<u>kélé</u>	14a		<u>dẹ</u>	14a
	kèlùkpò	21		ò dâkpà	21
	0á	elision			elision
		22 (oblig.)			22 (blocked)
		15a (oblig.)			15a (?)
	ò 0á kèlùkpò	15d (oblig.)			15d (?)

In the present (23a, b), the post-elision rules are all obligatory, and the rising tone created by the application of (21) is correctly simplified to low when a high tone follows and to high otherwise. The same is true with respect to the future (23c-e), except that (15a) is optional when the verb stem is monosyllabic (but obligatory when it is bisyllabic). I am not at all happy about this consequence of my analysis, but it is at least an improvement over the previous analyses in that it does in fact generate both variant forms, unlike Leben's, and does not require the strange mode of rule application illustrated in (17), as Elimelech's analysis entails. In any event, a parallel kind of situation must apparently be recognized with respect to (15c) in the past tense for all three analyses where this rule must be optional in the case of bisyllabic verbs, but blocked for monosyllabic verbs. If any one of these analyses is correct, then, this kind of formal mechanism must be recognized in phonological theory.

It might be suggested that all of these analyses are in fact incorrect by virtue of being excessively abstract, and that a more surface-oriented approach is required. Such an approach might make use of the rules given below in the troublesome tenses:

- (24) a. $H]_V \rightarrow L$
 b. $H H \dots]_N \rightarrow L H \dots$
 c. $L_1 \#]_N \rightarrow L H_1$

- d. $L_2 H \dots]_N \rightarrow H L H \dots$
 e. $L H \dots]_N \rightarrow H D \dots$

Under this approach, the tone patterns on the verb-object combinations would be generated directly by first converting the tone of the verb stem to low (24a), and then altering the first part of the tone pattern of the noun in accord with the other rules in (24). If elision is formulated so as to delete the tone on the final vowel of the verb stem, roughly as in (25), and applies prior to the rules in (24), derivations such as those in (26) will result:

$$(25) \begin{array}{c} V \\ | \\ T \end{array} \rightarrow \phi \# \text{---} \begin{array}{c} [V \\ \text{Noun} \end{array}$$

Verb

(26) a. $\begin{array}{c} \circ \# \text{de} \# \text{utsade} \# \\ | \# \# \# \# \# \\ L \# H \# H \# H \# L \# \end{array} \xrightarrow{25} \begin{array}{c} \circ \# \text{d} \# \text{utsade} \\ | \# \phi \# H \# H \# L \end{array} \xrightarrow{24a} (\text{inapplicable})$

$\xrightarrow{24b} \begin{array}{c} \circ \# \text{d} \# \text{utsade} \\ | \# \# \# L \# H \# L \end{array} \xrightarrow{\text{WFC}} \begin{array}{c} \circ \# \text{d} \# \text{utsade} \\ | \# \# \# L \# H \# L \end{array}$

b. $\begin{array}{c} \circ \# \text{kele} \# \text{akpa} \# \\ | \# \# \# \# \# \\ L \# H \# L \# \# \end{array} \xrightarrow{25} \begin{array}{c} \circ \# \text{kel} \# \text{akpa} \\ | \# H \# L \# \# \end{array} \xrightarrow{24a} \begin{array}{c} \circ \# \text{kel} \# \text{akpa} \\ | \# L \# \# \# L \end{array}$

$\xrightarrow{24c} \begin{array}{c} \circ \# \text{kel} \# \text{akpa} \\ | \# L \# \# L \# H \end{array} \xrightarrow{\text{WFC}} \begin{array}{c} \circ \# \text{kel} \# \text{akpa} \\ | \# L \# \# L \# H \end{array}$

This analysis generates precisely the same outputs as Leben's analysis; as such, of course, it fails to generate the second variant in the cases in the future and past habitual where there are alternative surface forms. One way of accounting for these forms within this approach is to posit an optional rule which follows those in (24), but precedes (22), and raises noun-initial low tones to high in the future and past habitual--but only for monosyllabic verb stems (it would have to be inapplicable in the case of bisyllabic verb stems). The necessity for making this rule optional (sometimes) makes it almost as suspicious as the corresponding "optional" rule in Leben's analysis, since, although it does not have quite as dramatic an effect as the latter rule, it too is subject to unusual morphological conditions on its applicability. Thus, the only analyses of those discussed which can account for the variant forms in the future and past habitual by means of an optional rule of the type likely to be optional are the first analysis discussed in this subsection and Elimelech's original analysis. More generally, any analysis which has this property must apparently recognize an intermediate stage with a rising tone on noun-initial vowels in such cases.

It is of considerable theoretical interest that all four of the analyses discussed here require a kind of constraint on the mode of application of a rule (i.e., whether it is optional, obligatory, or inappli-

cable) which is, to the best of my knowledge, unparalleled in descriptions of other languages, in that all require reference to the number of syllables in the verb stem in order to determine the mode of application of at least two rules (which, moreover, affect tones associated with a vowel that is part of a noun). I can conceive of no alternative analysis in which such requirements are unnecessary.

4. Implications. In this section, I would like to discuss briefly the implications of the preceding discussion concerning the relationship between diachronic tonal changes and synchronic tonal systems, and some related issues.

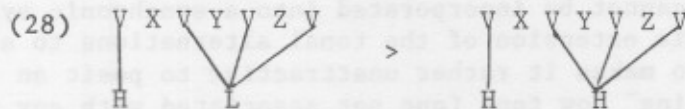
First of all, it should be clear that doing an internal reconstruction of the type performed in section 2.2 can have considerable heuristic value with respect to discovering a possible synchronic rule; rule (21) is in fact the result of a consideration of various stages in the proposed diachronic development. The synchronic account does not mirror exactly the diachronic account, however, and in this particular case it is impossible in principle that it could, due to the diachronic occurrence of an analogical leveling which cannot be incorporated into a synchronic system of phonological rules. This extension of the tonal alternations to all tenses in the negative also makes it rather unattractive to posit an abstract underlying "floating" low tone (one not associated with any vowel) as a marker of present tense: while one could make the synchronic analogue of the diachronic account work in the case of the present tense, similarly positing an abstract floating low would not work in the negative, for the same reason that attempting to attribute the diachronic developments in the negative to the influence of a lost low toned vowel did not (see section 2.2). Thus, in order to include both the present tense and the negative in the same synchronic generalization, it is necessary to refer to morphological categories, and not a low tone, as the trigger for the relevant rules--as was in fact done in all of the synchronic analyses discussed above. Furthermore, since such morphological conditioning is required, there is no good reason for attempting to find phonetically plausible rules--and, again, all of the analyses considered are similar in not requiring phonetic plausibility in the cases at issue. Etsakø thus appears to provide good evidence that even a fairly abstract synchronic phonological analysis can differ quite strikingly from a diachronic account of the same facts, even if the latter is arrived at solely on the basis of data available to the synchronic analyst.

The second issue to be considered here is the apparently different nature of the rules which express diachronic changes and those that account for synchronic alternations. All of the synchronic analyses considered contain rules which change tones associated with two or more different vowels (cf. rules (14a, 24c), for example); the diachronic account of section 2.2 incorporates no such rules. The latter fact is no accident; a deliberate attempt was made to include only phonetically plausible rules (i.e., "diachronically natural rules" in the sense of Hyman and Schuh 1974). The rationale behind this attempt is really nothing more than the traditional assumption that sound changes occur for a good reason--they are a response to some kind of physiological difficulty, either articulatory or perceptual. It is hard to see how changing a sequence of low tones to high, for example, could be interpreted in this way. The changes posited in section 2.2, on the other hand, are straightforwardly interpretable (with the exception of "elision"--see below) as cases of either "horizontal

assimilation" (Hyman and Schuh) or contour tone simplification, both of which have a clear articulatory motivation. Diachronic rules which do not have a physiological motivation are suspicious because in the realm of tone, as in other areas of phonetics/phonology, they explain nothing (though they may provide quite elegant statements about cumulative changes that have occurred from one stage in the history of a language to another).

It might be suggested that, however desirable a priori this assumption about the nature of sound change might be, the evolution of nouns which bear exclusively low tones in the problematic tenses indicates that it must be abandoned. (Recall that a change which, for reasons not explained, converted **all** of the low tones in such cases to high prior to the operation of the changes in (6) was required.) If tone is represented on a separate tier, and a single tone may be associated with more than one tone-bearing unit, then a change of (what appears to be) a sequence of tones is quite easy to state; the simple diachronic rule given in (27) would have the apparently impossible effect illustrated in (28):

(27) $L > H / H _$



Given the input representation in (28), such a change can be interpreted simply as the total assimilation of one tone to a neighboring tone. But recall that there is, at least in the case of Etsakɔ, more to it than this, since grammatical conditioning would have to be added to (27) in order to restrict it appropriately. And even further restrictions would have to be placed on this rule, since only nouns composed exclusively of low tones (but not, e.g., L L H nouns--cf. note 7) were affected by the change in question. Even if such an approach were adopted, then, the rule required would not be a simple one; in view of this fact, abandoning an otherwise well-supported conception of the nature of sound change seems clearly to be insufficiently motivated.

If there is in fact a diachronic-synchronic asymmetry of this sort, then of course an explanation for it is in order. I would like to suggest that the relevant explanation can be found in Leben's (1973) proposal that tones, although represented underlyingly on a separate "suprasegmental" (to use his term) tier, are at some point in a synchronic derivation "mapped onto" tone-bearing units and so become features of them on a par with "ordinary" segmental features. As long as we assume that sound change can affect only post-mapping representations (a very conservative assumption, since there is good evidence that sound change affects only phonetic representations--cf. Jeffers 1977), the asymmetry in question follows naturally.

There are other advantages to accepting this proposal, as well. First, as Leben points out, it implies that all rules that make no mention of associated segmental material will precede those that do; I know of no counterexamples to this prediction. Secondly, making a distinction between pre- and post-mapping rules allows for the possibility that there are other characteristic differences between these two types of rules. It seems that this is in fact the case, in that the former are invariably obligatory (cf. (14), (18), (21), and (24) above), while the latter may be either obliga-

tory or optional. Leben's proposal about mapping thus appears to be quite strongly supported, since there are three universal properties of tone rules which otherwise appear to be quite unrelated, but which can be seen either to follow from this proposal or to be naturally statable in terms of it.¹³

Despite the advantages of this proposal, however, there are also some apparent disadvantages. Probably the most glaring problem is the existence of what Goldsmith calls tonal "stability", which refers to the fact that tones frequently are not lost when the vowel with which they are associated is lost (as in the elisions discussed in sections 2 and 3)--both in synchrony and diachrony. If tones are not features of vowels at the level at which the loss of the vowel occurs, then this is just what would be expected. But the mapping proposal requires that they be features of vowels at the level which is affected by sound change, so another explanation for stability must be found if this proposal is to be maintained. It should be noted that there is further reason to seek, if not a different explanation, at least a deeper one, since the universal concerning the direction of reassociation of (synchronically) "stranded" tones mentioned in section 2 does not appear to follow from rejecting the mapping proposal alone. Moreover, underlyingly "floating" tones, which presumably arise from the diachronic analogue of the synchronic process in question (cf. Hyman 1978, Hyman and Tadadjeu 1976), appear to behave differently than we would expect on the basis of this universal, since they do not always surface (when they do) as being associated with the vowel which historically conditioned the "floating" of these tones (cf. Hyman and Tadadjeu 1976, Clements and Ford 1979).

The basis for the explanation to be argued for here is that segments are **not** deleted (at least in diachronic change and in the case of Stampeian natural processes in synchronic systems); rather, they₄ are assimilated, either partially or totally, to neighboring segments.¹⁴ In the case of the "elision" of vowels discussed above, the full diachronic sequence of events would be as sketched below:

- (29) a. (partial) assimilation of one vowel to another with respect to all oral cavity features;
- b. conversion of the resulting sequence of (nearly) identical vowels to a long vowel (but retaining the tonal features of the formerly distinct vowels);
- c. shortening of the long vowel.

Schematically, if the first vowel bears high tone and the second low tone, the development is that given in (30):

$$(30) \quad \hat{v}_i \hat{v}_j > \hat{v}_j \hat{v}_j \text{ (or } \hat{v}_i \hat{v}_i) > \hat{v}_j: > \hat{v}_j$$

If the assimilation in (29a) is total, and not only partial, in that tonal features are also assimilated, the result is "deletion" of the tone borne by the "deleted" vowel.¹⁵ This phenomenon is also occasionally found, although partial assimilation seems to be more common--just as assimilation with respect to point of articulation only (which is, in essence, exactly what is happening here) is more common than total assimilation of a nasal consonant to a neighboring oral stop.

With respect to synchronic alternations, it is frequently the case that only the input and the output of (29) are readily identifiable and it is

cases of this nature that Goldsmith uses as the basis for arguments which depend on "tonal stability". But the intermediate stages required by this proposal are also frequently to be found. An optional synchronic rule of vowel shortening (i.e., process (29c)) has been proposed for Etsako by Elimelech (1978:33), for example, and cases which show the operation of both (29a) and (29b) are commonplace (cf., for example, Welmers 1973:41-2, Elimelech 1978:22, and many of the papers in George 1972), though often (mis)described as cases of "compensatory lengthening".¹⁷ It is not crucial for present purposes that the output of (29a) exist as a stage independent of the output of (29b), and although I believe that there is a fair amount of evidence that this is in fact the case, I will not pursue the matter here.

Further evidence for this account of the existence of stability comes from phenomena which are not at all expected, given only the formal separation of tones and segments. Consider, for example, the following often cited Lomongo example, whose theoretical interest was first noticed by Lovins (1971a, b):

(31) bǎlǒngó bǎkǎé → bǎlǒngǎkǎé 'his book'

Although I know of no explicit proposals concerning the autosegmental representation of the input form in this example, (32) is the most obvious candidate:

(32) balongo # bakae
 L H # L H

"Deletion" of the second b and the word-final o (and its association line; as Clements and Ford 1979: 195n point out, this "follows from the binary nature of the association relation") would then yield:

(33) balongakae
 L H L H

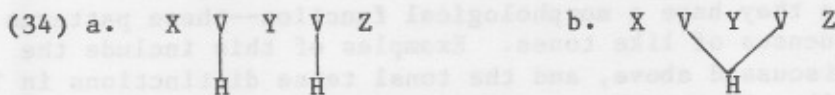
This is incorrect, since the H formerly associated with the deleted o should show up on the following a. Nothing in autosegmental theory as formulated by Goldsmith (1976), however, indicates that it should. Neither does the proposal of Clements and Ford (1979: 207n) that an element "set afloat" by deletion of an associated segment reassociates to the segment that conditioned this deletion, since the H of the lost o has not been set afloat. It is true, of course, that having two different H's associated with the two o's would result in a H which has been set afloat, and therefore appropriately reassociated to the a, but given the lack of any independent motivation for such a representation, suggesting this as an explanation for these facts is clearly ad hoc. Leben (1978: 182), who is apparently the only one who has recognized the problematic nature of such cases, even given Clements and Ford's proposal, suggests that association lines are **not** deleted when one of the elements they associate is lost, and that "these lines are transferred to the segment that occasions the deletion". But this entails the existence of "association" lines that

don't associate anything (cf. Clements and Ford's position cited above). Thus, autosegmental phonology has so far failed to provide an adequate account of some aspects of tonal stability--which is claimed to provide strong support for the theory.

The assimilation proposal, on the other hand--which depends neither on tones being set afloat nor on association lines that associate something with nothing--provides a straightforward account. The δ assimilates (after "loss" of the \underline{b}) to the following \underline{a} with respect to oral cavity features only; the tone necessarily remains because it was not assimilated along with the other features (If it had been assimilated, it would have given the appearance of also having been deleted.) Thus, not only is this proposal plausible in its own right, as evidenced by the independent existence of the stages required to give the appearance of vowel deletion, it also provides an **explanation** for a phenomenon which otherwise cannot even be given a reasonable formal characterization.

It might be questioned whether all synchronic cases of apparent vowel loss should be characterized as the synchronic analogue of the diachronic account just proposed. It is quite likely that not all of them should be; in Etsako, for example, whether the first or second vowel of a sequence is lost depends on morphological information, which at least suggests that the elision rules no longer correspond to the diachronic processes that brought them about. It is not unlikely, that is, that there are quite genuine cases of deletion rules ("rules" in Stampe's sense--i.e., not natural processes) in synchronic phonology. If this is so, however, then we are very much back where started from as far as an explanation of stability phenomena in such cases is concerned, although as long as vowel deletion precedes tonal mapping there is at least a reason why tones are not **always** deleted. Cases such as the Lomongo example discussed above remain problematic, however, unless we can somehow guarantee that representations like (32) are never the input to rules of vowel elision (since if two H's--one for each of the last two vowels in bàlóngó--are present, then the second will be set afloat by elision and reassociated by Clements and Ford's convention).

There is a fairly trivial way of doing so--by requiring that a single tone be associated with only one tone-bearing unit. Such a constraint seems to be presupposed by Elimelech in his treatment of Etsako, and, despite the fact that all other investigators seem to assume that it should not be maintained, I feel that it warrants serious attention. In addition to allowing for a characterization along the lines of Clements and Ford of the full range of synchronic (Stampeian rule) facts it constrains the theory of tonology in an interesting, and apparently appropriate, way. Within Goldsmith's (1976) framework, there is a formal difference between the two representations given in (34):



It has been argued that such a difference is necessary. Cheng and Kisseberth (1981), for example, have proposed a tone rule which depends on the possibility of making such a distinction; Odden (1982), following Goldsmith, proposes that this distinction has a phonetic correlate, with (34a) corresponding to what is traditionally termed a sequence of high and downstepped high tones, and (34b) to a sequence of simple high tones; and

Goldsmith (who refers to downstepped high tones as "mid") makes use of a similar distinction concerning the association of mid tones to "explain" why a mid that follows a mid has a lower pitch than the preceding tone. Analyses which depend on the existence of both such representations are rather rare, however (I know of no others), which at least suggests that this distinction is of questionable validity. In addition, such analyses require something which is very closely related to (and very much in the spirit of) what Kiparsky (1973) calls the "diacritic use of phonological features" in his argument against excessive abstractness in phonology. The abstractness required in analyses of this sort is perhaps sufficient reason for rejecting the tonal distinctions in question, but there is additional counterevidence as well. Neither Goldsmith nor Odden explains or even mentions the fact that no similar distinction with respect to low tones (and, in Goldsmith's case, high tones as well) plays a role in the language being described. If such a distinction is legitimate, then the total lack of any motivation for it in the tones not involved in downstep would be quite unexpected; if it is being used essentially as a diacritic to indicate downstep, on the other hand, this is not at all surprising.

There are, of course, arguments in favor of the position that a single tone can be associated with more than one tone-bearing unit; I will now attempt to counter these arguments. The first (cf. Leben 1973 and the references cited there) has to do with the putative existence of a restricted number of "tonal patterns" in a number of languages which occur in morphemes irrespective of the number of tone-bearing units they contain. Perhaps the paradigm example of a language of this type is Mende, in which patterns such as L H L--but not, e.g., H L H--are said to exist. That is, while there are monosyllabic morphemes with a triple L H L contour, disyllables with a low-falling sequence, and trisyllables with a low-high-low sequences, the corresponding H L H patterns, Leben claims, do not occur. But the latter do occur, at least for morphemes with more than one syllable (cf. Dwyer 1978: 184-5), and recent borrowings seem to show no tendency whatsoever to acquire the "permissible" tonal patterns (Dwyer 1978: 192); the apparent existence of such a restriction is explained by the occurrence of a double accident, one of history (Dwyer 1978: 185-91) and one of data sampled by Leben. Interestingly, languages claimed to have such restrictions (aside perhaps from "pitch accent" languages such as Japanese) always seem to have exceptions to them, even in cases where these patterns provide no support for allowing tones to be associated with more than one syllable (cf. Edmonson and Bendor-Samuel 1966 for an explicit statement concerning the existence of such exceptions in Etung, another often cited language which is claimed to have tone patterns that are associated with a unit larger than the tone-bearing unit).

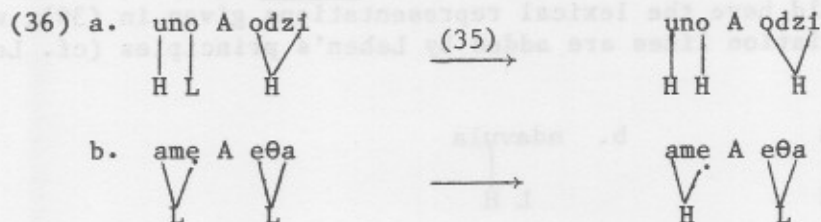
Furthermore, in some cases where it is fairly clear that tonal patterns which do not change depending on the number of tone-bearing units involved are needed--because they have a morphological function--these patterns of necessity have sequences of like tones. Examples of this include the Etsako rule (21) discussed above, and the tonal tense distinctions in Tiv (cf. Goldsmith 1976). In the absence of any genuine (nonaccidental) examples of languages that require tones to be associated with more than one tone-bearing unit, it can safely be concluded that arguments of this sort are without basis.

Leben (1978) has given two further arguments in favor of this position, however. First, in Etsako, there is a rule which raises low tones to high in "associative" constructions. Leben formulates this rule (p. 181) as

follows (where "A" represents the associative morpheme, which is not realized on the surface):

$$(35) L \rightarrow H / \text{ ___ } A$$

If tones can be (or must be, for Leben) associated with more than one vowel, then phonetic sequences of low-toned vowels in other contexts should show up as sequences of high-toned vowels in the associative. This is exactly what occurs, as illustrated in (36).



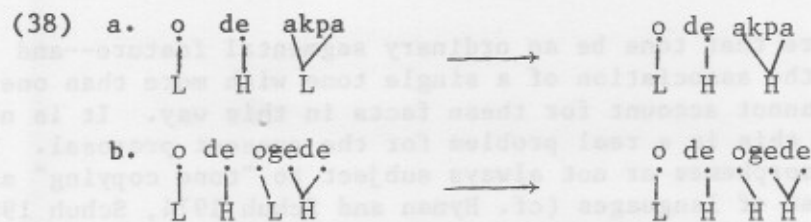
As Leben points out (p. 182), it is possible to account for these facts even if we do not accept this position by rewriting (35) as (35'):

$$(35') L_1 \rightarrow H_1 \text{ ___ } A$$

However, Leben continues, such an analysis "would yield no prediction, on the basis of [facts such as (36a)], that this process would also apply in [cases such as (36b)]" (pp. 182-3). In fact, such a prediction, though correct in this case, should not be made, since it gives incorrect results in other cases--even within Etsakɔ itself. Recall that rule (14a) above affects nouns which contain exclusively low tones. Solely on the basis of such forms, we would have no way of knowing how forms with final high tones would behave; we would therefore presumably choose (as Leben did) the simplest rule that accounts for the facts, something like:

$$(37) L \rightarrow H / H]_V N [\text{ ___ }$$

This would result in the following derivations in the present tense:



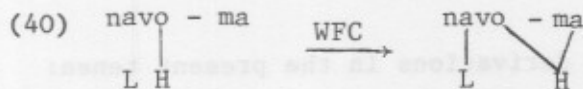
(38b) is incorrect, since the first e should bear low tone (see section 3 for discussion of this form). Thus, the prediction made by Leben's theory, while correct for the cases he discusses, is incorrect for others--even in the same language.

Leben's second argument involves a putative increase in simplicity allowed by maintaining this position in accounting for certain Mende data. Due to the existence of forms such as those mentioned above and other problems pointed out by Dwyer (1978), Leben (1978) has revised his earlier

analysis considerably. These revisions, it is argued, when coupled with a "convention on tone melodies" (p. 201) which deletes the second of a sequence of like tones (optionally separated by a word boundary), allow one to do without two rules which are apparently required by other analyses, as well as accounting for the problematic facts that Dwyer pointed out. The first revision involves allowing high tones to be associated lexically with a vowel in "exceptional words" (i.e. those that do not receive the correct phonetic tone pattern on the basis of Leben's mapping principles and his restricted set of underlying tone patterns, or otherwise behave unexpectedly from the standpoint of his analysis). Thus lèlèmá 'praying mantis' and ndāvúlá 'sling' would have the lexical representations given in (39), where the remaining association lines are added by Leben's principles (cf. Leben 1978:200):



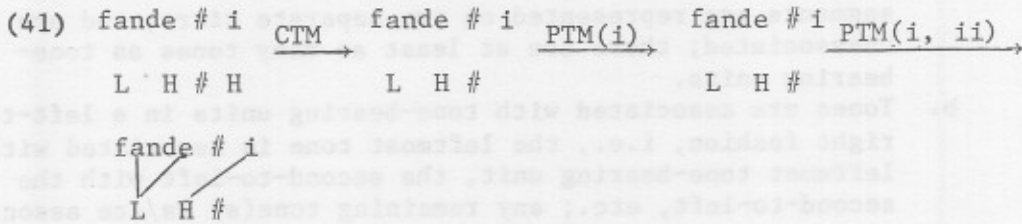
Given such representations, forms like lèlèmá can be said to have an underlying LH melody rather than LLH, which is inexpressible in Leben's system. This revision also allows for a lexical distinction in the case of bisyllabic words with a L H melody between those in which there is no lexical association of tones and those that have a high tone associated with the final syllable,¹⁶ although such a distinction is not required to account for surface isolation tone patterns (only low-high is found). But it is required if we want to do without the rules alluded to above. Consider, for example, constructions involving the postposition -ma 'on', which is apparently underlyingly toneless, and is assigned the immediately preceding tone. If tones may be associated with more than one tone-bearing unit, there is no need for a Mende-specific rule to account for such facts, since the (universal) Well-Formedness Condition predicts just this kind of behavior, as illustrated for nāvó 'money' (which has a lexically associated high tone in Leben's account) in (40):



Theories which require that tone be an ordinary segmental feature--and those that prohibit the association of a single tone with more than one tone-bearing unit--cannot account for these facts in this way. It is not clear, however, that this is a real problem for the present proposal. For one thing, toneless morphemes are not always subject to "tone copying" as in this case; in a number of languages (cf. Hyman and Schuh 1974, Schuh 1978) such morphemes undergo "tone polarization", whereby they receive a tone opposite to a neighboring tone. That is, the universalist account is suspect, since the facts for which it purports to account are not in fact universal. If it should turn out that tone polarization is highly marked as a synchronic phenomenon (it is clearly not a natural diachronic tone rule in Hyman and Schuh's sense), then a universal convention which has the effect of copying a neighboring tone onto a toneless tone-bearing unit can be added to the theory. It might be objected that adding a new universal convention on top of those that make up the Well-Formedness Condition is somewhat ad hoc, but, as will be seen below, since much, if not all, of the

work done by the WFC is not necessary within the present framework, this new convention is not an additional one, but rather a (perhaps partial) replacement for the WFC. Thus, if automatic assignment of tones to toneless segments is desirable, it can be done even without appealing to the WFC.

It is also possible to account for the behavior of bisyllabic L H definite forms that do not have a lexically associated H (unlike the example in (40)) without a Mende-specific rule by making use of the convention on tone melodies (CTM) mentioned above and a revised set of principles of tone mapping (PTM), which are as follows (cf. Leben 1978: 200): (i) a final H is associated with the rightmost syllable; (ii) unassociated tones and segments are paired up from left to right; and (iii) a toneless syllable is associated with the tone of the preceding syllable. Any further association required is done according to the WFC. The operation of these principles is illustrated in (41) for fândé 'cotton', where the high-toned f is the definite article:



That is, Leben's analysis accounts automatically for the fact that the final syllable in fândé-like forms bears low tone when followed by a high tone. Other analyses would presumably require a Mende-specific rule to do this--either one that converts a H to L when it is between a L and a H, or one that simplifies a rising tone to low when followed by high (in analyses that treat fândé-like forms as having, in essence, an underlying rising tone on the final syllable--cf. Dwyer 1978, Singler 1980, Szamosi et al. 1982). While it is true that such a language-specific rule is not necessary in Leben's account, it is also true that both CTM and PTM are themselves language-specific. That is, the cost of eliminating one rule (two in the case of the rising tone analyses, since the rising tone would have to be simplified to high when it is not followed by high) is adding two principles/conventions. Furthermore, Leben's account creates problems with respect to association of boundary symbols on the different tiers that Leben resolves (p. 202) by reassociating them on the basis of an unspecified principle (or set of principles) that require look-ahead global power (in that they refer whether a given reassociation **would** result in a new violation of WFC if it was performed), and, as Leben (1982) himself points out, his earlier account requires that at least one rule crucially precede the operations performed by the WFC--an impossibility in most versions of autosegmental theory. Thus, the Mende facts offer no support for the position that tones may be associated with more than one tone-bearing unit; neither do, to the best of my knowledge, any other natural language facts.

There is one final objection to the approach advocated here. It concerns the existence of contour tones on (phonetically) short vowels, which, Goldsmith (1976) argues, is the basis of a good argument against Leben's mapping proposal. Since there is almost incontrovertible evidence that such tones are composed of two (or more?) level tones (cf., for

example, Woo 1969, Leben 1973, Hyman and Schuh 1974), the existence of mapping entails that a **sequence** of tonal features is part of a single column of the matrix of phonological features. But this is impossible, at least if the mathematical sense of "matrix" is intended. A number of possible solutions to similar problems in the case of nontonal phenomena have been suggested (cf., for example, Krohn 1972, Campbell 1974b, Anderson 1976, and Herbert 1977), and the virtues of the present proposal are sufficiently in evidence that I feel it is worthwhile pursuing these suggestions (or others) with respect to the problem of contour tones, although I will not do so here for reasons of space.

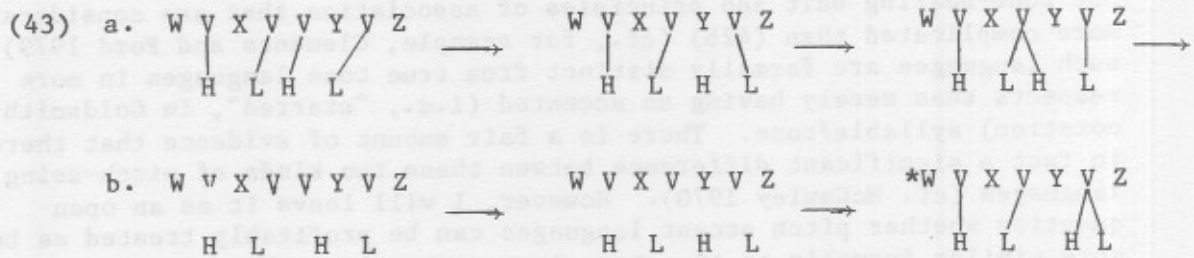
Having completed my discussion of objections to the revisions of autosegmental/suprasegmental theory suggested above, I would now like to make these revisions, which have so far been given only in rather broad outlines, more specific. The phonological component of the grammar of a true tone language is as follows:

- (42) a. In the underlying representation, tones and "traditional" segments are represented on two separate tiers, and are unassociated; there are at least as many tones as tone-bearing units.
- b. Tones are associated with tone-bearing units in a left-to-right fashion, i.e., the leftmost tone is associated with the leftmost tone-bearing unit, the second-to-left with the second-to-left, etc.; any remaining tone(s) is/are associated with the same tone-bearing unit as the tone to its/their left.
- c. Rules that do not require mention of both tiers apply, subject to Clements and Ford's convention on reassociation of stranded tones (e.g., when vowels are deleted).
- d. Tones are mapped onto tone-bearing units, i.e. they become features of segments in exactly the same sense as any other segmental feature.
- e. Rules other than those in (c) apply.

Many aspects of this model have already been discussed, but other reasons for, and implications of, these requirements also deserve some attention. First of all, (42a, b) entail that contour tones in such languages will be found only morpheme-finally, unless they are derived by rule. Whether this is in fact the case is of course an empirical question, and, for the present, an open one. If it should turn out to be false, then the restrictions specified in (42a, b) will have to be relaxed somewhat; allowing for right-to-left association, for example, would permit morpheme-initial contour tones, and allowing some tones to be lexically associated would permit contour tones in any position. Since I know of no clear evidence that such a weakening of the theory is required, I will retain the stronger version as stated.¹⁸

Furthermore, condition (42b) might appear somewhat surprising--or even unnecessary--given the existence of (42d). Why not just **map** the tones in the first place (in left-to-right fashion), and do away with association completely? The answer is that doing so would create severe difficulties with respect to tonal stability when vowels are deleted (not assimilated). Note first of all that such deletions must take place prior to tonal mapping, since otherwise the tones would also be deleted. However, in order for Clements and Ford's convention to work, it is necessary to be

able to distinguish tones that have been "set afloat" from those that have not. Unless tones have been associated with tone-bearing units prior to vowel deletion, it is not clear how this could be done. This convention is necessary, furthermore, in order to insure appropriate reassociation, as in the hypothetical example in (43a), rather than inappropriate mapping, as in (43b):



(The final representation in (43b) would of course actually be in strictly segmental representation, but I will ignore this complication.) It thus appears that, however redundant having both association and mapping might seem, each plays a distinct, and necessary, role in the theory.

It is worth devoting some space to a discussion of the relationship between (42b) and Goldsmith's Well-Formedness Condition (cf. note 2). Given both (42a) and (42b), this condition is unnecessary, since it follows from them: all vowels must be associated with at least one vowel because of the requirements specified in (42a) and the first clause in (42b); all tones are associated with at least one vowel by the joint effect of the two clauses in (42b) (and what is more, with at most one vowel by (42a) and (42b) combined); and association lines may not cross by virtue of the (only) association procedure specified in (42b). That is, with the exception of the questionable automatic association of toneless segments discussed above, all of the work of the WFC is done by independently motivated principles of the theory advocated here--further evidence for this theory.

A further aspect of this proposal may strike some as somewhat suspicious: the second clause of (42a), especially when joined by (42d), makes it appear to be quite similar to a strictly segmental theory, and since representing tone directly as a segmental feature would allow us to do away with all of the conventions in (42), it might be suggested that tone should be represented in just this way. But the theory of "quasi-segmental" tonology advocated here has a number of advantages over a purely segmental theory. First of all, as pointed out above, it allows for a characterization of the differences between pre-mapping and post-mapping tone rules, and requires that the former universally precede the latter. More importantly, it accounts for the existence of tonal stability. As Goldsmith has pointed out, this phenomenon--which shows up both in "ordinary" linguistic phenomena such as vowel deletion and in the more unusual movements of tone-bearing units in language games such as those described in Hombert (1973) and Surinramont (1973)--is extremely problematic for strictly segmental theories. It thus appears that, although (underlying) tones in true tone languages are more closely related to the segments with which they are eventually associated than current non-segmental theories would have it, they are also less closely related to them than is maintained in strictly segmental approaches.

One further point concerning the present proposal merits some discussion. This is the fact that it has been explicitly restricted to "true tone languages" (and, though not explicitly, to African languages). Nothing has been said about the treatment of "pitch accent" languages, which are formally quite similar to true tone languages in most current non-segmental theories. If these accounts of pitch accent are essentially correct, then, because they require the association of tones with more than one tone-bearing unit and principles of association that are considerably more complicated than (42b) (cf., for example, Clements and Ford 1979), such languages are formally distinct from true tone languages in more respects than merely having an accented (i.e., "starred", in Goldsmith's notation) syllable/tone. There is a fair amount of evidence that there is in fact a significant difference between these two kinds of pitch-using languages (cf. McCawley 1970). However, I will leave it as an open question whether pitch accent languages can be profitably treated as being more similar formally to true tone languages, since the existence of languages which appear to be intermediate types (cf. McCawley 1970, 1978; Voorhoeve 1973), at least suggests that the formal differences in question should not be as great as they now are. I will also not attempt to answer the question whether the distinction made by Pike (1948) between languages with a "register tone" system (typically, African and Amerindian languages) and those with a "contour tone" system (Asian languages) is a linguistically significant one, although the language game evidence at least suggests that it is (cf. Hombert 1976 and Churma 1979, ch. 5; but cf. also Yip (1980, 1982) for arguments that it is not). But the situation with respect to true tone languages seems much clearer. Here, tonal representations and rules must conform to the conditions specified in (42), which taken together constitute a theory of tonology that is significantly more restrictive than most other current theories. Moreover, previously proposed analyses allowed for by such less restrictive theories, but not by that advocated here, appear to be undesirable on independent grounds.

Footnotes

*An earlier version of sections 1 and 2 was presented at the Thirteenth Conference on African Linguistics under the title "On the diachronic development of tone in Etsako". I would like to thank Larry Hyman for helpful comments on that paper; Brian Joseph, David Stampe and Greg Stump also deserve a vote of thanks for discussion of the issues raised and/or comments on an earlier version of the full paper.

¹The dialect described is that spoken in Ekpheli; there are twelve other dialects (cf. Elimelech 1978:2). Etsako is a Niger-Congo language of the Kwa subgroup (more narrowly, Edo), and is spoken in Nigeria. (It is maintained by Elugbe (1980) that the proper name for the language is not Etsako--which is the name of the division in which the language is spoken--but Yekhee. Since my data come exclusively from Elimelech, I will use his term, but this should not be interpreted as an endorsement of it.)

I will adopt the following conventions for marking tones. High tone is indicated by an acute accent, low tone by a grave accent, falling tone by a circumflex, and rising tone by a háček; "downstepped" high tone--a tone which behaves exactly like an "ordinary" high tone, except that it has a slightly lower pitch than a preceding high (or downstepped high) tone

(following high tones bear the same pitch as this downstepped high tone, and following lows show the same drop in pitch that they would show if they followed "ordinary" highs)--is indicated by preposing a tick before a syllable whose vowel has an acute accent. Dotted vowels are lax, and g is a voiced velar fricative.

²There is a general problem with the Etsako data since, as Clements (1979) has pointed out, Elimelech's synchronic grammar "overgenerates". In particular, several rules described as being optional must apparently apply (or fail to apply) obligatorily for some forms, despite the fact that optional application correctly accounts for variant pronunciation in other cases. Since failure to record variant forms would not be an unexpected occurrence in a fieldwork situation, it is not clear whether this over-generation is real or merely apparent, but it will have to be given some attention. Because of this problem, one cannot be certain about what those forms not given by Elimelech actually are, so I have been forced to make inferences based on the output of his rules, together with the behavior of other nouns in the same tense. Aside from the problem of optionality, I am reasonably certain about the correctness of the forms given.

³Elimelech (1974:63, 70) gives ò dá'tásà, which is a further indication that the overgeneration problem may be more apparent than real, since this form can be obtained by applying one of Elimelech's "optional" rules to the form given in his later work.

⁴Elimelech (1978:93-102) treats these morphemes as having underlying falling tones (or high plus a "floating" low tone) in the affirmative. While such an account has some synchronic support in the case of the future and the past habitual, the only reason for maintaining it for the past perfect appears to be preserving the "generalization" that negatives are characterized by "a complete tone reversal of any aspectual morphemes..." (i.e., falling in the affirmative, rising in the negative) in their underlying form. Since there are only three tenses involved, proposing an abstract analysis solely in order to bring one tense into line with the (two) remaining tenses is highly suspect, and I will not attempt a diachronic explanation of this putative generalization.

⁵Elimelech gives no explicit discussion of the development of the development of the customary (although he does give a schematic reconstruction (p. 70)), but it is clear that an account like that in (3) is intended. My version of the development of the past is somewhat simplified, in that Elimelech (pp. 65-6) treats the two vowel deletions as occurring at separate stages; nothing here depends on whether this is in fact the case.

⁶As noted above, Elimelech (1978) refers to what he earlier called simply "present" as "present progressive". It appears that his earlier description was more accurate, since Etsako has no "past progressive", although it does have both a present and a past habitual. That is, there is a separable habitual morpheme (at least semantically--it does not seem possible to give a unique phonological form for this morpheme), but there is no evidence for a morpheme which can be interpreted as progressive aspect. There are exceptions to the generalization mentioned, including

present day Etsako, so the violation of this statistical universal can be only suggestive and not conclusive.

⁷ Elimelech does not state, either formally or informally, exactly which structures the rule of high tone spread is meant to affect, although it is clear that it must affect only nouns composed exclusively of low tones when they follow a (high toned) verb in the present tense converting all of the low tones to high. (No noun which contains a high tone is affected--see Table 1.) Elimelech (p. 73) suggests that this is evidence that low tone is a feature of the word (rather than each tone-bearing unit) in the case of low-low-(low) words, since "high spreads over one low tone only". The evidence for the latter claim is the behavior of words like àtásà, which shows up later as á'tásà in the present, as compared with something like ògèdè, which appears as ògèdè. But the high tone on the initial vowel in these forms has nothing to do with "high tone spread", as evidenced by the appearance of initial high tone (sometimes downstepped) in **all** tenses, and not just in the present as one would expect if high spread operated to prevent the present from merging with the customary. (The high tone in these latter cases is pretty clearly the effect of the loss of the vowel of the high toned verb stem.) Thus, the "sound change" in question must make reference not only to specific tenses (or perhaps a preceding tautosyllabic low tone--although this has nothing to do with preventing merger) but also to whether or not a word contains exclusively low tones (or, putatively, has low tone as a word-level feature).

⁸ The universal in question is claimed in these works to be valid only synchronically. Clements and Ford in fact argue (cf. also Hyman and Tadadjeu 1976) that for synchronic "floating" tones, which presumably have the same diachronic source as other stranded tones, the direction of association is unpredictable. While this is undoubtedly true, I would maintain that this lack of predictability would not carry over to a fully diachronic account, since the apparent counterexamples to the diachronic analogue of this universal should be amenable to explanation as being the result of tonal changes (e.g., spreading, in the sense of Hyman and Schuh) which operated prior to the loss of the tone-bearing element, as in the account presented in the following subsection. For discussion of a related concern, see section 4.

⁹ It might be supposed that the necessity of positing (8) indicates that (6d) is rather suspicious. However, there is some fairly strong evidence that the latter did in fact occur. If it did, then for word-internal tonal sequences *L H H ..., we should get *L \widehat{LH} H... While the latter do not occur, neither, with the exception of two forms out of the 542 listed in Elimelech's appendix (and one cited in the text but not listed there), do the former. This near gap could be explained if (6d) had occurred, and was followed by another change simplifying \widehat{LH} to L before H (cf. (6b)). Furthermore, the two forms in question are themselves suspicious in that one (ògbédè 'morning') does not appear to be cognate with five out of the seven other dialects of Etsako represented in Elimelech's appendix, and that the other (òkáfíkái 'gin') is clearly reduplicative in nature and is the kind of lexical item that could quite well be borrowed (note that no forms are given for two of the dialects for this item). While no comparative information is available for òtsédè 'sunrise', it is not unlikely that it too is a relatively late borrowing.

¹⁰It is also possible to offer an alternative account of the origin of the initial low tone on the verb-noun sequence in the case of the present tense, namely that the present has no overt marker, and that the low tone thus is spread from the pronoun onto the verb. (This would, of course, require a boundary adjustment in the reconstructed forms in (7).) Although this possibility cannot be entirely ruled out, it seems rather unlikely. If this were so, then the present and the past would be expected to merge, since (6a) would convert the past tense morpheme to v, which would then be simplified to ṽ by (6b), thus not allowing derivation of the rising tone on the pronoun. As far as I can see, the only way to prevent this merger is to claim that low-spreading is blocked in the past tense in order to prevent merger with the habitual. This would require in addition that the vowel qualities of the past and the habitual markers be identical.

¹¹Since high tones are automatically downstepped (i.e., "downdrifted") after a low tone, it is not really necessary to specify the following high as being downstepped, as long as (15e) does not precede the rule of downdrift, as seems to be universally the case (cf. Churma 1982). Because the latter rule would add irrelevant details to the derivations, I indicate its application only in cases of non-automatic (contrastive) downstep.

¹²This condition contains the following provisions (Goldsmith 1976:27): (i) all vowels are associated with at least one tone; (ii) all tones are associated with at least one vowel; and (iii) association lines do not cross. Together with the apparently universal convention (it is at least unmarked; Clements and Ford 1979, argue rather unconvincingly that it must be violated in Kikuyu) that leftmost tones are associated with leftmost tone-bearing units, the WFC will correctly specify the tone patterns given as the output of (21) in the derivations in (22) below.

¹³No explanation has been given for the invariably obligatory nature of pre-mapping rules, but it seems that the mapping proposal will undoubtedly play a large part in such an explanation: given that sound change affects post-mapping representations, and assuming further that only possible sound changes may have synchronic analogues that are optional, it follows that only post-mapping rules may be optional, and thus that pre-mapping rules are necessarily obligatory. Pre-mapping rules, on this view, are the result of "telescoping" two or more diachronic changes (cf. also Hyman and Schuh 1974).

¹⁴This account draws heavily on the work of David Stampe and his students (cf. especially Semiloff-Zelasko 1973, although her account is of course concerned with a somewhat different phenomenon). Dwyer (ms) has, apparently independently, come to quite similar conclusions. I believe that it is possible to make a stronger claim than that made by these authors, namely, that **all** (or almost all--cf. note 16) apparent deletions (in diachrony and in "natural" synchrony) are due to assimilation, although supporting such a claim is beyond the scope of this paper. Kay Williamson, in her first paper in George (1972), gives an analysis of a case of tonal stability that presupposes an approach to this phenomenon which is essentially identical to that advocated here, but gives no arguments to support it; in her later paper in the same volume, she makes explicit

theoretical claims which preclude her earlier analysis, but are falsified by the data given there (as well as other cases of stability).

¹⁵ Oral cavity features need not be totally assimilated, either. Of special interest here are cases in which the first vowel is assimilated to the second with respect to some oral cavity features, and the latter is assimilated to the former with respect to the remaining features, thus giving the appearance of a third (long) vowel that is a "coalescence" of the two vowels in question. This phenomenon is quite common; Sanskrit is one well-known example of a language with such a process (e.g., a + i → e;, a + u → o:).

¹⁶ Not all cases of apparent compensatory lengthening are the result of such a set of events. DeChene and Anderson (1979) discuss another common source, and examples such as Kikuyu /mo-aná/ 'child' → [mwa:ná] (cf. Clements and Ford 1981:318), which presumably mirrors exactly the diachronic development of this form, indicate that there is (at least) another. In the latter case, the explanation appears to be that the moraic structure of words is maintained even when there is a devocalization, reflected in this case by the retention of the original two morae in the lengthened a:. Length, unlike tone, is truly autosegmental (in that it is never mapped onto segments). Such examples show that the claim made by deChene and Anderson that apparent compensatory lengthening is always "composite" cannot be maintained, although they are correct that many cases described otherwise are in fact composite (as in those discussed here).

¹⁷ As Leben notes, this account entails the existence of an "accidental gap", in that there are no bisyllabic LH words with the H associated lexically with the first (rather than the second) syllable. Since there are also no trisyllabic words which require a lexical first-syllable H, one might well question the accidental nature of this gap. The fact that all trisyllabic words have a lexically associated H is further evidence that Leben's account is considerably less than optimal.

¹⁸ Clements and Ford's (1979) analysis of Kikuyu, which employs a rule that associates the leftmost tone with the tone-bearing segment that is second from the left, would be, if accepted, a counterexample to even the weakest version of these conditions. There is considerable evidence, however, that it should not be accepted. For one thing, it is exceedingly abstract, in the sense of Kiparsky (1973), and there are other problems with Clements and Ford's general framework which I have discussed elsewhere (cf. Churma 1982).

References

- Anderson, Stephen R. (1976). Nasal consonants and the internal structure of segments. *Language* 52:326-44.
- Campbell, Lyle. (1974a). On conditions on sound change. In Anderson, John M. and Jones, Charles, eds. *Historical Linguistics II*, pp. 89-97. Amsterdam: North Holland.
- Campbell, Lyle. (1974b). Phonological features: problems and proposals. *Language* 50:52-65.
- Cheng, Chin-Chuan, and Charles W. Kisseberth. (1981). Ikorovere Makua tonology (part 3). *Studies in the Linguistic Sciences* 11.1:181-202.
- Churma, Donald G. (1979). Arguments from External Evidence in Phonology. Ohio State University dissertation.
- Churma, Donald G. (1982). Cross-linguistic evidence for universal principles of rule interaction: downstep. Paper presented at the Winter Meeting of the Linguistic Society of America.
- Clements, George N. (1979). Review article of Elimelech (1978). *Journal of African Languages and Linguistics* 1:95-108.
- Clements, George N., and Kevin C. Ford. (1979). Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10:179-210.
- Clements, George N., and Kevin C. Ford. (1981). On the phonological status of downstep in Kikuyu. In Goyvaerts, Didier L., ed., *Phonology in the 1980's*, pp. 309-357. Ghent: E. Story-Scientia.
- deChene, Brent, and Stephen R. Anderson. (1979). Compensatory lengthening. *Language* 55:505-535.
- Donegan, Patricia and David L. Stampe. (1979). The study of natural phonology. In Dinnsen, D., ed. *Current Approaches to Phonological Theory*, pp. 126-73. Bloomington, IN: Indiana University Press.
- Dwyer, David J. (1978). What kind of tone language is Mende? *Studies in African Linguistics* 9:167-208.
- Dwyer, David J. (ms.). Deletions and pseudo-deletions. Michigan State University.
- Edmondson, T., and J. T. Bendor-Samuel. (1966). Tone patterns of Etung. *Journal of African Languages* 5:1-6.
- Elimelech, Baruch. (1974). Tone alternations in the Etsako verb. *UCLA Working Papers in Phonetics* 27.
- Elimelech, Baruch. (1978). A tonal grammar of Etsako. Berkeley: University of California Publications in Linguistics, Vol. 87.
- Elugbe, Ben O. (1980). Reconstructing the lenis feature in Proto-Edoid. *Journal of African Languages and Linguistics* 2:39-67.
- Fromkin, Victoria A., ed. (1978). *Tone: a linguistic survey*. New York: Academic Press.
- George, Isaac, ed. (1972). *University of Ibadan Research Notes* 5 (2,3).
- Goldsmith, John. (1976). *Autosegmental Phonology*. Bloomington: IN: Indiana University Linguistics Club.
- Herbert, Robert K. (1977). *Language Universals, Markedness Theory, and Natural Phonetic Processes: The Interactions of Nasal and Oral Consonants*. Ohio State University dissertation.
- Hombert, J.-M. (1973). Speaking backwards in Bakwiri. *Studies in African Linguistics* 4:227-236.
- Hombert, J.-M. (1976). Word games: some implications for analysis of tone and other phonological processes. *UCLA Working Papers in Phonetics* 33:67-80.

References

- Anderson, Stephen R. (1976). Nasal consonants and the internal structure of segments. *Language* 52:326-44.
- Campbell, Lyle. (1974a). On conditions on sound change. In Anderson, John M. and Jones, Charles, eds. *Historical Linguistics II*, pp. 89-97. Amsterdam: North Holland.
- Campbell, Lyle. (1974b). Phonological features: problems and proposals. *Language* 50:52-65.
- Cheng, Chin-Chuan, and Charles W. Kisseberth. (1981). Ikorovere Makua tonology (part 3). *Studies in the Linguistic Sciences* 11.1:181-202.
- Churma, Donald G. (1979). Arguments from External Evidence in Phonology. Ohio State University dissertation.
- Churma, Donald G. (1982). Cross-linguistic evidence for universal principles of rule interaction: downstep. Paper presented at the Winter Meeting of the Linguistic Society of America.
- Clements, George N. (1979). Review article of Elimelech (1978). *Journal of African Languages and Linguistics* 1:95-108.
- Clements, George N., and Kevin C. Ford. (1979). Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10:179-210.
- Clements, George N., and Kevin C. Ford. (1981). On the phonological status of downstep in Kikuyu. In Goyvaerts, Didier L., ed., *Phonology in the 1980's*, pp. 309-357. Ghent: E. Story-Scientia.
- deChene, Brent, and Stephen R. Anderson. (1979). Compensatory lengthening. *Language* 55:505-535.
- Donegan, Patricia and David L. Stampe. (1979). The study of natural phonology. In Dinnsen, D., ed. *Current Approaches to Phonological Theory*, pp. 126-73. Bloomington, IN: Indiana University Press.
- Dwyer, David J. (1978). What kind of tone language is Mende? *Studies in African Linguistics* 9:167-208.
- Dwyer, David J. (ms.). Deletions and pseudo-deletions. Michigan State University.
- Edmondson, T., and J. T. Bendor-Samuel. (1966). Tone patterns of Etung. *Journal of African Languages* 5:1-6.
- Elimelech, Baruch. (1974). Tone alternations in the Etsakɔ verb. *UCLA Working Papers in Phonetics* 27.
- Elimelech, Baruch. (1978). A tonal grammar of Etaskɔ. Berkeley: University of California Publications in Linguistics, Vol. 87.
- Elugbe, Ben O. (1980). Reconstructing the lenis feature in Proto-Edoid. *Journal of African Languages and Linguistics* 2:39-67.
- Fromkin, Victoria A., ed. (1978). *Tone: a linguistic survey*. New York: Academic Press.
- George, Isaac, ed. (1972). *University of Ibadan Research Notes* 5 (2,3).
- Goldsmith, John. (1976). *Autosegmental Phonology*. Bloomington: IN: Indiana University Linguistics Club.
- Herbert, Robert K. (1977). *Language Universals, Markedness Theory, and Natural Phonetic Processes: The Interactions of Nasal and Oral Consonants*. Ohio State University dissertation.
- Hombert, J.-M. (1973). Speaking backwards in Bakwiri. *Studies in African Linguistics* 4:227-236.
- Hombert, J.-M. (1976). Word games: some implications for analysis of tone and other phonological processes. *UCLA Working Papers in Phonetics* 33:67-80.

- Hyman, Larry M. (1978). Historical tonology. In Fromkin, pp. 257-269.
- Hyman, Larry M. and Russell G. Schuh. (1974). Universals of tone rules: evidence from West Africa. *Linguistic Inquiry* 5:81-115.
- Hyman, Larry M., and Maurice Tadadjeu. (1976). Floating tones in Mbam-Nkam. In Hyman, Larry M., Ed., *Studies in Bantu Tonology*. (Southern California Occasional Papers in Linguistics 3), pp. 57-111.
- Jeffers, Robert J. (1977). A comment on 'theories of language change'. Second Annual Linguistic Metatheory Conference Proceedings, pp. 28-46. East Lansing, MI: Department of Linguistics, Michigan State University.
- Kiparsky, Paul. (1973). Phonological representations. In Fujimura, Osamu, ed., *Three Dimensions of Linguistic Theory*, pp. 1-136. Tokyo: TEC Company, Ltd.
- Krohn, Robert. (1972). Underlying vowels in modern English. *Glossa* 6:203-224.
- Leben, William R. (1973). Suprasegmental phonology. MIT dissertation.
- Leben, William R. (1978). The representation of tone. In Fromkin, pp. 177-219.
- Leben, William R. (1982). Mande tone in lexical phonology. Paper presented at the LSA Winter Meeting.
- Lovins, Julie B. (1971a). Lomongo tonology. Unpublished paper, University of Chicago.
- Lovins, Julie B. (1971b). Melodic conspiracies in Lomongo. *CLS* 7:469-78.
- McCawley, James D. (1970). Some tonal systems that come close to being pitch accent systems but don't quite make it. *CLS* 6:526-532.
- McCawley, James D. (1978). What is a tone language? In Fromkin, pp. 113-131.
- Odden, David. (1982). Tonal phenomena in KiShambaa. *Studies in African Linguistics* 13:177-208.
- Pike, Kenneth L. (1948). *Tone languages*. Ann Arbor, MI: University of Michigan Press.
- Schuh, Russell G. (1978). Tone rules. In Fromkin, pp. 221-256.
- Semiloff-Zelasko, Holly. (1973). Syncope and pseudo-syncope. *CLS* 9:603-614.
- Singler, John V. (1980). The status of lexical associations and the obligatory contour principles in the analysis of tone languages. *BLS* 6:442-56.
- Stampe, David L. (1973). A dissertation on natural phonology. Unpublished University of Chicago dissertation. Distributed by Indiana University Linguistics Club.
- Surinramont, A. (1973). Some aspects of underlying syllable structure in Thai: evidence from Khamphuan--a Thai word game. *Studies in the Linguistic Sciences* 3.1:121-142.
- Szamosi, Michael, et al. (1982). A reanalysis of Mende tone. Paper presented at the Thirteenth Conference on African Linguistics, Université du Québec à Montréal.
- Tiersma, Peter M. (1982). Local and general markedness. *Language* 58:832-49.
- Voorhoeve, Jan. (1973). Safwa as a restricted tone system. *Studies in African Linguistics* 4:1-21.
- Welmers, William E. (1973). *African language structures*. Los Angeles: University of California Press.
- Woo, Nancy. (1969). Prosody and phonology. MIT dissertation. Distributed by Indiana University Linguistics Club.
- Yip, Moira J. W. (1980). The tonal phonology of Chinese. MIT dissertation. Distributed by Indiana University Linguistics Club.
- Yip, Moira. (1982). Against a segmental analysis of Zahao and Thai: a laryngeal tier proposal. *Linguistic Analysis* 9:79-94.

An Argument Against Reconstructing Glottalized Stops in PIE*

G. Michael Green

0. Introduction.

Recently, several linguists, including Bomhard (1979), Gamkrelidze and Ivanov (1973:150f.), and Hopper (1973:141f.), have proposed that the stops of Proto-Indo-European which are traditionally reconstructed as a plain voiced series be replaced by a series of glottalized stops. I present an argument here that such a move creates at least one serious problem and therefore ought not to be made.

1. The proposed analysis.

Those advancing the theory involving glottalic consonants point out that the proposed change to the traditional system solves several problems. First, a language with the following three series of stops (the labiovelar and/or palatal series are irrelevant here) seems to be typologically unnatural (Bomhard 1979:78 and the references cited there):

p	t	k
b	d	g
bh	dh	gh

However, the above system of stops is exactly what the traditional reconstruction has proposed. On the other hand, the proponents of the new theory argue, a system of stops like the following has typological parallels (Bomhard 1979:78):

p	t	k
p'	t'	k'
bh	dh	gh

And the above system is just the type that the newly proposed reconstruction suggests. Thus, the new system is argued to be typologically more probable than the traditionally reconstructed system.

The second argument for the new system is that there is very little indication that Proto-Indo-European possessed what would be reconstructed as a voiced bilabial stop in the traditional system (Bomhard 1979:78 and the references cited there). Those proposing to change the traditional system point out that it is not typologically natural for a language with the stop series of traditional Proto-Indo-European to lack a voiced bilabial stop phoneme, and the traditional reconstruction apparently leads us to claim that PIE had just such a gap. On the other hand, they argue, it is quite natural and often found that languages with glottalized stops lack

the glottalic bilabial (Bomhard 1979:78 and the references cited there). Thus, the new system is argued to be more natural in this regard, because the gap in this system is at p', which now replaces the gap at b in the traditional system.

These typological arguments are not without force, and certainly, *ceteris paribus*, we should prefer a reconstruction that is typologically natural to one that is not. I believe though, that the proposed change of the traditional voiced stops to glottalized stops, while seeming to have the advantage of greater typological probability, at the same time creates other problems which are serious enough to call its correctness into question.

2. The argument.

Bomhard (1979:68) has correctly pointed out that a proposal to revise the traditionally reconstructed system "must not only be typologically acceptable but also historically probable [*italics mine/GMG*], that is to say, it must be able to account for developments in the daughter languages at least as well as, if not better than, the old system." I propose that the new system must not only be historically probable with reference to reflexes in the daughter languages, but also with reference to facts about the proto-language itself which we can recover through internal reconstruction. To the extent that internal reconstruction can be shown to be a valid method in historical linguistics, by yielding results that are independently supportable, evidence uncovered by this method must be recognized as admissible in argumentation about comparative reconstruction. I wish to submit just such evidence bearing on the question of whether the PIE series of stops under discussion was voiced or glottalized.

There is substantial evidence, based on surface alternations in the proto-language, that PIE possessed a regressive assimilation rule for sequences of two stops. In the traditional system, this rule is a regressive voicing assimilation rule. From forms in the daughter languages, we can reconstruct PIE surface alternations like:¹

*werg- ~ *w_grk-to-

*leg- ~ *lek-to-

Gothic provides evidence for the alternating surface forms of PIE *werg-. The Gothic form waurkjan indicates PIE *werg-, with g, because k is the regular reflex of PIE *g in Gothic, but -waurths points to a final k in the root, thus *w_grk-to-, because h is the regular reflex of PIE *k in Gothic.² Thus, we must reconstruct two surface forms of this root in PIE, with an alternation in the two forms between g and k. For the alternating surface forms of *leg-, we may note that the reflexes in all of the daughters that retain this root point to a final *g in it, but no daughter shows any evidence for PIE *leg-to-; rather, all of the evidence points to *lek-to-. Therefore, we can firmly establish that PIE had surface alternations between voiced (in the traditional system) and voiceless stops, due to an assimilation of voiced stops to following voiceless stops. There are also cases where a voiceless stop assimilates to a following voiced stop:

full grade *ped- ~ zero grade *bd-

Such alternating forms can be reconstructed on the basis of reflexes like Avestan fra-bd-a 'fore part of foot,' Sanskrit upa-bd-a 'act of trampling, stepping on something,' and perhaps Greek ἐπιβδα 'day after the holiday,' belongs here also. -bd- in each case can be argued to be the reflex of PIE *ped- in zero grade, with *p having assimilated in voicing to *d in the proto-language. If PIE did have surface alternations like those above, and if PIE had voiced stops, not glottalized stops, then we can use internal reconstruction to argue that PIE also had a rule of regressive voicing assimilation for sequences of two stops. If, however, PIE had, not voiced stops, but glottalized stops, the above types of alternations lead us to claim that PIE had a regressive glottalization assimilation rule; that is, now the reconstructed alternations are the following:

*werk'- ~ *w_ork-to-

*lek'- ~ *lek-to-

*pet'- ~ *p't'-

A rule that assimilates voicing in two stop consonants is not an unnatural one, and in fact, is so natural, that we would not be surprised to find such a rule in any language. A glottalization assimilation rule, on the other hand, does not seem to be a particularly natural rule, and in fact, I would claim that it is an extremely unnatural kind of rule, so much so that I have been unable to find an example of any language that has such a rule. This lack of examples of a glottalization assimilation rule is really what we expect when we consider what such a rule would actually entail phonetically. A voicing assimilation rule and a glottalization assimilation rule would be formally quite similar, and we might formalize the two rules in the following way (the two subrules that each rule combines are given below the rule.

- (1) Voicing assimilation [-continuant] → [α voice] / ___ $\left[\begin{array}{l} \text{-continuant} \\ \alpha \text{ voice} \end{array} \right]$
- (1) a. First subrule of 1 [-continuant] → [-voice] / ___ $\left[\begin{array}{l} \text{-continuant} \\ \text{-voice} \end{array} \right]$
- b. Second subrule of 1 [-continuant] → [+voice] / ___ $\left[\begin{array}{l} \text{-continuant} \\ \text{+voice} \end{array} \right]$
- (2) Glottalization assimilation [-continuant] → [α glottalic] / ___ $\left[\begin{array}{l} \text{-continuant} \\ \alpha \text{ glottalic} \end{array} \right]$
- (2) a. First subrule of 2 [-continuant] → [-glottalic] / ___ $\left[\begin{array}{l} \text{-continuant} \\ \text{-glottalic} \end{array} \right]$
- b. Second subrule of 2 [-continuant] → [+glottalic] / ___ $\left[\begin{array}{l} \text{-continuant} \\ \text{+glottalic} \end{array} \right]$

The formal similarity between rules 1 and 2 obscures the significant phonetic differences between them. Voicing is a feature that we would expect to assimilate across clusters, because such an assimilation would eliminate the need to readjust the glottis during the articulation of the stop cluster, and thus accomplishes a genuine simplification of articulation. Therefore, subrules (1a) and (1b) both accomplish exactly the same kind of result, and are both natural assimilations for precisely the same reason, and thus can be naturally combined as a single rule. On the other hand, I

would claim that rules (2a) and (2b) are quite different phonetically, even though both involve a regressive assimilation of glottalization. Rule (2a) does accomplish a simplification of articulation, because it eliminates glottalization, a complex articulatory feature, from the articulation of a glottalized stop when it is in a cluster with another stop. However, rule (2b) would achieve a very different kind of effect. This rule introduces clusters of glottalized stops, which actually increases articulatory complexity. The complication is due to the fact that articulation of a cluster of glottalized consonants would require the repetition of the glottalization process twice⁴ within a very short period of time, and glottalization is a relatively complex feature even when involved in the articulation of a single consonant. In addition, Jeffers and Lehiste (1979:6) state that glottalized consonants, because of articulatory complexities associated with them, are among the segments especially likely to be involved in dissimilations.⁵ This fact makes the assimilation of a plain voiceless stop to a glottalized stop, as in (2b), even less likely. It is quite clear then, that the newly proposed reconstruction leads us to claim that PIE had a rule like (2) above, which presents the problem of being unnatural in two respects. First, the rule combines two subrules which achieve very different kinds of results.⁶ Second, one of the subrules is quite implausible phonetically. The traditional system, on the other hand, can account for exactly the same facts that cause serious problems for the newly suggested reconstruction, and can do so with a single rule which is very plausible phonetically and which combines two subrules which achieve exactly the same result and which can thus be argued to genuinely be instances of the same rule. Thus, the internal reconstruction of an assimilation rule for sequences of two stops in Proto-Indo-European provides at least one strong piece of evidence against accepting the replacement of the traditional PIE voiced stops with glottalized stops.

3. Conclusion.

As outlined in the beginning of this paper, there are typological arguments which can be made for the proposal that what have traditionally been reconstructed as voiced stops in Proto-Indo-European should actually be reconstructed as glottalized stops. As I have argued above, there is good reason to believe that such a proposal is actually incorrect. We must ask then, since arguments exist both for and against the proposal, how we are to decide which arguments are to be given the most weight. Actually, I believe that there are other reasonable arguments which could be advanced against this proposal, though I will not go into these here. I suspect that conflicts between typological arguments and other types of arguments may eventually cause a reassessment of the relative weight that typological evidence should be given in deciding questions about reconstruction. I cannot provide sufficient argumentation at this time to make this suspicion any more than a suspicion, but I would nevertheless maintain that the main argument that I have given in this paper must be reckoned with in deciding whether to reconstruct voiced or glottalized stops in Proto-Indo-European.

Footnotes

*I wish to thank Brian Joseph for much helpful discussion of the ideas presented in this paper. I would also like to thank Rob Fox for his comments concerning the phonetic issues involved.

¹I am indebted to Brian Joseph for these and the following examples of alternating forms in PIE.

²Though it is clear that we should consider the Gothic verb waurkjan to continue PIE *werg-, because neither the semantics nor the required sound changes are problematic, Gothic waurhts is not actually attested in simple form, but only in compound verb forms, for example, fra-waurhts and us-waurhts (Feist 1923:422). Even these compound forms are clear evidence for the alternation in PIE, however, because the forms fra-waurkjan and us-waurhts also occur (Feist 1923:422), and in these compound forms, there is no question that -waurkjan and -waurhts derive from the same PIE root. Thus, though Gothic waurhts does not actually occur in simple form, the evidence for the alternation in the proto-language is still secure.

³I have treated glottalization as being expressed by a single feature for the purposes of these rules. Whether such a treatment is actually correct or not makes no difference for the argument that I am giving.

⁴Some kind of repetition of the articulatory movements involved in glottalization would be required whether the proposed glottalic stops were ejectives, as Gamkrelidze and Ivanov (1973:150f.) and Hopper (1973:141f.) propose (a proposal with which Bomhard (1979:68) apparently agrees), or were voiceless stops followed by a glottal stop, or articulated with a simultaneous glottal stop, or even if they were some other type of glottalized stop (if other types actually exist). The point here is that having glottalization in both stops in a two stop cluster does not simply involve holding some articulatory factor constant throughout the articulation of the cluster, which is all that is involved in having the same voicing value for both stops in a cluster. Assimilation of voicing eliminates the need for an articulatory readjustment in the middle of a cluster; assimilation of glottalization, when it produces two consecutive glottalic consonants, creates the need for an extra readjustment. It is this fact that makes rule (2b) so phonetically implausible.

⁵Actually, Jeffers and Lehiste make this remark in talking about sound change, not about synchronic rules. However, there is every reason to believe that sound change should reflect natural synchronic rules. Therefore, it seems reasonable to conclude that if glottalized consonants are often prone to dissimilation in sound change, then they should exhibit the same kind of behavior in synchronic rules.

⁶In order to resolve this particular difficulty, it might be said that the two rules (2a) and (2b) should not be combined, but should be independent. This move does not really solve the problem though, for then the claim would be that PIE had two independent rules which achieve virtually opposite kinds of results, which is really no better (or only trivially better) than saying that the two were subrules of the same rule. Even if this problem could be resolved, the fact that rule (2b) is so implausible phonetically seems to be irresolvable.

References

- Bomhard, Allan R. (1979). The Indo-European Phonological System: New Thoughts about its Reconstruction and Development. Orbis 28.66-110.
- Feist, Sigmund. (1923). Etymologisches Wörterbuch der Gotischen Sprache. Halle (Saale): Max Niemeyer.
- Gamkrelidze, T. V. and V. V. Ivanov. (1973). Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse. Phonetica 27.150-156.
- Hopper, Paul J. (1973). Glottalized and Murmured Occlusives in Indo-European. Glossa 7.141-166.
- Jeffers, Robert J. and Ilse Lehiste. (1979). Principles and Methods for Historical Linguistics. Cambridge, MA: MIT Press.

How Ad Hoc is Phonology?
Evidence from Tocharian*

George Michael Green

1. Introduction
- 1.0. Background Assumptions.

Evidence exists which can be used to construct arguments against generative phonology as a correct theory of the phonologies of human languages. This evidence, from the language Tocharian B, is presented in section 1, along with the arguments that it provides. The validity of these arguments depends crucially on what one takes to be the goal of a linguistic theory; therefore, the view of this goal which has been assumed is stated explicitly here. Though most phonologists (and in fact, most linguists) probably share this view, justification is nonetheless provided for taking this particular view rather than others that might be taken.

The particular view of the goal of linguistic theory which is assumed here has been stated by, among others, Joseph (1980:345),¹ who states that "Among the primary concerns of theoretical linguists are the problems of characterizing what constitutes a 'possible' and conversely an 'impossible' human language." An assumption which is related to this view, seldom stated explicitly, but which most linguists seem to make ahead of time, is that the class of possible human languages is a proper subset of the class of conceivable languages. Put another way, it is assumed that there are some conceivable languages which are not possible human languages. Thus, the goal of linguistic theory might be viewed as separating the class of languages that are possible human languages from the class of languages that are not. There are reasons for supposing that the class of human languages is considerably restricted, but even if this assumption turns out to be incorrect, it is more fruitful, until the question is definitely decided, to assume that there are restrictions than to assume that there are not. The former assumption is more fruitful because it has empirical consequences, whereas the latter does not, and it can be argued that such consequences make it possible to compare theories of language in a principled manner.

If the class of possible human languages is in fact restricted, a theory which correctly specified this class might involve two different kinds of conditions. On the one hand, the theory might specify some property that every human language must have, thus narrowing the class of possible human languages to only those that have this property. On the other hand, the theory might specify some property that no human languages can have, thus limiting the class of possible human languages to those that do not have this property. Of course, a theory could incorporate only one or the other of these two types of conditions, or a combination of both types. Whichever of these three possibilities is chosen, it can be demonstrated that the theory has empirical consequences;

that is, the theory is subject to support or disconfirmation by the facts of human languages. If the theory claims that all languages will have a certain property, then an example of a language without the property falsifies the claim. Similarly, if the theory claims that some feature will be absent from all human languages, then an example of a language that possesses that property falsifies the claim. Of course, a theory which incorporated both types of restrictions would be subject to empirical disconfirmation of either sort. In contrast to theories that attempt to restrict the class of possible human languages, theories that do not attempt to do so have no empirical consequences, because they make no predictions about the properties of human languages as a class; thus, no matter what is found to be true with reference to the facts of human languages, these facts will have no relevance to an evaluation of the theory.

It can be seen then, that a linguistic theory which restricts the class of possible human languages is testable, whereas one that does not restrict this class is not testable. Clearly, then, it is desirable to assume that the class of possible human languages is in fact restricted, and to construct only theories which make claims about what these restrictions are. Since the correctness of theories which are constructed in this way can be evaluated by referring to the facts of human languages, it becomes possible to compare any two or more such theories. If there are no counterexamples to any of the theories which are being compared, then the theory which places the greatest restrictions on the class of human languages is to be preferred. The most restrictive theory is to be preferred, because it is the theory most likely to be too restrictive; that is, it would be the theory most likely to incorrectly require the inclusion or exclusion of some property for all human languages. Such a theory would be the most readily falsifiable one, for example, by the discovery of some language that does not meet with the predictions of the theory. Therefore, if none of the competing theories has suffered any counterexamples, then the most restrictive theory is the most likely to be correct.

Though many phonologists hold the same basic view of linguistic theory as the one which is stated above, it is still important to be explicit about these matters, because in practice, few phonologists are. Quite often, phonologists have even failed to consider their proposals in light of the requirements that linguistic theories must meet in order to be falsifiable, yet most of these same phonologists would probably accept these requirements as desirable, because they do wish to be able to evaluate phonological theories empirically. As a result of failing to consider fully enough questions of falsifiability, generative phonologists have come to be in the strange position of holding a phonological theory which is inconsistent with their views of linguistic theory as a whole. In the next chapter, the ways in which generative phonology is inconsistent with the goal of linguistic theory is demonstrated. First, though, the data on which this demonstration is based is presented and analyzed.

2. Generative Phonology in Light of the Goal of Linguistic Theory.
2.0. Principles of Generative Phonology.

In order to demonstrate that generative phonology is inconsistent with the goal of linguistic theory, it is first necessary to state the important principles of this phonological theory. There are a number of different versions of the theory of generative phonology, which differ on varying points, but which at the same time are all alike in certain basic ways. The points on which these versions differ do not have the same theoretical status as the points on which they agree. To differentiate the status of the two, a distinction is employed here, which is discussed by Zwicky (1972:151), between methodological principle and theoretical or substantive principle. Zwicky says that "substantive principles are theoretical requirements, methodological principles are theoretical biases." To apply this distinction, it would be said that the substantive principles of generative phonology are those principles which a theory must subsume in order for that theory to be a version of a generative phonological theory. If the theory omits or alters any of the substantive principles of generative phonology, then it is not a version of a generative phonological theory, but rather, it is a version of some different theory of phonology. On the other hand, different versions of generative phonology may vary freely on methodological principles, and still remain instances of the same general theory of phonology.

The single theoretical requirement, or substantive principle, of generative phonology can be stated as follows (together with references to phonologists who have stated this principle in some form):

- (1) The phonological component of a grammar maps input strings from the syntactic component onto systematic phonetic strings. The phonological component accomplishes this mapping by applying phonological rules, which utilize a set of phonological features, boundaries, and other pieces of notation which the theory provides (e.g., parentheses, angled brackets, etc.), and which may mention morphological or morphosyntactic information (cf. Householder 1979:253; Anderson 1979:3; Chomsky and Halle 1968:9-12, 14, 295-298).

In addition to this substantive principle, three methodological principles which are commonly employed by generative analysts can be given as follows:

- (2) Other things being equal, phonological rules which mention only phonological features, boundaries, and notation are to be preferred to rules that mention some nonphonologically defined class of lexical items (cf. Zwicky 1972:156).
- (3) Other things being equal, every morpheme in a language should have a unique underlying shape; that is, *ceteris paribus*, there is one and only one representation of a given morpheme in the lexicon, and only a single shape of any given morpheme is the input to the phonological component. Suppletion or listing of the allomorphs of a given morpheme in the lexicon is to be resorted to only when no acceptable phonological account

is available (cf. Householder 1979:253; Zwicky 1972:156; Kenstowicz and Kisseberth 1979:46-57).

- (4) "Other things being equal, an occurrence of a segment not involved in alternations should be represented underlyingly in its surface form" (stated by Zwicky 1972:158).²

In the next section, evidence is presented which can be used to argue that principle (1) above cannot be maintained. Specifically, it is argued in a later section that a major part of the mapping of lexical representations onto phonetic representations should not be accomplished by "phonological" rules, that is, by rules that mention phonological features and other phonetic information, but rather, by a set of morphological rules. It is further argued that the class of rules which should be allowed to mention phonological features must be very narrowly specified, and that therefore, the class of rules which are actually phonological is much smaller than the class of such rules which is allowed by generative theory.

Arguments can also be presented against the methodological principles given above. In particular, it can be argued that the strongest form of any of the three principles that can be defended allows ad hoc analyses of linguistic data. In general, it is argued that the set of four principles which are given above cannot be interpreted in a way which places any nontrivial restrictions on the class of possible human languages. Before this claim can be defended, the evidence on which it is based must be presented.

2.1. A Generative Analysis of the Tocharian Data.

In this section, a generative phonological analysis of data from the language Tocharian B is presented. Tocharian B (or West Tocharian) and Tocharian A (or East Tocharian) are the two languages which constitute the Tocharian branch of the Indo-European family of languages. The data from Tocharian B is the evidence upon which the arguments in this work are based. Before a generative analysis of this data is undertaken, however, the phonetic inventory of Tocharian B is given below, and several points are made concerning stress.

2.1.1. The Sounds of Tocharian B.

Krause and Thomas (1960:39) give the following inventory of sounds for Tocharian B:³

CONSONANTS

	Stops (all voiceless)		Nasals		Liquids		Sibilants (all voiceless)	Glides
		palat- alized		palat- alized		palat- alized		
Bilabial	p	py	m	my				w
Dental	t		n		l, r	ly	s	
Alveopalatal							ʃ ⁴	
Palatal	c		ɲ				ʃ	y
Velar	k	ky	ŋ					

VOWELS

	Front		Central	Back	
	Long	Short		Long	Short
High	ī	i	ɨ	ū	u
Mid		e	a		o
Low			ɤ		

Tocharian B also has three diphthongs, āu, au, and oy, and nasalized vowels, which are represented by the symbol m following the vowel (i.e. $V\bar{m} = [\bar{V}]$). The interpretation of the consonants which is given here, as well as the interpretation of the front vowels and the back vowels, is standard, and requires no further comment. The interpretation of the central vowels, however, is more difficult to be certain about, and the interpretation given here follows that of Jasanoff (1978:30-31). Jasanoff proposes the following phonetic values for the central vowels:

$\ddot{a} = [ɨ]$

$a = [ʌ]$

$\bar{a} = [a]$

Jasanoff gives two different arguments in favor of this interpretation. First, Jasanoff observes, the vowel \ddot{a} is best interpreted as a phonetically high vowel, because it often fluctuates with the other two high vowels, i and u, in the spelling of certain words. Specifically, \ddot{a} is sometimes spelled as i in the environment of labial and palatal consonants. For example, cāñcare 'lovely'⁵ is sometimes ciñcare (Jasanoff 1978:30. Cf. Krause and Thomas 1960:49 for the further example of śānmalñe 'das Kommen', sometimes spelled as śinmalñe). In addition, \ddot{a} is sometimes spelled as u in the environment of velar or labiovelar consonants (Jasanoff 1978:30). For example, kwālypelle, gerundive of the verb kulyp- 'desire', is sometimes spelled as kulypelle (Cf. Krause and Thomas 1960:50 for the further example of kwārsarwa, plural of 'Vehikel, Meile,' sometimes spelled as kursarwa⁶). Thus, since \ddot{a} seems to be phonetically high, yet clearly distinct from the front vowel i and the back vowel u (since it is usually differentiated from them), it seems correct to interpret it as a high central vowel.

Second, Jasanoff argues, once \ddot{a} has been established as a high central vowel, it seems best to interpret a and \bar{a} as central vowels also, because this interpretation allows a particular alternation in which these three vowels are involved to be viewed in a highly natural way. Specifically, \ddot{a} alternates with a, and a alternates with \bar{a} , the first vowel in each pair appearing unstressed, and the second vowel appearing under stress. For example (Krause and Thomas 1960:43) tārkār 'Wolke,' but plural tārkārwa; pārna 'draussen', but related adjective, pārnaññe 'ausenstehend'; śārsa 'wusste', but 3rd person plural śārsāre; tāka 'wurde', 3rd person plural takāre. Jasanoff argues that if the three vowels involved in these alternations are interpreted as central vowels of three heights, with \ddot{a} the highest and \bar{a} the lowest, then the alternation of \ddot{a} with a and a with \bar{a} can be interpreted as the lowering of \ddot{a} (= [ɨ]) and a (= [ʌ]) under stress, which is a very natural phonological phenomenon. Jasanoff's interpretation is accepted here, though it should be noted that whether this interpretation is actually correct or not, the arguments presented in this thesis are in no way affected.

2.1.2. Stress.

Krause and Thomas (1960:43) give the following rule for stress in Tocharian B:

"Die Hauptregel lautet: In den meisten zweisilbigen Wörtern ruht der Akzent auf der ersten, in drei- (und teilweise vier) silbigen auf der zweiten Silbe."

Krause and Thomas do note that there are exceptions to this rule; therefore, wherever exceptions to this rule occur in the data given below, they are specifically noted. Otherwise, all forms can be assumed to follow this rule.⁷

2.1.3. The Analysis.

Before a generative analysis of the Tocharian data is given, several points should be made. It can be argued that it is desirable to assume the strongest forms of principles 2, 3, and 4 of generative phonology given above (pp. 58-59), until some reason can be found, based on the data, to weaken the strongest forms of these three principles. That is, it would be most desirable to be able to claim that phonological rules never mention any information other than phonological information (the strongest form of principle 2), and that every morpheme in a language (other than, of course, clear cases of suppletion) always has exactly one underlying representation (the strongest form of principle 3), and that any surface segment not involved in alternations is always represented underlyingly in its surface form (the strongest form of principle 4). It is most desirable to be able to claim the strongest form of each of these principles, because the strongest form places the greatest restriction on the class of possible human languages, and thus most effectively pursues the goal of linguistic theory. It might turn out that the strongest form of a given principle is inconsistent in some way with the facts of human languages; however, if that is the case, the principle should be weakened only as much as is necessary to make it consistent with the known facts, in order for the weakened form of the principle to still place the greatest possible restriction on the class of possible human languages. Moreover, it might even turn out that a given principle must be weakened so much that it no longer places any nontrivial restrictions on this class; in such a case, the principle must be abandoned. As the following generative analysis is carried out, then, one of the major questions to be answered is the degree to which principles 2, 3, and 4 must be weakened to make them consistent with the data, if in fact they must be weakened at all. This issue could be seen as equivalent to specifying the exact meaning of the condition other things being equal in each of these three principles as stated in section 2.0; that is, if the strongest forms of these principles do not hold unequivocally, then the exact conditions under which they fail to hold must be specified.

A second important point concerning the analysis should be made. This analysis is not simply one of a number of possible generative analyses of this data; rather, it can be claimed to be the best account that it is possible to provide for this data within a generative framework. Arguments to support such a claim are given in the course of analyzing the data, as well as in the remainder of this chapter.

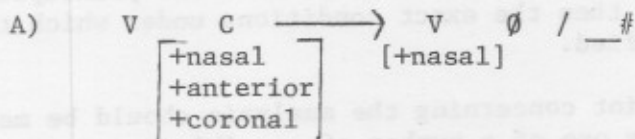
In answer to a possible objection to the methodology that is used, one final point deserves attention. Several verbal paradigms are examined in succession, one at a time, in order to construct the details of the generative account. It might be thought that the order in which the paradigms are presented determines the particular details of the account, and that if they were examined in a different order, a different account would result. This objection is not valid, however, because at each point in the analysis, all of the viable alternative analyses are considered, so that no possibility is omitted simply because of the order in which the data is presented. Therefore, the order in which the data is presented is based entirely on considerations of clarity and ease of presentation, and has no affect whatever on the resulting analysis.

The analysis begins with the active paradigm of the verb pälk-8 'leuchten' (Krause and Thomas 1960:262):

Singular 1	palkau	Plural 1	pälkem ⁹
2	palkät	2	palkcer
3	palkäm	3	palkem

First, it should be noted that the 1st person plural form is stressed on the second syllable; thus, given the alternation between ä and a, which is conditioned by stress (discussed above, p. 6), the stress in this form accounts for the appearance of ä in the first syllable, since it is unstressed, in contrast to the appearance of a in the first syllable of all the other forms, where the vowel is stressed. There is no way to determine at this point which of the two vowels involved in the alternation is underlying, but for the moment ä can arbitrarily be taken as the underlying vowel. Forms which can be used to decide this question definitely appear later in the data.

Second, the source of the nasalized vowels in the 3rd person of both numbers can be determined, and it can be shown that these vowels are derived, not underlying; therefore, a discussion of this matter simplifies the remainder of the analysis of these forms. All nasals but n and ñ occur word finally on the surface in Tocharian B, and ñ occurs on the surface only before velar stops. Thus, given these distributional restrictions, final nasalized vowels can be derived from an underlying sequence of a vowel followed by n; that is,¹⁰



Now the forms must be divided into component morphemes. The first four segments of each form (other than the alternation between ä and a, which has already been accounted for) are invariant, so at least this much can be taken to be part of the underlying verb stem.

It is possible, then, that the underlying verb stem is simply /pälk/. Other viable possibilities are /pälkä/ or /pälke/. If the verb stem is taken to end in ä or e underlyingly however, several facts about the surface forms are difficult to account for. Specifically, ä of the 3rd person singular and e of the 3rd person plural appear in exactly analogous phonetic contexts; thus, there is no straightforward way to take one of these vowels as the final vowel of the underlying verb stem and to derive the other one by phonological rule. Further, if the underlying verb stem is taken to end in a vowel, the fact that no vowel surfaces in the 2nd person plural form would have to be explained. Therefore, tentatively, it seems best to take /pälk/ as the underlying verbal stem, and the following as the underlying representations of the active verbal endings:

Singular 1	/au/	Plural 1	/em/
2	/ät/	2	/cer/
3	/än/	3	/en/

Next, the deponent-passive paradigms of two different verbs, mäsk- 'sich befinden, sein,' and plänt- 'vernügt sein,' may be considered:

Singular 1	mäskemar	Plural 1	mäskemt(t)är
2	mäsketar	2	mäsketar
3	mäsketar	3	mäskentar
Singular 1	plontomar	Plural 1	plontomt(t)är
2	plontotar	2	plontotar
3	plontotar	3	plontontär

In order to have a single underlying representation for the stem of each of these two verbs, along with a single underlying representation for each of the deponent-passive endings, the following representations are required:

STEMS

/mäske/ /plonto/

ENDINGS

Singular 1	/mar/	Plural 1	/mt(t)är/
2	/tar/	2	/tär/
3	/tär/	3	/ntär/

The above solution is the best one, because if the e in the mäsk- forms and the second o in the plänt- forms were analyzed as part of the endings, each and every deponent-passive ending would exhibit two different surface allomorphs, for example, 1st person singular -emar and -omar,

2nd person singular -etar and -otar, etc. Further, the distribution of these allomorphs could not be predicted on phonological grounds, so it would not be possible to have one as underlying and to derive the other by phonological rule. Thus, each of the deponent-passive verbal endings would have two underlying representations, and this solution would thus violate the strong form of principle 3, and therefore ought to be avoided if possible. This solution and its undesirable consequences can be avoided by analyzing the e and the o vowels as part of the verbal stem, which also means tentatively accepting the underlying representations given above as the correct ones for the deponent-passive verbal endings.

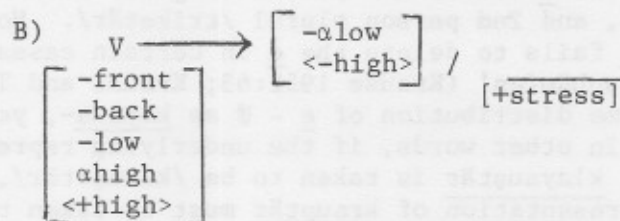
To return to the question of which vowel underlies the alternation between ä and a, the deponent-passive verbal endings above provide evidence that bears on this question. If the strong form of principle 4 (p. 59 above) is adopted, as it is argued above that it should be, then the surface forms of the deponent-passive endings lead to a determination of the underlying vowel in the a ~ ä alternation, as well as in the ā ~ a alternation (both discussed above, p. 61). The vowels a and ä both occur in the surface forms of the deponent-passive endings, for example, in the 2nd person singular -tar and the 3rd person singular -tär. These endings always constitute the last syllable in a word of at least two syllables; therefore, it follows from the main stress rule (given above, p. 62) that the vowels in these endings are never stressed. Thus, the a of the 2nd person singular deponent-passive ending does not alternate with ä, because the vowel is always unstressed in this ending. Similarly, the ä of the 3rd person singular deponent-passive ending is always unstressed, and therefore, this ä does not alternate with a. Therefore, the underlying representation of every a in the deponent-passive endings should be /a/, by the strong form of principle 4, because these segments never alternate. By the same reasoning, every ä in these endings should be represented as underlying /ä/.

Moreover, it can be argued that, if a and ä are taken as the underlying vowels in these cases, where the vowels do not alternate with ā and a respectively, then a and ä must be chosen as the underlying vowels for the alternating cases as well. To demonstrate this, let the alternative solution be adopted; that is, let it be supposed that the vowel underlying the nonalternating a in -tar is /a/, and that the vowel underlying the nonalternating a in -tar is /a/, and that the vowel underlying the nonalternating ä in -tär is /ä/, but that the vowel underlying alternating ä, as in pälk- (p. 63 above) is /a/, and that the vowel underlying a ~ ā is /ā/. Let it also be noted that no ā occurs unstressed on the surface, and that no ä occurs stressed on the surface.

If the alternative solution proposed here is adopted, then the only place that ä would occur in underlying representations in Tocharian B would be in the final syllables of endings, which is an unnaturally limited distribution for any segment. The reason that the distribution would be so limited is that the final syllables of endings are the only syllables where the a ~ ä alternation (and the ā ~ a alternation) never

occurs. Therefore, if this unnaturally limited distribution of ä in underlying representations is to be avoided, then ä must be taken as the underlying vowel both for surface ä that does not alternate, and for surface alternating a ~ ä. In this solution, the occurrence of underlying ä is entirely unrestricted; that is, the vowel may occur underlyingly in any position.

A similar argument can be made for the ā ~ a case. If nonalternating a is derived from underlying /a/, but a alternating with ā is derived from underlying /ā/, then the occurrence of ā in underlying representations is restricted to syllables other than the final syllables of endings, but again, this restriction is a strange one.¹¹ If, on the other hand, /a/ is taken as the underlying representation not only for nonalternating surface a, but also for surface alternating ā ~ a, then the occurrence of /a/ in underlying representations is entirely unrestricted. Therefore, the underlying representation for surface a ~ ä is /ä/, and the underlying representation for surface ā ~ a is /a/, and the rule that governs these two alternations can be stated as follows:



2.1.3.1. The Present Palatalizing Verbs.

Certain verbal paradigms in Tocharian B exhibit several interesting alternations, and necessitate a number of revisions to the previous analysis. As a first example of these verbs, the deponent-passive paradigm of the verb klyaus- 'hören' may be examined:

Singular	1	klyausemar	Plural	1	klyausemt(t)är
	2	klyauṣtar		2	klyauṣtär
	3	klyauṣtär		3	klyausentär

This verb is similar in some respects to the verb mäsk- (p. 64 above), but in two respects it is different. First, the vowel -e- is present in the 1st person singular, and 1st and 3rd persons plural, but is absent in the other persons, whereas in the verb mäsk- the -e- vowel is present in all persons. Second, s appears in the stem in the persons with the -e- vowel, but alternates with ṣ which appears in the forms without a following e. These two alternations, e ~ ∅ and s ~ ṣ, must be accounted for. There are two possibilities for accounting for the e ~ ∅ alternation; namely, either the vowel is present underlyingly and is deleted in some cases, or it is not present underlyingly, and is inserted in some cases. Considering the deletion alternative first, this possibility may be ruled out based on verbs like trik-¹² 'in die Irre gehen,' which shows 2nd person singular triketär, 3rd person singular triketär, and 2nd person plural triketär (Krause 1952:66; cf. Krause and Thomas 1960:200). If klyauṣtar, for example, were underlyingly /klyauṣetar/, with a vowel,

and the vowel were deleted by a phonological rule, then the rule would be expected to delete the -e- in triketar also, but the e in this form is not deleted, but surfaces. There is a way that a rule could be written which would delete the e in the klyaus- forms but not in the corresponding trik- forms; specifically, in the klyaus- forms where e is deleted, it is preceded by a continuant, but in the corresponding trik- forms, the e is not preceded by a continuant. Thus, a rule can be written that deletes e when preceded by a continuant and followed by a non-nasal (in order to keep the e in the other forms of klyaus- from being deleted):

C) $V \rightarrow \emptyset / [+continuant] _ [-nasal]$
 $\left[\begin{array}{l} -back \\ +front \\ -high \end{array} \right]$

This rule correctly deletes the e in the 2nd person singular /klyaušetär/, 3rd person /klyaušetär/, and 2nd person plural /klyaušetär/, but also correctly fails to delete the e in the 2nd person singular /triketär/, 3rd person singular /triketär/, and 2nd person plural /triketär/. However, this rule also incorrectly fails to delete the e in certain cases. For example, the verb kraup- 'anhäufen' (Krause 1952:63; Krause and Thomas 1960:188), exhibits the same distribution of e ~ \emptyset as klyaus-, yet kraup- ends in a [-continuant]. In other words, if the underlying representation of the 3rd person singular klyaušetär is taken to be /klyaušetär/, with an e, then the underlying representation of kraup- must be taken to be /kraupetär/, with an e also. Rule C deletes the e of /klyaušetär/ to derive klyaušetär, but fails to delete the e of /kraupetär/ and incorrectly derives kraupetär. The correct surface form of the 3rd person singular of kraup- is kraupetär, without e. Therefore, the loss of the vowel in the klyaus- forms clearly has nothing to do with the fact that the vowel follows a [+continuant] in these cases, because there are also cases where the vowel is lost following a [-continuant]. Moreover, other than the difference between a preceding [-continuant] versus a preceding [+continuant], the klyaus- forms and the trik- forms are phonologically undifferentiable, yet in the klyaus- forms, there is an alternation between e and \emptyset , but in the trik- forms, there is not.

The first serious challenge to the strong forms of principle 3 has now been encountered, for Tocharian B possesses two different classes of verbs, one of which exhibits surface allomorphy (verbs like klyaus-), and one of which does not (verbs like mäsk- and trik-). Thus, it appears that the surface allomorphs of klyaus-, namely, klyause- and klyaus-, must simply be listed, for there appears to be no way to distribute them phonologically. However, though the deletion analysis does not work, the insertion alternative has not yet been examined, and perhaps this alternative provides a way out of abandoning the strong form of principle 3. Unfortunately, the insertion analysis does not work either, and this fact can be demonstrated readily. If an insertion analysis were proposed, the underlying representation of the 1st person singular of klyaus- would be /klyausmar/, and the application of the rule would correctly derive klyausemar. However, a verb like kaläk- 'folgen,' is of the same type as pälk- (p. 63 above), and thus, has a consonant final stem, to which endings are added, as has been demonstrated for pälk- above. The first person singular

deponent-passive of kalāk-, however, is kolokmar, without e. If Tocharian B had an insertion rule which derived klyausemar from /klyāusmar/, then it would be expected to apply to an underlying form like /kolokmar/ as well, yet no such insertion rule applies, because /kolokmar/ surfaces as kolokmar. Thus, neither an e deletion nor an e insertion analysis works for the forms of verbs like klyaus-, and the strongest form of principle 3 must be modified. This question is taken up again later in this section, but now, the alternation between s and š must still be accounted for.

Since š occurs only before t, and s never occurs before t in the klyaus- forms, it might be hypothesized that it is this factor which causes the alternation, and that s becomes š preceding t. Though such a rule would work for the klyaus- paradigm, it does not work for other paradigms that exhibit alternations very similar to those in klyaus-, for example, the active paradigm of āk- 'führen':

Singular	1	ākau	Plural	1	akem ¹⁴
	2	āšt ¹³		2	āšcer
	3	āšäm		3	ākem

It can be seen that āk- shows an alternation between k and š in exactly the same persons that klyaus- shows an alternation between s and š. Therefore, unless some way can be found to predict the distribution of s and k as opposed to the palatal alternates š and ś, phonologically, then the allomorphs of all of these verbs must be listed in the lexicon.

Returning to klyaus-, it might be hypothesized that the forms with the š alternate really do have a vowel that follows the stem in underlying representation, which palatalizes the s and then is deleted. It has already been shown above that this vowel cannot be identical in underlying representation to the vowel that surfaces as e in the non-palatalized forms of a verb like klyaus-, and in all of the forms of a verb like māsk-, but perhaps there is a palatal vowel in the underlying representation in the š forms of klyaus- which is different from the underlying vowel in the forms of the verb that do not palatalize. If such a solution can be made to work, then the distribution of the palatal and nonpalatal consonants in a verb like klyaus- does not have to be stated in nonphonological terms, but instead can be predicted by rule.

The first step in such a solution is to determine what the underlying vowel in the palatal forms is. A very natural hypothesis is i or ī, since either of these vowels could quite naturally condition a palatalization of the type found in klyaus-, and since it would also be quite natural for a high vowel to be lost in a medial unstressed syllable. Considering the possibility of i first, it can be seen that this vowel clearly conditions palatalization of just the sort found in klyaus- (and also in āk-), for example, rošicer, second person plural imperfect of rok- 'leuchten' (Krause 1952:104). However, this same form also demonstrates that i cannot be the underlying vowel that is being deleted in the palatalizing verbs, because here i is retained in a medial syllable; in āšcer, second person plural of āk- above,

on the other hand, it would have to be claimed that the i was deleted if an analysis with an underlying i in the palatal forms of the verb is to be defended.

Turning to the possibility of an underlying i that palatalizes and is then deleted, problems exactly analogous to those with i arise. The 3rd person singular and 2nd person plural imperfect of täs- 'setzen' is tašitär (Krause 1952:105), again with i showing palatalizing effects, but also being retained, where it should be deleted, if i is to be defended as the underlying vowel in the palatal forms of verbs like klyaus-. tašitär compared directly with klyauštär shows that i cannot be the underlying vowel in the palatalized forms of klyaus-.¹⁵

Another vowel that might naturally be thought to palatalize consonants is ä, but this possibility can be eliminated straightforwardly, since there are many examples of ä not causing palatalization, for example, nesäm, 3rd person singular active of nes- 'sein' (Krause 1952:61), in addition to the 3rd person singular active of pälk-, palkäm, given above.

The only palatal vowel left in the surface inventory of Tocharian B is e, but it has already been shown above that this vowel cannot be underlying in the palatalized forms of klyaus- because there is no phonological way to explain why the vowel would be retained in the nonpalatalized forms, but deleted in the palatalized forms. Further, if it is claimed that an underlying e in the palatal forms of the verb conditions palatalization, then it should condition palatalization in all of the forms. In other words, if the underlying representations of the forms of klyaus- all have an /e/, then the surface e in the 1st person singular and in the 1st and 3rd persons plural acts as though it were not an /e/, but the surface \emptyset in the other persons behaves as though it were an /e/, because it is associated with palatalization. Thus, if there were a vowel underlyingly in all six forms, it could not be the same vowel in every form. This problem is not new, however, for it has already been shown above that there is no way to have a single underlying representation for verbs like klyaus-. However, it can be seen that, although these verbs cannot have a single underlying representation, the palatalization in certain forms of the verb could be predicted phonologically if the palatalized forms had an underlying palatal vowel, and the nonpalatal forms had an underlying nonpalatal vowel; thus, surface e must be derived from a nonpalatal vowel, and surface \emptyset from an underlying palatal vowel. It has been demonstrated above that e is the most likely of the Tocharian palatal vowels to be underlying in the forms of the verb that have palatalized consonants. Using this vowel as underlying in these forms of the verb requires that surface e be derived from something other than underlying /e/, but this fact fits perfectly with the independent observation that surface e does not act like a palatal vowel.

If the palatal forms of verbs like klyaus- are underlyingly /klyause-/, it remains to be determined what the underlying representation is of the forms that do not have palatalization of the surface. In other words, the exact vowel that underlies surface e in the 1st person singular and 1st and 3rd persons plural must be determined. First, it can be seen that surface e in these cases cannot be derived from underlying i or ī, because these vowels always cause palatalization of a preceding

consonant (if the consonant has a palatalized form), and surface e clearly does not condition palatalization. Second, it would not be possible to derive surface e from any of the vowels involved in the a ~ ä or ā ~ a alternations, because it would not be possible to predict, for example, if /ä/ were chosen as the underlying representation of e, which instances of underlying /ä/ would surface as ä, which as a, and which as e. In other words, not all occurrences of underlying /ä/ could be converted to e, because some of them must surface as ä, and some others as a, and there is no phonological way to predict when /ä/ is converted to e, and when it is not.

The next closest vowel to e in terms of features is o, and since o is not involved in any surface alternations, it would be possible to simply convert all occurrences of underlying o into e. However, it can be seen that it is not necessary to convert every underlying /o/, including those that do not figure in any alternations, into surface e, but only those which appear in nonpalatalizing forms of the palatal verbs. In other words, most occurrences of surface e can be derived from underlying /e/, since they are not involved in any alternations, and most occurrences of surface o can be derived from underlying o, since they are not involved in alternations. Only occurrences of surface e that act as though they were underlyingly nonpalatal need to be derived from underlying /o/. By converting only those occurrences of underlying /o/ that appear in palatalizing verbs into surface e, one can account for the alternations in these verbs, and the strongest form of principle 4 can still be maintained. The underlying /e/ vowels that must cause palatalization and then be deleted, as well as the underlying /o/ vowels that must be converted to surface e, always occur at the end of a verbal stem, and therefore, always before a morpheme boundary. Thus, the rules that are required convert /e/ to \emptyset and /o/ to e preceding a morpheme boundary. All occurrences of /e/ and /o/ not preceding a morpheme boundary are immune to the rules, and surface as e and o, respectively. One other rule, however, must be mentioned. Specifically, if the palatalization rule is used to account not only for the palatal alternations in verbs like klyaus-, but also in verbs like äk-, then the underlying representations for these two verbs are as follows:

Singular	1	/klyauso+mar/	Plural	1	/klyauso+mt(t)är/
	2	/klyause+tar/		2	/klyause+tär/
	3	/klyause+tär/		3	/klyauso+ntär/
Singular	1	/ako+au/	Plural	1	/ako+emo/
	2	/ake+t/		2	/ake+cer/
	3	/ake+än/		3	/ako+en/

It can be seen that in the 1st person singular, 1st person plural, and 3rd person plural of äk-, the stem vowel, o, precedes an e vowel which occurs initially in the ending, but there is no trace of the stem vowel in the surface forms. Thus, a rule is needed to delete a vowel before another vowel, and further, this deletion rule must apply before stress is assigned, because the stress pattern of the surface forms indicates that the underlying vowel at the end of the stem does not count as a

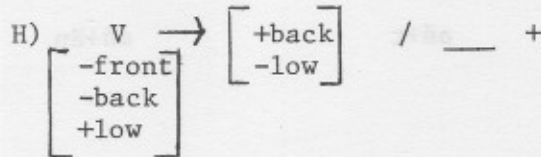
syllable when stress is assigned, so the vowel must be deleted before stress assignment. The rules that are required for the derivation of the forms of klyaus- and āk- are as follows:

- D) $\left\{ \begin{array}{l} \left[\begin{array}{l} -\text{vocalic} \\ +\text{back} \end{array} \right] \longrightarrow \left[\begin{array}{l} -\text{back} \\ +\text{continuant} \end{array} \right] \\ \left[\begin{array}{l} +\text{anterior} \\ +\text{continuant} \end{array} \right] \longrightarrow \left[\begin{array}{l} -\text{anterior} \\ +\text{high} \end{array} \right] \end{array} \right\} / \overset{V}{-} \left[\begin{array}{l} +\text{front} \\ -\text{high} \\ -\text{low} \end{array} \right]$
- E) $\overset{V}{\left[\begin{array}{l} +\text{front} \\ -\text{high} \\ -\text{low} \end{array} \right]} \longrightarrow \emptyset / _ +$
- F) $\overset{V}{\left[\begin{array}{l} +\text{back} \\ -\text{high} \\ -\text{low} \end{array} \right]} \longrightarrow [+front] / _ +$
- G) $V \longrightarrow \emptyset / _ V$

The palatalization rule, D, must precede the deletion of e, which is accomplished by rule E, and which must precede rule F, because if F were allowed to precede E, it would feed it; that is, the underlying /o/ vowels which F converts to e would be deleted if E were ordered after F, but these derived e vowels actually surface, which demonstrates that F must counterfeed E (must be ordered after it). There are two restrictions on rule G. First, it must be ordered after D; otherwise, G would delete /e/ in the 3rd person singular of āk- before the e had a chance to palatalize the k. Therefore, G must counterbleed D (must be ordered after it). Second, as stated above, rule G must apply before stress is assigned. Rule E must also apply before stress is assigned, because, for example, the surface form of /klyause+tar/ is stressed on the first syllable, which indicates that it only has two syllables at the time that stress is assigned.

Before complete derivations for the forms of klyaus- and āk- are given, one last problem with the analysis must be solved. The previous analysis of the verb plānt- (pp. 64-65 above) takes the underlying representation of the stem to be /plonto/, but it can be seen now that rule F above would convert this underlying representation into the incorrect surface form plonte-. The problem here is that since underlying /o/ preceding a morpheme boundary is converted into surface e, surface o preceding a morpheme boundary must be derived from some vowel other than /o/. Since every vowel of Tocharian B but one occurs as an underlying vowel also, it would put an end to the shuffling around of the underlying vowels of the system if surface /o/ before a morpheme boundary could be derived from the one surface vowel that does not occur in underlying representations, namely, ā. Thus, the underlying forms of plānt- are as follows, and the new rule which is needed to convert /ā/ before a morpheme boundary to surface o is given as H:

Singular 1	/plontā+mar/	Plural 1	/plontā+mt(t)är/
2	/plontā+tar/	2	/plontā+tär/
3	/plontā+tär/	3	/plontā+ntär/



Rule H must counterfeed F, because the output vowel of H, o, must not be input to F, or else it would be converted into e, and underlying /ā/ would surface as e, which is incorrect. Thus, H must be ordered after G.

The complete derivations of the forms of klyaus- and āk- can now be given:

	Singular 1	2	3
<u>RULE</u>	/klyauso+mar/	/klyause+tar/	/klyause+tär/
D		klyause+tar	klyause+tär
E		klyaus+tar	klyaus+tär
F	klyause+mar		
G			
H			
B	klyausemar	klyau [•] star	klyau [•] tär
	Plural 1	2	3
<u>RULE</u>	/klyauso+mt(t)är/	/klyause+tär/	/klyauso+ntär/
D		klyause+tär	
E		klyaus+tär	
F	klyause+mt(t)är		klyause+ntär
G			
H			
B	klyausemt(t)är	klyau [•] stär	klyause [•] ntär

	Singular 1	2	3
	/ako+au/	/ake+t/	/ake+än/
<u>RULE</u>			
D		aše+t	aše+än
E		aš+t	aš+än
F	ake+au		
G	ak+au		
H	āk+au	āš+t	āš+än
B			āš+äṇ
	ākau	āšt	āšäṃ
	Plural 1	2	3
	/ako+em/	/ake+cer/	/ako+en/
<u>RULE</u>			
D		aše+cer	
E		aš+cer	
F	ake+em		ake+en
G	ak+em		ak+en
H		āš+cer	āk+en
B			āk+eṃ
	akem	āšcer	ākeṃ

Given the analysis to this point, certain forms which have not been considered yet in the paradigms of the palatalizing verbs present potential problems; specifically, the following forms require comment (Krause 1952:209):

Active Participle klyaušeñca

Deponent-Passive Participle klyausemane

Several points need to be made concerning these forms. First, the segmentation of these forms into morphemes should be carefully observed. The e in the active participle must be part of the ending, even though the stem shows palatalization in this form. This segmentation is required because if this e were part of the stem underlyingly, it would derive from underlying /o/, and therefore would not cause palatalization. Thus, this e must be part of the ending, and the palatalization of the stem is caused by underlying /e/ in the stem, which is deleted via rule E. The underlying form of the active participle, therefore, is:

/klyause-eñca/

In the deponent-passive participle, on the other hand, the e that appears on the surface must derive from underlying /o/, just as other surface occurrences of e which are not associated with palatalization derive from underlying stem final /o/. Therefore, the underlying representation of this form is:

/klyauso-mane/

The crucial point that must be made concerning the participial forms is that both forms exhibit surface e in the second syllable, but this e must be analyzed in two very different ways in the two forms. Moreover, the treatment of this vowel in each case is dictated entirely by whether or not the vowel is associated with palatalization of the verb stem.

2.2. The Argument against Generative Theory: Palatalizing Causatives.

The generative analysis of the data from Tocharian B is now complete, and it appears that the generative analysis is able to account for the Tocharian forms, except for a weakening of principle 3 that is required by alternations in the palatalizing verbs. There are, however, other forms in Tocharian B which the generative framework cannot give a satisfactory account of. Specifically, it is argued that these forms provide evidence that a correct account of the palatalizing verbs, which are analyzed above phonologically, is not a phonological account at all, but rather, a morphological one. It is further argued that these forms provide evidence against generative phonological accounts in general. In this section, this new evidence is presented, and the arguments against generative phonology which are based on this evidence are given.

The causative forms of some verbs in Tocharian exhibit palatalizations very similar to those already observed above in the (non-causative) forms of verbs like āk- and klyaus-. Specifically, some verbal stems with initial k- have causative forms with initial ś-, and certain stems with initial s- have causative forms with initial ś- (Krause 1952:217-309):

kātk- 'überschreiten'

Present Causative Stem śātk- (e.g., participle, śātkässeñca)

Preterite Causative Stem śātk- (e.g., 2nd sg. deponent-passive śātkatai)

kārs- 'wissen'

Present Causative Stem śārs- (e.g., 3rd sg. śārsäṣṣäm)

Preterite Causative Stem śārs- (e.g., 3rd sg. śārsa)

kāl- 'führen, bringen'

Present Causative Stem No forms attested

Preterite Causative Stem śāl- (e.g., 1st sg. deponent-passive śālamai)

kau- 'toten'

Present Causative Stem No forms attested

Preterite Causative Stem śau (e.g., 1st sg. śauwwa)

säl- 'springen'

Present Causative Stem şal- (e.g., deponent-passive participle şaläskemane)

Preterite Causative Stem şäl- (e.g., 3rd pl. şaläre)

spärtt- 'sich drehen'

Present Causative Stem şpartt- (e.g., 3rd sg. şparttaşşäm)

Preterite Causative Stem şpyart- (e.g., 3rd sg. şpyarta)

spänt- 'vertrauen'

Present Causative Stem şpant- (e.g., gerundive şpantäşälona)

Preterite Causative Stem No forms attested

spärk- 'vergehen'

Present Causative Stem şpark- (e.g., 3rd sg. şparkäşşäm)

Preterite Causative Stem şpyark- (e.g., 2nd sg. deponent-passive şpyarkatai)

The alternations in these forms are similar to the previously observed ones in that k alternates with s and s alternates with ş; however, they are very dissimilar in that there is apparently no way to analyze the palatalizations in the causative forms as being phonologically conditioned. In particular, the analysis of the non-causative alternations, as in äk- and klyaus- where underlying /e/ before a morpheme boundary causes palatalization and is then deleted, cannot be extended to the causative alternations. The reason that the previous analysis cannot be extended to the causative cases is that these alternations do not take place before a morpheme boundary, at the end of a stem, but rather, at the beginning of a stem, with no following morpheme boundary. Therefore, if /e/ appeared in the underlying representations of the causative forms of the above verbs, there would be no way to delete this vowel by using the rules that have already been formulated. For example, if the underlying representation of the present causative of kätk- were (keätk-/), the underlying /e/ would correctly palatalize the initial /k/ of the stem to ş, but there would then be no way to delete the e using the rules developed thus far, because it does not occur before a morpheme boundary; therefore, the e would incorrectly surface giving *şeätk- (or, if stressed, *şeatk-). Thus, the previous analysis cannot account for the causative alternations, and clearly, these alternations must be accounted for.

The most natural way to account for the causative forms is to revise the previous analysis; however, there does not seem to be any way to revise the analysis in order to account for these forms without creating irresolvable problems at the same time. If the causative stems have underlying /e/, as suggested above, and the e deletion rule is generalized so that it deletes every occurrence of underlying /e/, then the palatalizations in the causative forms are correctly predicted, and the underlying e vowels that condition these palatalizations correctly fail to appear in the surface forms. However, such a generalization of the e deletion rule has extremely undesirable consequences; specifically, other parts of the original analysis must also be revised in such a way that the strong form of principle 4 must be abandoned. In the original analysis, surface occurrences of e that are not involved in alternations are derived from underlying /e/, but with

the newly revised e deletion rule, this line of derivation for nonalternating e will no longer be possible, because now, all occurrences of underlying e are deleted. Therefore, all occurrences of surface e must now be derived from some vowel other than underlying /e/. Since some occurrences of surface e derived from underlying /o/ in the original analysis (by rule F), it seems simplest to now derive all occurrences of surface e from underlying /o/, which could be accomplished by removing the morpheme boundary from rule F. This move, however, means that surface occurrences of o that do not alternate, which were originally derived from /o/, must now be derived from some other underlying vowel, since all occurrences of underlying /o/ will now be converted to surface e. Since some occurrences of surface o derived from underlying /ā/ in the original analysis via rule H, this rule can now be generalized by removing the morpheme boundary, so that all surface occurrences of o, whether they are involved in alternations or not, are now derived from underlying /ā/. Thus, generalizing the e deletion rule requires that the o → e and a → o rules be generalized also, so that every occurrence of three different underlying vowels in Tocharian B now surfaces in a form different from its underlying form, which seriously violates principle 4 above. Of course, as previously argued, it would be desirable to hold to as strong a form of this principle as possible, but generalizing these three rules does not seem to allow holding to this principle in any form. What is clearly worse, though, than the necessity of abandoning principle 4 is the fact that the generalized form of the e deletion rule results in Tocharian B having an underlying segment which never surfaces in any segmental form. It leaves behind traces of its presence in underlying representation, by causing palatalization, but the segment itself never surfaces. Allowing languages to have segments in their underlying inventories that never actually surface considerably enlarges the class of the possible phonologies of human languages, and therefore ought not to be allowed if possible.

Beyond the general theoretical problems involved in trying to generalize rules E, F, and H, there are reasons internal to Tocharian for rejecting this move as a possible way of accounting for the palatalized causative forms. For these same reasons, it can be argued that no phonological account of these forms is possible. In particular, the first consonant of the stem appears in exactly analogous phonological environments in both the causative forms, where the consonant is palatalized, and in the noncausative forms, where the consonant is not palatalized. For example, the k in the non-causative forms of the verb kau- (p. 74 above) appears word-initially before the diphthong au, and the ś in the preterite causative stem appears in exactly the same environment. Similarly, the s in the non-causative forms of the verb spant- (p. 75 above) appears in exactly the same environment as the ś in the present causative stem. Presumably, in generative theory, the non-causative and causative forms of these verb morphemes should be the same underlyingly, since they are the same morpheme. Therefore, it follows that there could be no phonological difference between the underlying representations of the causative and the non-causative forms that could cause the difference between the palatal and nonpalatal consonants in the stem.

Using a line of reasoning employed by Halle (1959:21-23) to argue against classical phonemics as a correct theory of phonology, the Tocharian facts provide an analogous argument against generative phonology as a correct theory of phonology. Halle argued that, since some voiced obstruents

in Russian are allophones, while others must be considered phonemes, that two different obstruent voicing rules would be required by classical phonemics in order to account for the fact that in Russian all obstruents are voiced before a voiced obstruent. In effect, as Halle sees it, one obstruent voicing rule would be required to get from the morphophonemic to the phonemic level, which would operate only on those voiceless obstruents whose cognate voiced obstruents have the status of phonemes in Russian, and another obstruent voicing rule would be required to get from the phonemic to the phonetic level, and this rule would operate only on those voiceless obstruents whose cognate voiced obstruents have the status of allophones in Russian. However, Halle argues, this treatment breaks the obstruents up into two classes and requires an "extra" obstruent voicing rule, when there really should only be one voicing rule, since, whether the rule operates to derive segments that have the status of phonemes or the status of allophones, it achieves exactly the same effect, namely, making obstruents voiced before voiced obstruents. Halle concludes that if the "bi-uniqueness condition", and therefore also the phonemic level of representation, is dropped, then Russian obstruent voicing can be covered by a single general rule.

Accepting Halle's reasoning for the moment without argument, an analogous case can be constructed against generative phonology, using the Tocharian facts, as follows. In Tocharian, k alternates with ś, and s alternates with ṣ. Some of the instances of each of these two alternations can be accounted for within a generative framework by a phonological rule; however, other instances of these alternations (in the palatalizing causatives) cannot be accounted for by that same phonological rule. Therefore, a generative account of Tocharian must treat the instances of a single phenomenon in two different ways (however it is that the palatalizing causatives are handled, it has already been shown that it cannot be phonologically, and therefore, they must be treated differently than the palatalizing present verbs), and therefore misses a generalization about Tocharian.

A possible counterargument to the argument that a generative account must miss a generalization here would be simply to claim that there is no generalization to be captured here; that is, to claim that the instances of k, s alternating with ś, ṣ in the palatalizing causatives and the instances of these alternations in the palatalizing presents are not really the same phenomenon, and therefore need not be accounted for in a single way. There are several reasons why this counterargument is not satisfactory, however. First, each of the instances of k alternating with ś and s alternating with ṣ involves exactly the same two segments, not just the same two classes of segments, as in Russian, where all the alternations involved in the argument are cases of voiceless obstruents alternating with voiced ones. Thus, though it might be said that it is somehow counter-intuitive to claim that all of the cases of the k~ś and s~ṣ alternations involve the same phenomenon, it seems rather that it would be counter-intuitive to claim that they are not. Further, in Tocharian, ś alternates only with k¹⁶ and ṣ alternates only with s. Thus it seems strange intuitively to claim that these alternations are not all one phenomenon.

Since intuition seems to weigh in the favor of the argument against generative phonology, a more principled objection to the argument might be sought. In particular, it might be objected that although the substitutions of ś for k and ṣ for s involve the same segments, the substitutions have different causes. Specifically, the substitutions of ś for k and

s for s in the palatalizing presents like āk- and klyaus- seem to have clearly phonetic teleologies; that is, ś and s being substituted for k and s before the palatal vowel e (this also happens before ī in Tocharian) is clearly phonetically motivated. On the other hand, these same substitutions occurring in the palatalizing causatives apparently have no such phonetic motivation. Thus, the two sets of substitutions really do appear to have different causes; therefore, they could be claimed to be instances of different phenomena, and the fact that a generative analysis must account for them in two different ways is no longer a problem, and is in fact to be expected. However, such an argument is inconsistent with the generative point of view, for in generative theory, multiple causes for a given substitution are never seen as a reason to treat that substitution as several different phenomena. On the contrary, in such a case, all of the different causes for the substitution would be grouped together (using curly brackets), and treated as instances of the same rule. For example, the two rules

$$\begin{array}{l} a \rightarrow b / \text{---} x \\ a \rightarrow b / \text{---} y \end{array}$$

would be collapsed as

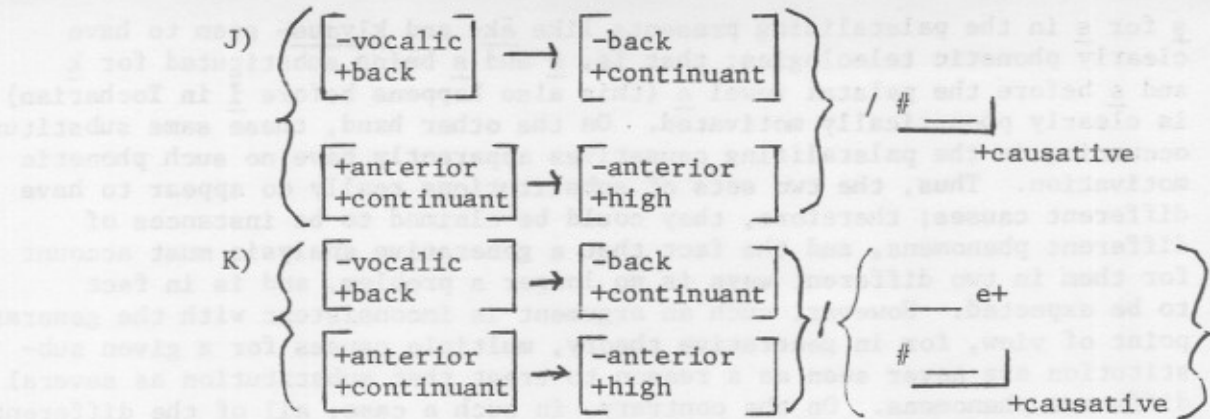
$$a \rightarrow b / \text{---} \{x, y\}$$

The fact that x causes the substitution of b for a in one case, and that y causes it in the other case is virtually irrelevant in generative theory. The two rules are formally similar in a way that makes them collapsible within the theory; therefore, they are to be collapsed. The difference in the causes of the substitution in individual cases is not relevant in generative theory; likewise, then, the different causes of the k~ś and s~s substitutions in Tocharian cannot be used to justify the claim that there is more than one phenomenon. Thus it has been demonstrated, using criteria entirely internal to generative theory itself (Halle's argument from Russian and the criterion of combining formally similar rules), that generative theory cannot provide a correct account of the k~ś and s~s alternations in Tocharian.¹⁷

2.3. On the AdHoc Nature of Generative Phonological Analysis.

It might be suggested that there really is a generative account for the palatal alternations in Tocharian, despite the fact that the alternations in the present can be accounted for phonologically, while the alternations in the causative are not phonologically conditioned. This account would consist of collapsing the two rules which would be needed to account for both of the two types of palatal alternations, in the present palatalizing verbs and in the palatalizing causatives, into a single rule. In other words, rules I and J below could be collapsed into rule K:

$$I) \left\{ \begin{array}{l} \left[\begin{array}{l} -\text{vocalic} \\ +\text{back} \end{array} \right] \rightarrow \left[\begin{array}{l} -\text{back} \\ +\text{continuant} \end{array} \right] \\ \left[\begin{array}{l} +\text{anterior} \\ +\text{continuant} \end{array} \right] \rightarrow \left[\begin{array}{l} -\text{anterior} \\ +\text{high} \end{array} \right] \end{array} \right\} / \text{---} e+$$



Such a move is possible because there is no distinction in generative phonology between rules that involve phonological conditioning of alternations, such as rule I, and rules that do not, such as rule J.¹⁸ That is to say, there is no reason in generative theory why two rules such as I and J cannot be collapsed using the curly brackets notation. Collapsing these two rules seems to get around the argument given in the last section that generative phonology cannot give a unified account of the palatal alternations in Tocharian, even though the criteria of the theory itself demand that a unified account be given. At the same time, however, collapsing rules I and J does not correctly account for all of the facts about the palatal alternations in Tocharian. In particular, there are verbs that have stems with initial k and s that do not palatalize in the causative (Krause 1952:217-309):

kātk- 'sich freuen'

Present Causative Stem kātk- (e.g., participle kāt₁kās₂señca)

kārp- 'herabsteigen'

Present Causative Stem kārp- (e.g., 2nd deponent-passive kārpasträ)

Preterite Causative Stem kārp- (e.g., 2nd sg. karpässasta)

kän- 'zustandekommen'

Present Causative Stem kan- (e.g., participle kans₁señca)

kärn- 'schlagen'

Present Causative Stem kärn- (e.g., participle karnnäss₁señca)

kälp- 'erlangen'

Present Causative Stem kalp- (e.g., 1st sg. kalpāskau)

kery- 'lachen'

Present Causative Stem ker- (e.g., 2nd pl. kerästrä)

krās- 'verdriessen'

Present Causative Stem krās- (e.g., 3rd sg. krāsäss₁äm)

sāk- 'zurückbleiben'

Present Causative Stem sāk- (e.g., 3rd sg. sākäss₁äm)

sätk- 'sich ausbreiten'

Preterite Causative Stem sätk- (e.g., 1st sg. deponent-passive
sätkasamai)

soy- 'satt werden'

Present Causative Stem soy- (e.g., 3rd sg. soyäṣṣäm)

Preterite Causative Stem soy- (e.g., 1st sg. soyäṣṣāwa)

staukk- 'müde werden'

Present Causative Stem staukk- (e.g., 3rd sg. staukkäṣṣäm)

swār- 'gefallen'

Present Causative Stem swār- (e.g., 2nd pl. swārāsträ)

Preterite Causative Stem swār- (e.g., 2nd pl. deponent-passive
swāräṣṣat)

The above forms clearly require a revision in rule J (and therefore in rule K), because, as it stands, this rule predicts that all verbs in Tocharian with initial k and s will have the corresponding palatal in the causative, but the above forms do not. Thus, beyond the fact that the k̃ṣ̃ and s̃ṣ̃ alternations in the causatives cannot be phonologically predicted, and must therefore be accounted for by a rule that mentions the nonphonological category causative, this rule must also be lexically restricted; i.e., it applies to some lexical items, but not to others. This lexical restriction is necessary because there is no other way, phonological or otherwise, to distinguish the causatives with palatalization of the initial consonant from those without palatalization. If the rule is restricted in this way, then rule K seems to correctly account for the palatal alternations in Tocharian.

Though it now appears that a generative account of all of the palatal alternations in Tocharian can be given, it must be pointed out that the devices which are needed for this account are extremely powerful; in fact, it can be argued that all of these devices together yield a phonological framework which is generatively omnipotent, and which allows for, in fact, necessitates, totally ad hoc analyses. If the methodology that is used in analyzing the Tocharian data is examined, the ad hoc nature of the analysis can be seen clearly; further, it can be seen why the four principles of generative phonology given above (pp. 58-9) do not really constitute any significant restrictions on, or predictions about, the phonologies of human languages.

First, in the analysis of the present palatalizing verbs (pp. 66-72), it is found that the alternations in these verbs cannot be accounted for by phonological means alone, but that, if the vowel at the end of the stem is underlyingly /o/ in the 1st person singular and in the 1st and 3rd persons plural, and /e/ in the other persons, then the alternations can be accounted for by phonological means. However, as noted before, this move means that the present palatalizing verbs like klyaus- and āk- cannot have a single underlying representation. Thus, just on the basis of the analysis of the present palatalizing verbs, it can be seen that principles 2 and 3 (pp. 58-9) are not absolute restrictions. Phonological rules that mention only phonological information and single underlying representations for every morpheme in the language are preferred, but generative theory makes no guarantees that these preferences will not have to be violated in some cases. Preferences guide an analyst in choosing from among a number of different possible analyses, but clearly they have

no value in making a theory more restrictive, for, though one type of language or another is preferred, the theory stills does not predict that the other types cannot exist. If the possibility of restricting phonological rules lexically, as is necessary if a generative account of the causative palatalizations is to be given, is considered, it can be seen how totally ad hoc analyses are possible. In the Tocharian case, as in all cases, a phonological account is considered first, but if none is available, the theory is still not disconfirmed, because morphological categories or other nonphonological information may still be used. This move of resorting to nonphonological information is the step that seems required for the Tocharian causatives; in other words, once a rule that mentions the category causative is formulated, the forms seem to be accounted for. However, when it is found that there are other forms that the rule does not account for, and further, that now lexical restrictions must be placed on the rule, the theory is still not disconfirmed. Of course, a lexical account cannot fail, because the forms that the rule should apply to, it is allowed to apply to, and the forms that it should not apply to are simply excluded from its application. If a phonological theory is allowed to condition phonological rules lexically, then it cannot be disconfirmed, but if generative theory is prevented from conditioning phonological rules lexically, then it is disconfirmed by the case of the Tocharian causative forms, because it cannot account for them in any other way.

The reason then, that generative analyses are necessarily ad hoc is that the analysis can always be tailored to fit the facts at hand. A restrictive theory of phonology, by contrast, would predict that certain facts will never be encountered, and would be falsified if such facts should be discovered in some language. In the next chapter, a proposal is outlined that would constitute a restrictive theory of phonology, which at the same time requires the formulation of a theory of morphology, which, it is argued, can also be restricted in significant ways.

3. A Morphological Account of the Tocharian Palatal Alternations. 3.1. Separating Phonology and Morphology.

It can be argued that one of the primary reasons that generative phonology is such a nonrestrictive theory is that it fails to make any distinction between purely phonological rules, that is, rules that contain only phonetic information, and morphological rules, that is, rules that may mention morphological categories and boundaries. When phonological rules are allowed to be morphologically or lexically conditioned, they are very powerful devices, and if no explicit restrictions are placed on them otherwise, they can potentially generate virtually any phonological output from any given input. However, if phonological rules are restricted so that they contain only phonetic information, that is, segmental features, and phonological boundaries, i.e., syllable, word, and phrase boundaries, then their output is considerably more restricted. Further, in a theory of phonology where the possible phonological substitutions are limited, such as in the theory proposed by Stampe (1979), the generative power of the theory is yet more restricted. Theories of phonology in this sense of phonology, however, since they deal only with phonetically conditioned substitutions, could not account for alternations of the type found in the Tocharian palatals. Since these substitutions are morphologically conditioned in some way, a device separate from phonological rules needs to be developed to account for them, and further,

in keeping with the goal of linguistic theory, it must be shown how this device can be restricted so that it excludes certain language types and therefore has empirically testable consequences.

3.1. An Account of the Tocharian Palatals.

The device that is used to account for the palatals in Tocharian must correctly predict in what places they occur, and it must account for which stems exhibit the alternations, and which do not. In order to do this, the proposal being made here utilizes a particular way of listing morphemes in the lexicon, along with a rule that distributes the allomorphs of each morpheme. The lexical entries required would be the following:

klyaus/ʃ : 'hören'

ak/ś : 'führen'

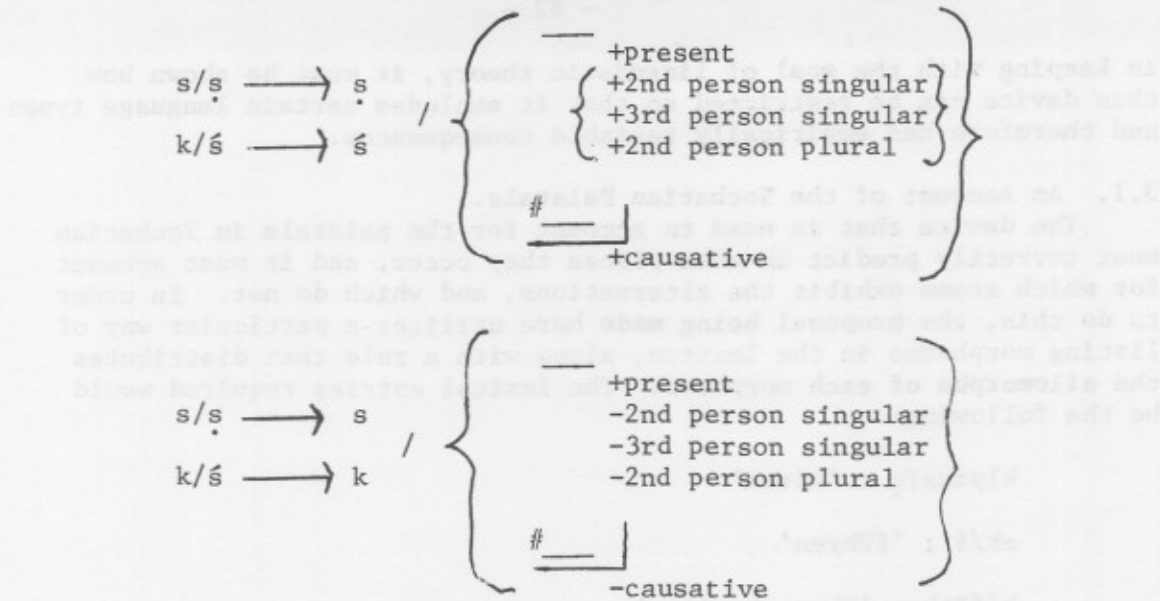
k/śätk : 'überschreiten'

s/śäl : 'springen'

katk : 'sich freuen'

sak : 'zurückbleiben'

The slash line is to be interpreted as meaning that the segment at that location in the morpheme can be any one of the segments given on either side of the slash line. This device seems to capture what speakers actually know, since for a given morpheme, speakers must know whether the morpheme has allomorphs, and if so, what those allomorphs are. Morphemes such as katk and sak are identified as nonalternating because they have only a single allomorph listed in the lexicon. The reason for not listing only one allomorph in the lexicon for morphemes that do alternate, and treating the other allomorphs as derived by rule, is that the choice of which allomorph is derived and which allomorph is lexical would be entirely arbitrary. Until some principled basis for treating one or the other of the allomorphs as derived can be found, it seems justified to list all of the allomorphs in the lexicon. A rule to distribute the allomorphs of the alternating morphemes would be required also, formulated as follows:



These rules, along with the lexical entries above, correctly account for which morphemes in Tocharian exhibit palatal alternations and which do not, and also correctly accounts for where the palatal segments occur, and where the nonpalatal segments occur. Now, though, it must be shown how these devices make predictions about possible language types. In order to do this, however, the phonemes of the language must be determined. The meaning of phoneme intended here is that of Stampe (1979). The restriction then, on the type of morphological rule proposed above is that the output of such a rule will always be a phoneme of the language, as determined by the phonology of the language. Further, the segments involved in alternations which can be accounted for by these rules will always occur in some positions in the lexicon where they do not alternate. As can be seen, this restriction holds in Tocharian; that is, the above rules account for a morphological alternation between k, s̄ and ś, s̄ in Tocharian, and k and s occur in morphemes in other places in the lexicon where they do not alternate, for example, in kātk- and sāk- (p. 82). ś and s̄ also occur nonalternating in Tocharian, for example, in śāmp- 'übermütig sein', and šāms- 'zahlen', where the ś and s̄ in the two stems do not alternate with any other segments.

The restriction proposed above for the type of morphological rule given on pages 82-83 constitutes an empirical hypothesis about human languages; therefore, this proposal is empirically falsifiable. Further, the sort of case that would falsify this proposal is clear. If a morphological rule were to produce a segment which did not occur as a nonalternating segment anywhere in the lexicon, then the rule would be a counterexample to the proposal. Therefore, this proposal is testable in a way that the principles of generative phonology are not. Further, since the output of these morphological rules is always a phoneme of the language, the final output of all the morphological rules, which all apply before any phonological rules, will always be in terms of exactly the same segments that occur in the lexicon. Thus, the morphological component cannot "create" any new segments; it is severely restricted. Further, this restriction is a significant one, because it limits the generative power of morphological rules individually and of the morphological

component as a whole. Further, as a result of the fact that the types of morphological rules proposed above manipulate only (classical) phonemes of the language, it follows that speakers are aware of the alternations in morphemes which are accounted for by morphological rules (since these rules substitute phonemes for other phonemes), and that they are never aware of alternations in morphemes which are due to phonological rules (in the narrow sense of phonology referred to above). This fact has some interesting empirical consequences. The most important of these consequences is that alternations due to morphology are subject to analogy, but alternations due to phonology are not. Thus, it is predicted that when clear cases of analogy are found, they will always be cases of speakers simplifying, extending, or reanalyzing alternations due to morphology. Again, this constitutes a further empirical prediction about human languages, because of analogies on phonological alternations were found, the prediction would be falsified.

4. Conclusion.

The sort of account that the Tocharian palatal alternations force a generative theory to give shows particularly vividly the ad hoc nature of generative analyses. Specifically, the nature of the Tocharian data forces a generative approach to abandon the attempt to provide a strictly phonological account and to retreat further and further from the position that all alternations can be accounted for using strictly phonological information. The Tocharian facts drive the generative analysis eventually all the way to the extreme position that phonological rules may be lexically restricted. Certainly, however, it must be concluded that if the theory is forced to allow phonological rules to be morphologically and lexically conditioned, then there is no value in intractably maintaining the position that all alternations can be accounted for by "phonological" means. As is suggested above, once the need for reference to morphological information in order to account for at least some alternations is conceded, then it seems reasonable to propose that morphology and phonology are in fact separate components of grammar, and that, in such a theory, phonological rules are not sensitive to any nonphonetic information. Such a restriction on phonological rules is a significant one, and as proposed above, a separation between phonological and morphological rules also allows restrictions on the morphological component of grammar. However, even if this restriction on morphological rules turns out to be incorrect, this fact would not lessen the validity of the arguments given here against generative phonology as a correct theory of the phonologies of human languages.

Notes

¹For other expressions of this view, see Chomsky 1965:6, 27; Chomsky and Halle 1968:4; Wall 1972:295-296.

²Actually, generative phonologists did not originally hold to the position expressed by this principle, but rather, to another position known as the "free ride" principle, also stated by Zwicky (1972:158):

"Choose the representation that results in the longest derivations."

Principle 4 is included here because it is a significant attempt to limit the possible phonologies of human languages. In fact, Zwicky (1972:158) gives this principle as one of two conditions imposed by Kiparsky (ms. 1968;1971) on abstract phonological analyses.

³Krause and Thomas do not give the inventory of sounds in the form of a chart, but this format is used here for clarity.

⁴Krause and Thomas (1960:39) call this sound "palatalisiertes s," but give no further description. It seems reasonable to interpret this sound as something in the area of a palatoalveolar, but whether this interpretation is exactly correct or not, it makes no crucial difference for the arguments given here.

⁵The gloss provided for this form, and for all subsequent forms, is taken directly from the work cited as the source of the form.

⁶The absence of w in kulypelle and kursarwa is apparently due to the fact that w is never written before u in Tocharian B.

⁷The best evidence for the placement of stress in Tocharian B is the alternation that occurs between a ~ ǣ and ā ~ a discussed above in section 1.1 (Krause and Thomas 1960:43; Krause 1952:10 ff.).

⁸The citation form used to refer to a given verb is taken from Krause and Thomas (1960) or Krause (1952), and is not meant to imply anything about the underlying representation of the verb. Unless otherwise noted, the paradigms of all verbs given here can be found in Krause and Thomas (1960).

⁹Krause and Thomas actually give this form as pälkem(o). Krause (1952:7-8) says, concerning the final o that sometimes occurs:

"Ein bewegliches -o findet sich häufig -und nach Bedarf des Metrums- in poetischen Texten im absoluten Auslaut da, wo im Indogermanischen ein (später apokopierter) Vokal vorhanden war."

Therefore, since the form with final -o occurs only in poetic texts, under the influence of the meter, the form without final -o is here taken to be the normal one.

¹⁰For the phonological feature system used here, see Chomsky and Halle (1968). One feature has been added to the system found there for the purposes of the analysis given here. Specifically, Chomsky and Halle have no way to differentiate central vowels unambiguously from all other vowels. Thus, the feature front is used here for this purpose.

¹¹Zwicky (1972:154) states that "distributional restrictions on phonological elements" are generally taken to be among the "data to be comprehended by a phonological analysis." Thus, phonologists typically assume that segments that have a limited distribution are derived rather than underlying; that is, if a surface segment occurs only in restricted contexts, it is generally assumed that the distribution of that segment is to be accounted for phonologically, and not lexically. From this it follows that, in "orthodox" phonology, the occurrence of a given segment in underlying representation should be unrestricted (except, of course, for the sorts of restrictions that are handled by morpheme structure constraints). Thus, possible restrictions like those discussed above on the underlying occurrence of ä and ā would usually be seen as unnatural, and therefore, a solution which results in such restrictions is to be avoided if possible.

¹²trik- is given here to demonstrate that the deletion analysis is not possible, because it might be claimed that the difference between a verb like māsk-, where the e always surfaces, and a verb like klyaus-, where the e sometimes does not surface, is that the e is preceded by two consonants in māsk-, but only by one in klyaus-. A verb like trik-, where a single consonant precedes the e, yet where the e always surfaces, shows that the difference between one and two consonants preceding the e is irrelevant to whether the vowel is retained or deleted.

¹³This form also has final -o in poetic texts.

¹⁴Again, this form shows final -o in poetic texts, and is stressed on the second syllable, which explains the a in the first syllable, in contrast to ā in the other forms.

¹⁵In addition to rošicer and täšitār, there are a large number of other cases of medial i and i causing palatalization but also being retained (Krause 1952:217-309):

aišimar imperfect of aik- 'wissen, erkennen'
aušimar imperfect of auk- 'aufwachsen, zunehmen'
krašiyate durative of krās- 'verdrriessen'
klyaušiyem imperfect of klyaus- 'hören'
našitar optative of näk- 'vernichten'
paššimar imperfect and optative of pāsk- 'hüten'
prekšitar imperfect of prek- 'fragen'
trisimar causative optative of trik- 'in die Irre gehen'

¹⁶ś also alternates with s, but this is always following another s, e.g., paṣṣimar, imperfect (and optative) of pāsk- 'hüten,' with -ṣṣ- from -sk- before i. This, of course, does not bear on the issue at hand, because ś usually alternates only with k, and s with s.

¹⁷David Stampe (1979:79, note 17) has also shown that Halle's argument applies to generative theory itself, but Stampe's argument is different than the one given here. Stampe has also argued (personal communication) that Halle's argument against classical phonemics is valid only if a "level" of representation is conceived of as a "natural break in a linear ordered set of 'rules'," but no one has apparently ever argued that this conception of a level of phonological representation is a correct one. Until this view has been established, Halle's argument against classical phonemics cannot be considered valid. On the other hand, the argument given above in this thesis does not depend on this conception of a level, and therefore is valid regardless of whether Halle's view of phonological representation is correct or not.

¹⁸Of course, as principle 2 above (p.58) states, rules with only phonological conditioning are preferred, but that does not mean that rules such as J are not permitted, only that they are to be avoided if possible. In this case, the data demands that a rule such as J be formulated.

References

- Anderson, Stephen R. 1979. On the subsequent development of the 'standard theory' in phonology. Daniel A. Dinnsen, ed., Current Approaches to Phonological Theory, pp. 2-30. Bloomington: Indiana University Press.
- Chomsky, Noam. 1965. Aspects of the Theory of Syntax. Cambridge: M.I.T. Press.
- Chomsky, Noam, and Morris Halle. 1968. The Sound Pattern of English. New York: Harper and Row.
- Halle, Morris. 1959. The Sound Pattern of Russian. The Hague: Mouton.
- Householder, Fred. 1979. How different are they? Daniel A. Dinnsen, ed., Current Approaches to Phonological Theory, pp. 252-264. Bloomington: Indiana University Press.
- Jasanoff, J. 1978. Stative and Middle in Indo-European. Innsbruck.
- Joseph, Brian. 1980. Linguistic universals and syntactic change. Language, Volume 56:345-370.
- Kenstowicz, Michael, and Charles Kisseberth. 1979. Generative Phonology. New York: Academic Press.
- Kiparsky, Paul. ms. 1968. How Abstract is Phonology?
- _____. 1971. Historical linguistics. In Dingwall, A Survey of Linguistic Science. Linguistics Program, University of Maryland.
- Krause, Wolfgang. 1952. Westtocharische Grammatik. Band I. Heidelberg: Carl Winter Universitätsverlag.
- Krause, Wolfgang, and Werner Thomas. 1960. Tocharisches Elementarbuch. 2 Volumes. Heidelberg: Carl Winter Universitätsverlag.
- Stampe, David. 1979. A Dissertation on Natural Phonology. New York: Garland.
- Wall, Robert. 1972. Introduction to Matematical Linguistics. Englewood Cliffs: Prentice-Hall.
- Zwicky, Arnold. 1972. The strategy of generative phonology. W. Dressler and F. V. Mareš, eds., Phonologica 1972, pp. 151-168. Munich: Wilhelm Fink Verlag, 1975.

The Role of Deixis in the Development of
Finno-Ugric Grammatical Morphemes*

Leena Tuulikki Hazelkorn

1. Introduction.

Finno-Ugrists have traditionally maintained that the reconstructed Proto-Finno-Ugric or Proto-Uralic¹ mother language had a series of person markers which occur in three different functions in all of the daughter languages: as personal pronouns, subject agreement markers in verbs, and as possessive suffixes (corresponding to English possessive pronouns). With the exception of a few reflexes in the present day languages, their common origin is transparent even to a non-Finno-Ugrist. Scholars have disagreed, however, about which grammatical category these person markers represented in the proto-language. Some maintain that they were personal pronouns which later become suffixes under certain conditions. Others assume that these elements were originally possessive suffixes, and that the independent personal pronouns represent a later development.

The aim of this paper is to present a critical summary of the reconstruction of person markers in the Finno-Ugric languages, as proposed in the available literature,² and to investigate the relationship of the morphemes used to mark persons to morphemes marking other grammatical categories in the daughter languages. This latter aspect has not been discussed in the literature, except in a few contradictory statements. Some scholars have pointed out the similarity between certain grammatical morphemes and the person markers, but to my knowledge an historical account of the similarities has not been offered.

The reconstruction of markers for certain grammatical categories in Proto-Finno-Ugric has been controversial because the scholars have not been able to agree on the form of these morphemes in the proto-language. The disagreements concern the Proto-Finno-Ugric accusative case marking (*-m), and the reconstruction of the plural morpheme (*-t) for the proto-language.

It is the purpose of this paper to demonstrate that the reconstructed person markers did not function in Proto-Finno-Ugric as markers of one of the inherited grammatical categories, but that they had a wider range of application. They were actually general deictic

particles which referred to the roles and locations of the participants in the speech situation. Their development as markers for a variety of grammatical categories may have been initiated in the proto-language, but it continued to evolve in the separate development of each daughter language. In the various languages, different deictic particles were generalized for particular functions. An analysis of these reconstructed person morphemes as deictic particles in the proto-language will also shed light on the controversy surrounding the reconstruction of the grammatical morpheme *-m for the accusative, and *-t for the plural as well.

If we interpret the aforementioned elements as general deictic particles we are able to account for the development of person markers per se, demonstrative pronouns, accusative markers, *-t plurals, and some other grammatical morphemes in a more satisfactory way.

The discussion which follows will offer an explanation of how the original deictic elements developed into markers of various grammatical categories. The major semantic feature shared by all of these grammatical categories is definiteness. Historically, the₃ deictic particles could be used either in the focusing function, when the speaker wished to focus or emphasize a constituent, or in the topicalizing function, to mark the given (old) information in the utterance. It is proposed here that in the history of the Finno-Ugric languages, the use of deictic particles in the focusing function is chronologically prior to the topicalizing function.

The historical development of the Finno-Ugric "person markers" indicates the importance of discourse notions, such as focus and topic, in the historical modification of the grammar of a language. Therefore, a morphological reconstruction has to consider a wider grammatical framework when establishing the semantic value of the reconstructed entities. We are dealing with a grammatical development which took place in the distant past, and with elements that are considered to be part of the oldest recoverable stage of the language. Conclusive evidence is therefore difficult to find. Some of the developments that are discussed in the context of Finno-Ugric languages may have counterparts in other language families, but the historical association of deictic elements with certain types of grammatical morphemes seems to be especially transparent in this particular language family.

2. Person Markers in the Finno-Ugric Languages.

2.1. General descriptions in the literature.

Proto-Finno-Ugric is traditionally assumed to have had one series of person markers that gave rise to three different sets of morphemes in the modern daughter languages--personal pronouns, possessive suffixes, and subject agreement markers in verbs.

The scholars who have reconstructed the person markers for the Proto-Finno-Ugric language have not agreed, however, about which grammatical category these morphemes represented in the proto-language. Finno-Ugrists have often claimed that the Proto-Finno-Ugric language had not developed a differentiation between nouns and verbs, but only had undifferentiated roots that could function as nouns or verbs (e.g., Hajdú 1975:78). There are, however, other scholars who argue that the finite verb-forms of the Uralic languages were originally nominal constructions, verbal nouns with suffixed person markers (Collinder 1960:243; Itkonen 1962:208).

The supporters of the first theory generally assume that the proto-language had one person marker category--personal pronouns. Personal pronouns were used in connection with nouns to mark possession, and in connection with verbs to mark the performer of the action, the agent. Hajdú, for example, claims that the proto-language could have alternate orders: Verb - Personal Pronoun, or Personal Pronoun - Verb. The latter word order was used when the pronoun had emphasis. In a later stage, the unemphasized personal pronouns lost their independent status, and became suffixes; possessive suffixes attached to nouns, and subject agreement markers attached to verbs (Hajdú 1975:85-87).

Serebrennikov (1973:72) explains the similarity of the person marker systems by their common derivation from demonstrative pronouns; according to him, both personal pronouns and possessive suffixes developed from demonstrative pronouns in Proto-Uralic.

Those scholars who interpret the verb forms in the proto-language as verbal nouns reconstruct the person marker category as possessive suffixes. Some Finno-Ugrists, for example E. Itkonen have presumed the reverse because the person markers in the verbal system are almost identical to the possessive suffixes, the proto-language must have a system of verbal nouns as finite forms of the verb (Itkonen 1962:208). It is suggested that personal pronouns developed under those conditions where the person marker was emphasized.

The actual processes through which the differentiation to three person marker categories took place have not, however, been explicated. There are references to the primitive thinking of the 'Urmensch' (e.g., Serebrennikov 1973:66) with the underlying implication that the present system with three person marker categories represents a step towards sophistication among the Finno-Ugric peoples.

If one wants to explicate the development and function of the person markers and their proto-forms, the starting point should be a context wider than only the three person marker categories in the modern languages. In order to explain the three similar reflexes in the modern languages, it is not, however, imperative to derive two of the categories from the third one. If we hypothesize that the proto-forms of the person markers were actually general deictic particles in Proto-Uralic--the position I will take in this paper--and not

specified as markers for the specific grammatical categories in the modern languages, we are able to understand the development of these elements as markers in a variety of grammatical categories in the daughter languages besides those associated with person. The deictic particles referred to roles in the communication act: first, anything connected with the speaker or in the proximity of the speaker ('speaker deixis'); second, anything connected with the addressee or his location ('addressee deixis'); and third, anything that is not connected with the speaker's or the addressee's location ('audience deixis') (I am using Fillmore's terms; Fillmore 1975).

The following section discusses the traditional reconstruction of the proto-forms associated with person markers on the basis of the modern reflexes. The reconstruction considers reflexes from nine daughter languages (see Tables 1-3). The Lappish dialects and Samoyed languages have been excluded even if they would give very useful information for the reconstruction. The dialect differences in these languages are very great, and the available material does not provide a coherent description of the deictic/person elements in any dialect. The information given would therefore be quite unreliable, being gathered from different dialects without any systematic analysis. The study of the person markers in the Samoyed languages might contribute significantly to the explanation of the original deictic elements, because these languages make a greater number of distinctions than the traditional three deictic categories. Therefore references are made throughout the paper to Samoyed forms even though they are not systematically discussed. Finnish and Estonian are used as representatives of the Balto-Finnic group. Finnish is considered a conservative language, "Finno-Ugric Sanskrit" (Anttila 1973:318) while Estonian is an innovative one. The other Balto-Finnic daughter languages fall between these two in respect to relative archaism.

A lengthy discussion on the reconstruction of the "person markers" is included in this paper because this information is not readily available in the handbooks. Authors have often noted that one proto-form can be reconstructed for all three reflexes, but detailed discussions of how the modern reflexes have developed from the proto-forms have not been presented. Some of the changes seem to represent universal or near-universal developments; others are based on language-specific developments. All of the details have not been discussed; the emphasis has been on those points where considerable controversy exists. Some suggestions have been given concerning the directions which further studies should take.

2.2. Functions of the person markers in modern languages.

To clarify the basic functions of the person markers in the modern languages, an example is given from Modern Finnish:

	<u>mennä/mene-</u> 'to go'	<u>koti</u> 'home'	<u>personal pronouns</u>
Sg. 1	<u>mene-n</u> 'I go'	<u>koti-ni</u> 'my home'	<u>mi-nä</u> 'I'
2	<u>mene-t</u> etc.	<u>koti-si</u> etc.	<u>si-nä</u> etc.
3	<u>mene-e</u> (vowel length)	<u>koti-nsa</u>	<u>hän/se</u>
Pl. 1	<u>mene-mme</u>	<u>koti-mme</u>	<u>me</u>
2	<u>mene-tte</u>	<u>koti-nne</u>	<u>te</u>
3	<u>mene-vät</u>	<u>koti-nsa</u>	<u>he/ne</u>

In modern Finnish, the subject pronouns of the first and second persons, singular or plural, do not have to be overtly expressed; the person is marked in the verb. In the third person, however, the pronoun must be expressed. For example:

Löysi-n kirja-ni 'I found my book'
Löysi-t-kö kirja-ni? 'Did you find my book?'
 *Löysi kirja-nsa 'He found his book'

This is the situation in Standard Finnish, but in dialects (or in colloquial speech in general) where the agreement markers have disappeared from the verbs, the subject is marked overtly by the pronoun in all persons.

Vértes (1967) reports from Ob-Ugric languages that the use of the personal pronoun is sometimes stylistically conditioned. There seems to be, however, a general tendency that in coordinate structures, the first sentence has a personal pronoun, but the second does not (in her examples [pp. 16-17], both sentences have the subject in the same person). If the sentence structure is Subordinate Clause-Main Clause, the subordinate clause does not have a pronoun, but the main clause does. The subordinate clause maintains the older state; the main clause innovates.

The possessive suffixes, whose main function is the same as possessive pronouns in English, have developed various other functions in some daughter languages. In Permian languages (i.e., Zyrian and Votyak), Cheremis, Ob-Ugric (i.e., Vogul and Ostyak), as well as in Samoyed languages, the second and third person singular possessive suffixes can be used as definite articles (in Cheremis and in Permian languages, even the first person suffix can be used for the same function) (Collinder 1960:204). In a later section of the paper, I will discuss the significance of this usage for the proposed analysis.

In Finnish, genuine possession is actually indicated in the third person only in cases where the genitive of the personal pronoun is used, for example, Se on hänen hattunsa 'It is his hat'. One can say Hän otti hattunsa 'He took his hat', but not *Se on hattunsa 'It is his hat'. In the sentence Hän otti hattunsa 'He took his hat', the third person singular possessive suffix -nsa acts much like a definite article used reflexively. In Lappish, the possessive suffixes are chiefly reflexive possessive pronouns:

gawdni-m gâpperâ-m 'I found my cap', but gawdni-m dú gâpperâ
found-I cap-my found-I your cap
'I found your cap' (Collinder 1960:203).⁴

The possessive suffixes do not generally occur with subjects in Lappish, mú akso lä lappum 'my axe is lost' (instead of akso-m).

In the Balto-Finnic group, there is a general tendency to lose the possessive suffixes; Estonian and Votic have lost them completely, and other languages have limited their use. Finnish represents the archaic state, but as the colloquial language indicates, the possessive suffixes are gradually disappearing. Interestingly, those forms where the possessive suffixes have survived the longest, for example in Estonian, are vocative forms (Mark 1925).

The modern Balto-Finnic languages have developed possessive pronouns which are genitive case forms of the personal pronouns, e.g., Finnish minä/ minu- 'I', minu-n 'my' (genitive case ending n). Serebrennikov (1973) regards this as the younger formation type. The older formation of possessive pronouns in Finno-Ugric languages consists of the stems of the personal pronouns with attached possessive suffixes, e.g., Erza dialect of Mordva sonze 'his/her/its', (son 'he/she/it' and -ze 'possessive suffix, 3rd singular'). The Ugric branch uses other ways to express possession. At this point let it suffice to say that the category of possessive pronouns is a secondary development in Finno-Ugric languages, and represents a relatively young formation.

2.3. Reconstruction and historical development of the person markers.

The handbooks on Finno-Ugric languages treat the reconstruction of the proto-forms of the person markers as established. Consequently, they do not provide a thorough discussion of the development of the present-day reflexes. Mark (1925) discusses the development of possessive suffixes in the Balto-Finnic languages, but a great deal of his information is controversial, and he overlooks a number of central issues. This section of the paper treats those issues that must be investigated in determining the proto-forms of the person markers. The discussion is based on the data provided in the handbooks and in those articles that were available during the preparation of this paper. Many details have been omitted because of a lack of complete information. Finnish developments have been treated in more detail, because more source information is available about Finnish and because Finnish evidence may be particularly significant considering the archaic character of the language.

2.3.1. Shape of the proto-forms.

Most scholars agree that there existed two basic syllable types in Proto-Finno-Ugric: V and CV. Most roots were bisyllabic of the type CVCV; only the pronominal roots have been reconstructed as monosyllabic, CV (in some cases V).

The investigation of the reconstructed "person markers" indicates that these might not have been "roots" in the proto-language, but enclitic particles of the type CV. As will be seen later in greater detail, the independent pronouns, both personal pronouns and demonstrative pronouns, were formed from these CV type clitics by attaching them to a 'neutral base' (cf. Forchheimer 1953:8). 'Neutral' means neutral as to person, the deictic particle indicating person in these pronouns (see section 2.3.4 for further discussion).

As becomes obvious from Tables 1-3, the modern reflexes of the prehistorical morphemes are not always CV or C (with the loss of the final vowel), but some forms are of the type VC or VCV. What has happened in the history of some languages, especially in Hungarian and Ostyak, is that word-final vowels are lost. Therefore, there was a morphophonemic alternation in the stem morphemes: the final vowels were preserved in those cases where a suffix (e.g., a person marker) followed. In the course of the time, the vowel came to be reinterpreted as a part of the person marker, and certain vowels came to be associated with certain person markers. Certain suffixes combine with the verb stems as a result of vowel contraction. (The handbooks interpret these cases as a preservation of the stem vowels in the cases where there would be a consonant cluster.)

2.3.2. Vowels

As tables 1-3 indicate, the vowels in the present day person markers vary greatly, and the whole range of vowels (a, e, i, o, u, ü, ä, ö) can be found. One of the problems for the reconstruction of proto-forms is whether one should reconstruct one vowel for all persons, and if so, which one.

Most Finno-Ugrists have established the vowel as being [+Front] without specifying its exact quality. The basic reason for this decision may have been that the archaic member of the language family, Finnish, has a front vowel in these morphemes, either e or i, depending on the form. (Note that the third person subject agreement markers are of secondary origin, as will be discussed later.) The other reflexes have been generally considered vowel harmony variations in the present day language, not derived from the proto-forms through other phonological changes.

In my opinion, we should consider three major sources for these vowel alternations: vowel harmony either in the proto-language or in the individual daughter language, dialectal variations in the proto-language; different reflexes going back to different proto-dialects; and the possibility that if we reconstruct these elements as originally deictic particles, the proto-language may have had more distinctions based on the proximity or distance of the participants in the communication situation.

On the basis of comparative evidence vowel harmony can be reconstructed for the proto-language. It has continued more or less

intact in all of the daughter languages except for the Lappish dialects, Estonian, and Zyrian and Votyak (Hajdú 1975:94). The vowel harmony in the proto-language is reconstructed as the same type as in modern Finnish, i.e., front-back vowel harmony. Hungarian and some other languages (Hajdú mentions, e.g., the eastern dialect of Cheremis and Selkup) have acquired an additional distinction, namely the rounded-unrounded vowel harmony. When deciding which vowels to reconstruct for the proto-forms of the person markers/deictic particles one has to consider the status of these elements in the language. Were they independent words or were they suffixes? Some evidence has been presented that indicates the status of the person markers as independent words because the initial consonants have undergone the changes of word-initial consonants, not word-internal (Barczi 1963). If the elements were independent words, they would not participate in vowel harmony, and therefore they would have either had only one shape, one allomorph, in the proto-language, the present reflexes being due to different proto-forms, either dialectal variations or different deictic particles. It seems reasonable to assume that the elements were not suffixes but enclitic particles in the proto-language, which explains the development of the initial consonants as word-initial segments. Because the daughter languages can be shown to have inherited vowel harmony from the proto-language, the present vowel harmony alternations cannot be used to argue for the status of person markers/deictic particles as suffixes in the proto-language. The suffixation can represent a parallel development in the daughter languages.

The reason the status of the "person markers" as suffixes or as independent words has caused discussion is the fact that the Proto-Finno-Ugric language has been assumed to have had a restriction with respect to which vowels could occur in a word outside the first syllable. All the reconstructed vowels could occur in the first syllable, but most scholars agree that the non-first syllables could have only a, e, or ä (ä being either low or mid lax front vowel). Hajdú (1972) for example assumes that if the first vowel in the word was palatal, then the following vowel had to be either e or ä ([ϵ]); if the vowel was [+Back], then the following vowel had to be a or e. According to Hajdu, e was a neutral vowel in the same way as in modern Finnish (in Finnish e and i act as 'neutral vowels' with respect to vowel harmony). Hajdu reconstructs therefore e.g., the first person singular marker as *me~ma. If, however, the person marker/deictic elements can be reconstructed for the proto-language as independent words, not as suffixes, then the restrictions on the second syllable (or non-first syllable) vowel cannot be used as a basis for determining the vowel of the proto-form.

One cannot exclude the possibility that the proto-language had dialectal variation in the vowel quality of the deictic particles and that the daughter languages generalized one or the other dialect variation. Obviously this variation cannot be reconstructed.

A third possible source of vowel variations in the person markers/deictic particles might be an original distinction in the semantic function of the deictic elements, i.e., the proto-language may have made more distinctions as to the respective location of the speaker and the action/objects. Some evidence from the present day languages is given in the section on demonstrative pronouns, but on the basis of the material I have used for this paper, I cannot convincingly argue for the vowel variations in the proto-language which would be explainable only by additional deictic distinctions.

In my opinion, a further study should be carried out to investigate the development of person markers as well as demonstratives in those Finno-Ugric languages where more than three distinctions are made in the present day demonstratives such as Lappish and Samoyed (Tauli 1966:141).

Whatever the source of the vowel variation in the person markers/deictic elements, the individual languages have generalized certain vowels to certain functions. For example in Votyak where the plural marker was lost because of a phonological change, the vowel *y* [ü] was interpreted as a plural marker in the possessive suffixes. Mordva has the mid back vowel *o* in the singular form, but high front *i* in the plural forms (see Tables 1-3). Finnish has *i* in the first and second person singular but *e* in the corresponding plural forms.

On the basis of the above considerations I have not been able to come to any positive conclusion with respect to the reconstructed vowel segments; the traditionally reconstructed [+Front] vowel seems to be a viable solution as long as there is not enough contradictory evidence; the Finnish data would indicate a mid front vowel for the proto-form.

2.3.3. Consonants

The evidence indicates that the consonantal segments in the person markers or deictic particles can be reconstructed as **m* (first person, 'speaker deixis'); **t* (second person, 'addressee deixis'), and **s* (third person, 'audience deixis'), but it appears necessary to re-examine some of the suggested processes that have resulted in the modern reflexes.

In this section I am going to comment on the development of those reflexes that have undergone changes of the consonant segment. This is not intended to be an exhaustive analysis of the historical changes, but a brief survey of the types of changes that have resulted in the present consonant segments in the person markers. Special attention is focused on the comments in the literature about the role of an '*n*-affix' in the development of person markers.

If we look at the forms of the first person markers in Table 1, we notice very similar reflexes in all daughter languages. We can reconstruct the initial consonant as **m*. In the Balto-Finnic languages,

the word-final vowel has been lost in the subject agreement markers, and the word-final consonant *m became n. A similar development is to be seen in Mordva. The problematic case is the reflex ni in these languages. Why -ni and not -mi? The first person plural form has the m-consonant. The occurrence of n in this form has been explained by the influence of an n-affix that was attached to the noun stem at an earlier stage of the language, and which preceded the actual person marker under certain conditions. Scholars disagree about what these conditions were. The existence of an n-affix in forms which contain person markers would explain several other modern reflexes in various daughter languages. Finnish scholars in particular argue for its existence. In the following I will summarize some of the arguments that have been presented in the literature.

Mark (1925) supposes that there was an n-element preceding the possessive suffix in all oblique cases (in all cases except nominative singular). In the genitive and accusative case the n was supposed to be the case ending; in the nominative plural, the suffix which indicated the plurality of the possessions. Mark explains the use of this affix in other singular cases but genitive and accusative, as an analogy from these two cases which had an n.

The interpretation of the n-affix as a marker of plurality has been accepted by many scholars, because that would support the reconstruction of *-n as a plural marker. The fact that the daughter languages show pairs such as Finnish tämä 'this' - nämä 'these'; tuo 'that' - nuo 'those' where the plurality is indicated in the segment n, has led several Finno-Ugrists to assume an *-n segment in proto-Finno-Ugric as a plural marker (e.g. Collinder 1965:130; Hakulinen 1957:60; Szinnyei 1922:52). As it is shown in a later section (3.3), the demonstratives do not support this type of reconstruction for the proto-language.

Szinnyei (Mark 1925:49) assumes that proto-Finno-Ugric had two series of possessive suffixes, one which marked singular possession ('my house'), the other which marked plural possession ('my houses'). This plural series had, according to Szinnyei, an n-affix. It is not evident, however, why this n would be generalized into singular forms.

Hakulinen (1957:73) agrees with this type of explanation, pointing out that there was a basis for the mixture because several cases ended in n (genitive, accusative, instructive, illative, and allative). The singular and plural were not distinguished in these case forms.

According to Hakulinen, the Proto-Balto-Finnic forms of the possessive suffixes were as follows:

singular possession

1. *-mi e.g. 'my house'
2. *-ti~ð̃i
3. *-sen~zen

plural possession

1. *-nmi > *-nni e.g. 'my houses'
2. *-nti
3. *-nsen

Plural

- | | |
|--------------------------------|---|
| 1. *-mek~mmek e.g. 'our house' | 1. *-nmek > *-nnek ~ *-mmek e.g. 'our houses' |
| 2. *-tek~ð̃ek | 2. *-ndek |
| 3. *-sek~zek | 3. *-nsek |

The Iitti dialect of Finnish has been presented as an example of the archaic state of Balto-Finnic (e.g., Wickman (1955:19)). The first person singular form 'my house' is tupam (tupa-m), but the accusative singular form and the nominative plural is tuvain (< *tuβa-n-me). Because of the consonant gradation operating on p, the nominative singular originally had a strong grade p because it occurred in the beginning of an open syllable: *tupa-mi 'my house'. The final vowel was lost, and it gave the form tupam. The accusative form had *tuβan-mi where the affix n closed the syllable, and resulted in the weak grade y from β.⁵ The weak grade form can only be explained by the occurrence of a syllable-closing consonant which was subsequently lost. Therefore, the first person singular possessive suffix in Balto-Finnic languages is assumed to have developed from *n-mi > *n-ni, finally becoming ni.

There is a problem with the n-affix that has not been explicated in the literature. Do we have to assume an n-affix for proto-Finno-Ugric or is it only a Proto-Balto-Finnic phenomenon? If we reconstruct an n-element for the proto-Finno-Ugric we have to be able to explain its function. The reconstruction of *-m accusative or *-n genitive for proto-Finno-Ugric is quite controversial, as is the reconstruction of an -n type plural. These are just the grammatical elements that have been used as an argument for the existence of n-affix in the proto-language. Unfortunately, at this stage of research the handbooks abound with the following type of statements: "One might suppose that the...n is ultimately identical with the genitive ending. This would imply either that the possessive suffixes give evidence of an early stage of Proto-Uralic [PU] when there were not yet local cases, or that the local cases...have got their co-affix through analogical influence from the genitive" (Collinder 1960:302). The sound correspondences in the possessive suffixes would be easier to explain if one could propose the existence of such a coaffix n, not only in the Balto-Finnic languages, but in the whole Finno-Ugric family. The function of this coaffix still remains to be established. That requires a reconstruction of the case system for the proto-language, one that is not based on the preconceptions of the scholars, but on firm linguistic theory.

As a last suggestion about the origin of the n-affix, I would like to revive M. Castrén's old hypothesis (cited in Mark 1925) according to which the n-affix was "somehow connected with the first person pronoun *-m[V]." This hypothesis sounded quite far-fetched to Mark, but if we can show that *mV represented a general deictic particle in the proto-language, and that it could serve several functions, this hypothesis is not absurd at all. I cannot in this paper argue for a definite function for an -n type affix, nor even argue with surety for its existence, but the comparative evidence certainly allows the possibility that there was an affix *-n in the Proto-Finno-Ugric language which was of deictic origin. It remains to be established if it can be derived from an older form *-mV or if it represents an additional deictic element at an earlier stage.

There is an additional factor to be taken into consideration. Similar phonological changes have taken place in both possessive suffixes and in subject agreement markers in verbs. Those scholars who maintain that the -n affix was a case marker, have to propose that verb forms were actually nouns, taking case suffixes, or find another explanation for the identical reflexes. If the n element had had its origin as a deictic particle, as it will be argued for other "person markers," both occurrences of n-influence would be accounted for. But this is just a vague suggestion; at this point I am not committed to any particular hypothesis of the origin of the n-affix.

Now I would like to return to the discussion of the reflexes in Table 1. Besides the Balto-Finnic possessive suffix -ni, the suggested existence of an n-affix conditioned the sound changes in Mordva and Cheremis subject agreement markers: nok-nuk, etc. (Mordva), and na-nä (Cheremis) (*nm *nn>n).

The first person plural subject agreement marker has lost the nasal element in Hungarian. The earliest documents from the 10th century have *mk forms (Károly:126). The j segment is a result of a levelling phenomenon; it will be discussed in the context of third person markers.

In Vogul and Ostyak, the *m segment has been denasalized in the plural forms: w occurs both in possessive suffixes and subject agreement markers, but the personal pronouns have the m element, as do the singular forms.

The second person shows more variety (Table 2). The proto-segment *t has become s in the Balto-Finnic languages, in the possessive suffix (singular), and in the personal pronoun (singular). This is a result of a regular sound change whereby *t > s / - Vowel [+Front, +High]. The Finnish second plural possessive suffix has been accounted for by the influence of the n-affix at an earlier stage: *nt > nn.

The various d reflexes in the second person are results of regular sound changes: *t > d / V - V. The Hungarian reflex d has

been used as an evidence for the position that the person markers were not suffixes in the proto-language, but independent words (e.g., Barczi 1963). The regular correspondence for Proto-Finno-Ugric *t in Hungarian is z (i.e., *t > z / V - V). Barczi explains the d reflex as due to the fact that t had become an accusative marker, and would have caused homonymy if the possessive suffix had remained unchanged. In the plural where this problem did not arise, the *t remained. Obviously there must have been other (phonological) factors that caused the change *t > d in the singular. The Finnish scholars, as well as Collinder, explain the Hungarian d reflex as a result of the preceding n: PFU *nti > *nt > d (Collinder 1960). Good historical method demands that conditions be established why the *nt cluster became d in the singular, and not in the plural, before it can be assumed that this particular change was due to the influence of the n-affix. As it has become obvious, a cross-language study is needed of the role or non-role of the n-affix in the history of person markers. The reader of the current literature on this topic gets the impression that the role of n-affix is used as a device to "explain" present reflexes without establishing the relevant conditions, and wonders, e.g., why it plays a role in some forms but not in others. "Mixed paradigms" as an explanation is not very satisfactory.

Ob-Ugric languages (Vogul and Ostyak) and Zyrian have an unexpected reflex in the second person subject agreement morpheme: n. Vogul and Ostyak have n in all person marker categories in this person. Two types of explanations have been given in the literature:

Szinnyei (1922) assumed that PFU or PU had a regular consonant alternation t~n, and these languages which show the n reflex have generalized the n alternant to all phonological environments. It has not been established, however, that the proto-language had any alternation of this type. This proposal seems to be quite arbitrary with respect to the reconstructed proto-language.

Vértes (1967) presents the opinion that n was introduced to the paradigm because phonological change in these languages had caused PFU *s to merge with *t, and thus the second and third person markers fell together. This does not, however, explain why the n was chosen as a new marker for the second person. Where did it come from? Is it possible that it was due to Altaic influence or can we explain it as an internal language change? If we could assume that the proto-language had another deictic element *n(V), the n reflexes in Vogul, Ostyak, Zyrian as well as Samoyed languages could be derived from this element directly. Scholars have suggested that the existence of the deictic element *n would explain developments of demonstrative pronouns (see section 3.3 for further discussion). Future research may clarify the history of the second person reflex -n, and its relationship with n-element in the demonstrative pronouns, as well as its possible connection with the coaffix -n.

The third person reflexes differ drastically from those for the first and second persons, because they do not seem to be derivable through phonological changes from the proto-segment *s. There is much

more variation across languages, and even inside one language, between the possessive suffix, personal pronoun and the subject agreement marker, as Table 3 indicates. Reasons for this variation are discussed in more detail in a later section (2.3.5). In this section I will comment on some regular sound correspondences of PFU *s.

PFU *s has an s reflex in initial position in Finnish, Mordva, and Permian languages (Zyrian, Votyak) and a z reflex intervocally, except in Finnish where it always appears as s. In Cheremis, *s occurs as s in initial position, ž intervocally. In Hungarian, initial s was lost, and because the stems were of the shape CVCV, a glide developed as a transition between the stem-final vowel and the vowel of the enclitic particle (by that time probably a suffix). In the third person, j results from a straightforward sound change, but j was later generalized to other persons; compare the first person plural. In Vogul, PFU *s has a t reflex in all positions. Ostyak has different reflexes depending on the dialect (Southern Ostyak has t; Northern, l; Eastern, j; λ (lateral spirant) and č also occur in some dialects. The third person personal pronoun occurs therefore as teu/luw/yoγ/λou and variations of these depending on the dialect group. (Vértes 1967 has a detailed listing.) These reflexes are repeated here because the handbooks sometimes give confusing information about the sound correspondences. I do not know how much dialect mixture is involved in the person marker systems in Ostyak, but at least we can distinguish the regular correspondences in each individual dialect.

2.3.4. Personal pronouns

In my opinion, the main function of the deictic particles in the proto-language was to mark the focus of the sentence. The particles could be attached to verbs or to nominal elements, or they could occur at the end of the whole sentence. At an earlier stage, pronominal subjects were not overtly expressed. In an SOV language (as PFU has been reconstructed), the agent or subject was indicated in the verb. In my analysis, the enclitic deictic particle attached to the verb served this function. Finno-Ugrists have pointed out that the personal pronouns were expressed only for emphasis. That would mean that in those cases where the agent, the person performing the action indicated by the verb, was the focus of the sentence, it was expressed by a corresponding deictic element. In the subject position, the "person markers" had the shape CVCV. In the course of the historical development, those deictic particles which had become suffixes were reinterpreted as inflectional affixes, as person markers, and those in subject position as personal pronouns. In the further development of the languages, the subject agreement markers disappeared through phonological changes, and the use of the personal pronouns/deictic particles became obligatory even in those cases when the agent was not emphasized. This change seems to be connected with the word order change in these languages; the 'western' languages in this family have become SVO.⁶ In Finnish, for example, the first and second person can occur without an overt personal pronoun, but this no longer

applies to colloquial language. Personal pronouns have become almost obligatory. In literary Finnish the subject agreement markers remain unchanged; in colloquial language they have largely disappeared.⁷

In his study The Category of Person in Language (1953), Forchheimer came to the conclusion that the affixed pronoun forms universally represented the pure pronominal elements, whereas the independent pronoun forms were formed by a base neutral with regard to person, attached to a pronominal affix. Finno-Ugric languages seem to follow this pattern, both in the formation of independent personal pronouns and in the formation of independent demonstrative pronouns.

The personal pronouns are formed by attaching the element n (+Vowel) to the "person marker" (deictic particle). This applies to the first and second person pronouns, and some third person pronouns as well, for example, Finnish mi-nä 'I', si-nä 'you'. In many languages the final vowel has been lost through a regular sound change, and only the final n is left from this suffix, e.g., Mordva mo-n 'I', to-n 'you'. It is quite probable that originally there were two ways of forming personal pronouns, either with the deictic particle and the suffix *na-nä, or with the deictic particle alone. This situation still prevails in Estonian where the personal pronouns have double forms mi-na ~ ma 'I', si-na ~ sa 'you' (Raun-Saareste 1965).

The origin of this *n(a) element is a somewhat controversial issue. Serebrennikov (1973) indicates that its meaning is not clear. Hakulinen (1957) claims that it is a deictic element meaning 'I here'. The same element is to be found in the locative ending na-nä, which is assumed to be one of the oldest case suffixes in FU languages. The same opinion is represented by Munkácsi (quoted by Vertes 1967) who also argued that n in the personal pronouns was a demonstrative pronoun or a deictic particle meaning 'I here', 'you there'. Hajdú (1975) suggests that this locative suffix na-nä, and also another locative suffix in PFU *-tt (or *-ta), may have been derived from demonstratives. The demonstrative attached to a noun had a variety of adverbial functions, and the suffix gradually developed a more specialized use--in this case, as a locative marker. I will return to this suffix in the section on demonstratives (3.3.).

Some languages have a different element attached to the deictic element, not *na-nä. Cheremis for example has -jo ~ jö: to-jo/töjö 'you', or Zyrian si-je 'he/she/it'. This same suffix can occur in the demonstratives. I would assume that it is of the same origin as jo- in Finnish, for example, in the pronoun joka 'which' (relative). The Samoyed languages show further patterns: in the Enets dialect, the personal pronouns are formed by attaching -či < *ti to the deictic particle: mo-či 'I', to-či 'you'.

The main pattern seems to consist of the deictic element and an additional CV sequence which is either to be analyzed as a locative suffix or a demonstrative element. It appears that the deictic or

demonstrative elements might have had a great role in the development of case markings and--as I attempt to show in this paper--other grammatical morphemes. This development was probably underway at the time of common development, in the system that is reconstructed as Proto-Finno-Ugric, but the same elements were still used for other functions as well. It is therefore possible that the Proto-Finno-Ugric language had the elements **na-nä* and **ta-tä* as demonstratives, but these elements were also used for the specific locative function. There are a great number of similar developments in pidgins and creoles where at some point in their development, a single grammatical morpheme can serve several functions.⁸ The Cheremis and Zyrian forms (with the *-jo/je* element) seem to represent language-specific developments, but they follow the basic pattern in the sense that the pronominal element (*-jo* etc.) is added to the personal/deictic element to form personal pronouns.

Some scholars have analyzed forms such as Enets *toči* 'you' as a reduplication. This would be a natural way to emphasize the deictic element, but the analysis is made difficult by the fact that the locative suffix had the same initial consonant **t* (which became *č / V - V*) as the corresponding deictic element. They would obviously consider forms such as *toči* as original, and *moči* as a generalization of the pattern. The analysis of **na-nä* and **ta-tä* as locatives (whether derived from demonstratives or not) gives a more consistent picture of the patterns of formation in both the categories of personal pronoun and demonstrative pronoun. Reduplication in some cases might have been a way of forming an emphasized form of a pronoun, but it was probably not the basic pattern of pronoun formation.

The Hungarian first person singular pronoun differs from the general pattern: *én* 'I'. Two explanations have been offered: it is derived from **e-me-n* where *e* represents a demonstrative root, and *me* the old Uralic first person marker (*n* would obviously be of the same origin as in the other pronouns). The second way to account for this form would be that it is a secondary Hungarian development where the demonstrative pronoun **e* has been suffixed by the pronominal suffix *-n* (Szinnyei 1922, Collinder 1960).

2.3.5. Third person

The special status of the third person has often been pointed out by linguists. Lyons (1971:276) comments on the fact that the third person has to be distinguished from the first and second persons in several respects. The speaker and hearer are necessarily present in the situation, whereas other persons and things to which reference is made may not only be absent from the situation, but they may also be left unidentified.

In a certain sense, the third person is the "primary" category; it is the unmarked category, and can only be called a "person" with

reference to the first and second persons. Many languages have no "personal pronouns" for the third person. Either the person is completely unmarked or a demonstrative pronoun is used for this function. Estonian and Cheremis in the FU family are examples of the use of demonstratives in the third person.⁹

The elements referring to the speaker and the addressee are inherently definite. But the third person may also be identified only by using the zero marker, to distinguish it from the first and second person, but not by marking it definite.

Hajdú (1972) describes the function of the third person subject agreement marker *se, which is different from *me and *te. The first person suffix *me and the second person suffix *te referred to the subject of the verb. In the third person, there was no need to specially indicate the subject. The zero suffix alone indicated that it did not refer to the speaker or the addressee. Therefore, whenever *se appeared, it did not refer to the subject of the verb, but indicated an indirect reference to the object of the verb. It was used to mark the definite object of the verb. Hajdú calls this attachment of *se "a pronoun with the value of an accusative" (p. 44).

If we interpret the grammatical elements under discussion as deictic particles which came to be used as focus markers, the differing patterns in the third person can be analyzed in a systematic way. The present Finnish system includes an interesting type of a sentence which has been called an 'impersonal' or a 'generic' sentence. (See Hakulinen and Karttunen 1973 for a detailed synchronic analysis of this type of sentences.) These sentences correspond to English sentences which have the impersonal 'you' or 'one' as the subject. Such sentences are still very common in Finnish. The subject constituent is missing in these sentences; the verb is in the third person singular. It is possible that this particular type of sentence corresponds to the PFU sentence type where no deictic element was attached to the VP constituent. It was a general statement that did not refer to any location in the situation of the utterance.

Based on the information in the handbooks, the reconstruction of tense markers for Proto-Finno-Ugric seems to be problematic. Although Serebrennikov's argument is speculative to say the least, he seems to be right in assuming that PFU did not have tense marking, but that it marked aspects instead (Serebrennikov 1973). One can reconstruct elements that marked continuous, iterative, accomplished, etc., action.¹⁰ When there was no deictic element attached to the VP (i.e., when the utterance was just a general statement, not defined in terms of the speech situation), the verb form consisted of the verb stem plus the aspect marker. After the person markers had developed as a category, the aspect markers were analyzed as person markers in those cases where it was felt that the person marker was "missing" (i.e., in those cases where there was no deictic particle). This led to two major trends--the development of subjective conjugation markers in some languages, and in others the development of the third

person subject agreement marker from aspect markers. There are languages, such as Zyrian and Votyak, which do not have two conjugation systems but still have two forms in the third person: one with the reflex of the deictic particle *sV and the other without it. In Zyrian the form without the deictic particle came to be interpreted as the present tense, while the form with the deictic particle was reinterpreted as the future tense.

The subjective/objective conjugation distinction is discussed in a later section. At this point I would like to discuss the development of the third person markers in the Balto-Finnic languages. As Table 3 indicates, these languages show the most divergent reflexes in the third person. I will first discuss the third person subject agreement markers, and then the personal pronouns.

The details have not been explicated, but it has been suggested that the Balto-Finnic *pa-pä which gave rise to both the third person singular subject agreement marker and to the present participle was originally used to mark a continuous, progressive action. In the present tense third person singular *a ~ ä > i /-- #, while the participles still have -a-ä in Finnish. Anttila (1972:351) calls it a grammatical conditioning of a sound change which took place in the predicate verbs.

There was another alternation at an earlier stage of Finnish: p ~ β. After a stressed syllable, the bilabial stop p occurred; after an unstressed syllable, the bilabial fricative β. This alternation occurred in both the third person singular present tense and in the participles. There are attested forms from Old Finnish (16th century) in which this alternation can be seen: sööpi 'eats'; kúmartäpi 'bows'; ántäβi 'gives'. The primary stress is on the first syllable, a secondary stress on the third (Ruoppila 1967:47). As a result of levelling, va-vä (in the participles) and vi (in the present tense third person) were generalized whatever the stress situation obtained. The texts also give examples of the next stages in the development of the third person forms: ottavi 'takes' becomes ottav and further ottau and ottaa, which is the current form in Finnish. Through this process, the lengthening of the stem-final vowel became the third person singular marker.

Estonian has the third person singular present tense marker b which developed from *pa-pä. The participle marker is v: tulev 'coming' (present participle); tuleb '(he) comes'. Here the different alternations were assigned to different functions; in Finnish, where the final vowels were retained, the vowels carry the distinction.

The third person plural forms have the noun plural suffix t attached to the singular forms. The plurals had the weak grade because t closed the syllable: βat-βät which became vat-vät. Estonian has kept the alternation: palub '(he) asks', but paluvad '(they) ask'.

Finnish has generalized this suffix, which originally occurred only in the present tense to the preterite as well; -vat ~ vät was interpreted as the third person plural marker. Finnish kysyy '(he) asks'; kysyvät '(they) ask'; kysyi '(he) asked'; kysyivät '(they) asked'. Estonian has palus '(he) asked' and palusid '(they) asked'. In Finnish the generalization of the third person plural present tense ending to the past tense is a very late phenomenon: 18th century literature still has he söit ja joit 'they ate and drank' (Ruoppila 1967).¹¹

As it was previously indicated, the PFU verb forms that had the deictic particle *sV attached to them developed differently in the daughter languages. In some cases the deictic element developed specialized functions because it was not required to mark the speaker or the addressee. In other cases it developed into a definite object marker in objective conjugation, a future tense marker in the third person, or it could become a regular third person singular marker, as *mV and *tV had developed into the first person and second person markers. At first sight, Finnish does not seem to have any reflex of this *sV particle as a verbal suffix, but in a further analysis paraphrases of the following type are found:

- a. Sen työn tekee helposti (impersonal, generic sentence)
that job does easily 'that job one does easily'
(acc.)
- b. Se työ tehdään helposti (passive sentence)
(nom.) 'that job is done easily'

The form (b) represents the "impersonal" passive in Finnish. Historically it is derived from *teke-tä-sän (verb stem teke-; causative suffix -tä; 3rd singular 'person' marker -sän). It is therefore possible that the PFU *sV deictic element developed into an "impersonal" passive suffix in Finnish. At this moment, I am unable to present historical evidence about the passive formation in other FU languages.

Some examples from doublets in Southern Estonian dialects where the third person singular can occur without any person marker are: jaga 'distributes'; and 'gives'; lät 'goes'; näge 'sees', or with a suffix -s: eläs 'lives'; kaes 'looks'; küzüs 'asks'. Posti (1963) has characterized the differences between these two groups as being due to the semantics of these verbs. *sen occurs with those verbs which refer to an action by which the subject has a personal involvement, e.g., kaes 'looks at' compared with näge 'sees'. The indicated action refers to the subject of the sentence. This corresponds to the meaning of the mediopassive in IE languages. Posti considers this Estonian situation to be the original one; other Balto-Finnic languages have developed passives from these forms, others such as Karelian developed reflexive forms.

The function of the -n in passive forms is problematic. It has not been fully explained. Hakulinen (1957:174) suggests that

the -n could be a dual marker that occurs in the suffixes, but why would it only occur in this one form? He does not bring any evidence from other languages. It could be of the same origin as the -n in personal pronouns. We have to determine, however, the historical development of these forms in greater detail.

If we look at the personal pronouns in the third person in Finnish, we find that the corresponding forms are hän/se. Finnish is the only Finno-Ugric language that differentiates human/non-human in the third person: hän 'he/she'; se 'it'. In colloquial speech, se is used for both functions.

Posti (1953:61) discusses the s-h alternation in Balto-Finnic. According to him, there was an s ~ z alternation in Baltic Finnic corresponding to the k ~ γ, t ~ č, p ~ β alternation (cf. p. 99). *z occurred in the beginning of a closed syllable intervocalically, and also after any unstressed vowel. *z became h by the end of the Late Balto-Finnic period. After the change *z > h, an s ~ h alternation arose, but at this stage levelling occurred in most paradigms because the speakers did not feel these sounds closely related. There are some relic forms of this alternation in Finnish: mies 'man' alternating with mieh-en 'of the man' (genitive).

The problem is how to explain the h in the independent position, since the s ~ z alternation only took place intervocalically. Posti's explanation is that it occurred because of the frequent use of the pronoun in unstressed position where the initial s became h in the same manner as the medial s became h after an unstressed vowel.

I would like to suggest that there was reinforcement from Swedish: Swedish has han/hon 'he/she', and it is possible that the strong Swedish influence reinforced the use of han for persons in the situation where the form was already in the system even if it did not occur independently in this form. The form -han~-hän occurred in the verbal paradigm as a marker of the impersonal passive. In the 18th century there were still forms anneta-han 'is given' (<*anneta-san); the present form is annetaan).

Vértes (1967) suggests that han was introduced because of the phonological change *t > s / - i, and *tinä had become sinä; therefore the third person *se/si- fell together with the second person. She assumes that -n in the modern form (hän) is the same -na-nä suffix as in the first and second person, the final vowel having been deleted.

There is one more problem that does not appear to be discussed in the literature: why does hän have an -n, but not se? Other FU languages have -n in the third person pronouns, for example, Mordva son. If we interpret the -n as a part of the local suffix that was used to emphasize the independent pronouns, we might conjecture that the third person deictic element differed from the first and second in that it could also be used attributively, e.g., se poika 'that boy'. It may be possible that the attributively used forms did not

have the *-n+Vowel suffix. The same phenomenon occurs in the demonstrative pronouns where tämä 'this' has the structure CVCV, but tuo 'that' does not have any additional suffix. Tuo may be a generalization from the attributive form.

2.3.6. Plural and dual forms of the person markers

The plural forms of the person markers/deictic particles can be reconstructed as having been formed by attaching the suffix *-k to the corresponding singular forms. This plural marker is found in the surface forms in Mordva, Ostyak and Hungarian, as well as in dialectal forms in Livonian. Several other reflexes can be explained in various languages by postulating this *-k element in the plural forms. For example, in Balto-Finnic, the vowel in the deictic element became [+High] in word-final position, but in the plural forms it did not change because of the final -k; therefore the singular form has mi-, but the plural me-. After the loss of the -k element, several languages morphologized a vowel alternation as a marker of the plural, for example Mordva has the singular forms, mon, ton, son 'I, you he/she/it', but the plural forms min, tin, sin 'we, you, they'.

Some scholars have tried to establish the origin of the plural marker *-k. They suggest either that it was derived from a dual marker *-ka which was derived from the numeral *kakte 'two', or that the plural marker is of the same origin as the derivational suffix *-kk (e.g., Finnish kuusi 'fir-tree', kuusi-kko 'a group of fir-trees'). Hungarian and Cheremis show a reflex of this *-k in the marking of noun plurals, but other FU languages mark plurality in different ways in noun plurals than they do in person markers (cf. section 3.1.3.). Comparative evidence indicates that *-k was used as a plural marker in connection with the deictic particles/person markers, but further research is needed to determine its origin.

The dual exists in the Lappish, Samoyed, and Ob-Ugric languages. Décsy (1965) denies its existence in PFU or PU, but Hajdú and Collinder reconstruct a dual marker *-ka~kã(n) (< *kakte 'two') (Hajdú 1975:84; Collinder 1965:131). It is not at all clear how the present forms would have evolved from these reconstructed forms, because most reflexes show no sign of a *-ka element.¹² Hakulinen (1957:57) assumes that PU had a dual in its system, but its use was limited to 'special cases'. He does not elaborate what these special cases might have been. Because the languages that would be crucial for the establishment of proto-duals, Samoyed and Lappish, have been omitted in this paper, and because the dual forms do not add to the major arguments, no stand is taken with respect to the status of the PFU dual.

3. Role of the Deictic Particles in the Marking of Non-Personal Grammatical Categories.

Reflexes of the PFU deictic elements discussed in section 2 are not restricted to the three categories discussed in the previous chapter--personal pronouns, possessive suffixes, and subject agreement markers--but as I have indicated above, other grammatical morphemes represent reflexes of these deictic particles. In order to better elaborate the development of the morphological marking of certain grammatical categories in the FU languages, it is proposed here that these deictic particles, which originally referred to the participants in the communication act and to their location, came to be used as definiteness markers, in order to indicate the focus of the utterance. In subsequent developments, these same elements came to be interpreted as, on the one hand, person markers, and, on the other hand, accusative markers, plural markers, etc. The major characteristic associated with the entire set of reflexes considered here is definiteness.

Definiteness is generally analyzed as an inherent feature in personal and demonstrative pronouns. Demonstratives and third person pronouns are universal definiteness markers. Definiteness indicates something that is identifiable: the addressee can identify the particular referent the speaker has in mind. In the case of the deictic particles that refer to the closeness of the object to the speaker and/or the addressee, the identification is established by the situation. Definiteness can also be established by linguistic phenomena in the domain of a discourse when reference is made to some object which has been previously mentioned (further discussion, e.g., Moravcsik, 1969; Chafe 1976).

We can distinguish two types of definiteness markers, both derived from deictic particles. The first group (which represents a chronologically older development) includes accusative morphemes, person markers in objective conjugation, the plural morpheme *-t, and the so-called 'definite declension' in Zyrlian.

These grammatical morphemes developed from deictic particles used as focus markers in the utterance. The particles were placed after the focused constituent, which had the strongest stress in the sentence. Through phonological processes, the particle, which had weaker stress than the preceding constituent, became suffixed to it, and in the subsequent development of the language(s) it came to be reinterpreted as an inflectional morpheme.

The second group represents younger formations. These definiteness markers arose in the individual developments of the daughter languages. Whether they are direct derivatives from the deictic particles or extensions of the morphemes in the first group, is difficult to determine. Examples of grammatical morphemes of this type are the definite article in Hungarian, morphemes used in the definite declension in Mordva, and some clitics in Finnish. The constituents to which the deictic particles were attached in the first group conveyed new information; in the second group the

constituents introduce given (old) information into the sentence. They can be topic markers, as are the Finnish clitics, or they can be general definiteness markers which can be used to establish the definiteness of the topic constituent.

3.1. Older developments in the marking of definiteness.

3.1.1. Accusatives.

There has been a controversy over whether an accusative case can be reconstructed for the PFU. Hajdú, for example, supposes that the accusative case in *-m existed in PFU (Hajdú 1975:80). Other scholars have attempted to find various explanations for the occurrence of an *-m marker in some languages, and for the lack of it in others.

It is a generally accepted view that *-m marked only definite objects; indefinite objects in PFU were unmarked morphologically (Wickman 1955). The word-order for PFU has been established as SOV, which means that in the unmarked order subjects were placed in initial position. There is a universal tendency for subjects to be definite; they are generally the topic of the sentence. Objects were part of the new information, the comment. They could, however, become the focus, the emphasized part of the sentence. In that case they had to be specially marked. The focused object came to be marked by the deictic particles.

The object marker *-m does not exist in Hungarian and Ostyak, but these languages mark objects by -t (Ostyak marks only personal pronouns), which can be derived from a deictic particle. (Wickman 1955:73 agrees with other scholars that the Hungarian -t accusative marker was originally a demonstrative element.) Finnish marks the accusatives of personal pronouns with -t. It is therefore reasonable to suggest that at some stage in PFU, there was variation in the marking of the definite (focused) object by a deictic element, either *mV or *tV or *sV. This variation would apparently depend on the location of the object relative to the speaker. These elements lost their deictic meaning, and came to be interpreted as accusative (i.e. object) markers. In each language, one of the deictic elements was generalized. Considering the central role of the speaker in the communication situation, it is understandable that the 'first person' element was most generally regularized for this function. One must also consider the possibility that if these deictic elements, having had a general definite-marking function in the proto-language, came to be interpreted as inflectional morphemes in individual daughter languages, each of the particles may have been interpreted as a different type of inflectional morpheme (e.g., one was interpreted as an object marker, another as a plural marker).

In the modern Uralic languages there is typically no single morphological form that marks all direct objects and only direct objects. Comrie (1975) claims that there is an operating principle that the subject and the direct object are distinguished from one another in those circumstances where confusion is likely, and not

(necessarily) otherwise. Comrie calls this type of languages 'anti-ergative', because they mark the object in those circumstances and not the subject, as ergative languages do. Comrie's analysis may account for the synchronic facts, but historically, the object marking seems to be related to the marking of the focus of the utterance.

In light of these general comments about the origin of the FU accusative, I would like to review some of the descriptions of the accusative markers of individual languages to show how these facts might be accounted for by a hypothesis that PFU marked focused, definite objects by deictic particles which also served other major functions in the language.

Balto-Finnic and Cheremis reflexes are straightforward: $-\underline{n}$ ($*\underline{mV} > \underline{m} > \underline{n} / - \#$), and $-\underline{m}$ respectively. Mordva has a palatalized nasal \underline{n} which is not a regular reflex of $*\underline{m}$. It has therefore been problematic for the Finno-Ugrists. Wickman, for example, considers this reflex unexplained (1955:39). If, however, we assume that \underline{n} developed from a deictic particle which had a palatal vowel [e.g., $*\underline{mi}$], the palatalization is a reasonable process, especially considering that palatalization is a very common process in Mordva.

In the Permian languages, Zyrian and Votyak, the final $-\underline{m}$ was regularly lost. By another rule, final vowels were lost in disyllabic words. However, the fact that final vowels occur in accusative forms suggests the earlier occurrence of final $*-\underline{m}$. Hence accusatives and nominatives are kept distinct even though the original accusative marker was lost, for example, Votyak murt 'man' (nom. sg.); murte (acc. sg.). The final e belonged originally to the stem, but it is now interpreted as an accusative marker. The use of this type of an accusative is limited to only a few nouns and pronouns (Wickman: 58).

There is another definite object marker in Votyak and Zyrian: $-\underline{es}$ (Votyak $-\underline{es} \sim -\underline{ez}$). This marker has been identified as the third person singular possessive suffix. Wickman points out that in many FU languages the third person singular possessive suffix is used as a kind of a definite article. The vowel has been generalized as a part of this suffix, although it was originally part of the stem. Therefore in Votyak ajiz 'the father or his/her father (nom. sg.)', but ajez (acc. sg.). Wickman also reports that some dialects have $-\underline{ti}$ or $-\underline{te}$ as the object marker in the plural, which has been generally interpreted as an old PU ablative marker. Both of these object markers, es and ti/te could, however, be derived from the deictic particles $*\underline{sV}$ and $*\underline{tV}$.

As it was indicated above, the Hungarian accusative marker $-\underline{t}$ can be assumed to have its origin in a deictic particle $*\underline{tV}$. In the modern language, it is used to mark all objects, not only definite ones. Károly (1972) maintains, however, that it was first used to mark only definite objects. When the objective conjugation developed, the definiteness was marked in the verb, and $-\underline{t}$ became a general object marker, both indefinite and definite. If there is a possessive suffix

-m (1st sg.) or -d (2nd sg.), the accusative suffix is not often used (Collinder 1962). The absence of -t in words with possessive suffixes goes back to the period when -t denoted only the definite object, and the possessive suffix could denote definiteness in itself, without -t. In modern Hungarian, there is a tendency towards regularization of the -t accusative to all direct objects.

According to Wickman (1955:63), North Vogul, and Pelym dialects of South Vogul, have no accusative marking. In the other dialects there is an accusative containing the element m, either alone or mostly as a part of a suffix -mV (-ma, -mä, -me, -mi). This marker is used to indicate definite objects in those cases where there is no possessive suffix. Those scholars who have interpreted this accusative marker as a reflex of a PFU *m have considered the final vowel problematic. Collinder (1960:285) for example identifies it as a third person singular possessive suffix, but does not indicate why it would be used and why the reflex is a vowel when the third person singular possessive suffix in Vogul has a consonantal element -t. In my opinion, the Vogul object marker is a reflex of the deictic particle *mV.

Hajdú's discussion of object marking in Forest Yurak (one of the Samoyed languages) offers interesting insights into the function of the elements that have been called deictic particles in this paper (Hajdú 1960). He claims that this language does not have a regular accusative marker at all, but that possessive suffixes are also used secondarily to mark objects as their function. Mainly, it is the third person suffix -ta which is used to indicate the object. It "replaces the lost accusative suffix *-m." The possessive suffix of the second person is used for the same function. Hajdú does not clarify if and how these suffixes differ in function. He emphasizes, however, that when used as object markers, these suffixes have lost the function of marking the possessor. Hajdú makes three points: first, the use of the possessive suffix as an object marker is not consistent, unmarked objects being quite frequent (he does not specify whether there is a semantic difference between marked and unmarked objects); second, these same elements can still function as possessive suffixes in other contexts; and third, these suffixes can serve to mark the genitive case if used as attributes of nouns.

It is interesting to note that Forest Yurak has object forms which have a grammatical marker containing an m (i.e., which is identical to the first person possessive suffix). Hajdú rejects vehemently the idea that this marker could be regarded as the same type of an object marker as the second and third person possessive suffixes. According to him, the indication of possession is often there, which is not the case with second and third person possessive suffixes. It seems to me that a more detailed analysis of the Forest Yurak data might clarify the connection between -m-suffixes that Hajdú considers possessive suffixes, 1st sg., and those m-suffixes

that he views as relic forms of the old PFU *m accusative. Hajdú's report on Forest Yurak indicates that the deictic elements (he calls them possessive suffixes in this case) can have various functions even at more recent stages of the daughter languages. The same elements can function as object markers, markers of possession, and as genitive markers. Either the context or the word order disambiguates the meaning.

It is interesting that Lappish has developed a definite object marker which is -ta~-tä, a marker that is identical with the Finnish partitive case marker -ta~-tä. The partitive case in Finnish is used among other things to mark indefinite objects. Wickman proposes that Lappish developed this definite object marking under Finnish influence. It had the *m accusative marker in the singular to mark definite objects, but no marker in the plural. According to Wickman, Lappish borrowed the -ta~-tä morpheme from Finnish, and used it to mark definite objects instead of indefinites, as its function is in Finnish. One could conjecture, however, that the definite object marker in Lappish is of different origin, that it is derived from the PFU *tV instead, and has the function of a definiteness marker from the proto-language.

Both the function of accusative morphemes (to mark definite objects) in FU languages and the actual phonological shapes of this morpheme in individual FU languages can be accounted for if we assume that the elements which are used for the marking of accusatives in the modern FU languages derived from the deictic particles *mV, *tV and *sV in the proto-language.

3.1.2. Objective (definite) conjugation

In some Finno-Ugric languages, there is another means of turning the attention of the addressee to the definiteness of the direct object: the speaker may indicate the definiteness of an object by the use of objective conjugation, i.e., the definiteness is marked in verbal forms instead of being marked in the object NP. In these languages, transitive verbs have two sets of person markers, on the one hand, those of an objective (definite) conjugation; on the other hand, person markers of the subjective (indefinite) conjugation.

An objective vs. subjective conjugation distinction exists in Hungarian, Vogul, and Ostyak (i.e., in the Ugric branch), as well as in Mordva and the Samoyed languages. The two-conjugation system in these languages is not an inheritance from the proto-language, but has been determined to be a parallel development.¹³

As an illustration of the differences between the person markers in the objective and subjective conjugations, consider the paradigm of the Hungarian verb lát 'see':

	<u>objective</u>	<u>subjective</u>
sg.	1. lát <u>-om</u> 'I see (something that can be identified, something definite)'	1. lát- <u>ok</u> 'I see'
	2. lát- <u>od</u>	2. lát- <u>sz</u>
	3. lát- <u>ja</u>	3. lát
pl.	1. lát- <u>juk</u>	1. lát- <u>unk</u>
	2. lát- <u>jatok</u>	2. lát- <u>tok</u>
	3. lát- <u>jak</u>	3. lát- <u>nak</u>

(Decsy 1965:176)

The singular forms of the person markers (deictic particles) in the objective conjugation result from straightforward phonological changes, as was indicated in section 2.3.3. The -j has been generalized from the third person singular to all plural forms, and it has become a marker of objective conjugation. The morphemes that are interpreted in the grammatical system of Hungarian as first and second person singular person markers in subjective conjugation were originally aspect markers (cf. p.105). The aspect markers were reinterpreted as person markers in those cases where a person marker was felt to be missing.

-Sz and -l (which occurs as second person singular marker with some verbs) have been determined to be reflexes of PFU markers for iterative action. The person marker -k in the first person singular has been analyzed as an analogy from the first person plural, or of indeterminant origin (Rédei 1966; Károly 1972). I do not see any reason why it could not be a reflex of the PFU aspect marker *-k which marked continuous action. This marker has been established on the basis of cross-linguistic evidence, and has played a role in the development of person markers in, e.g., Finnish. If the second person singular person marker developed from an aspect marker, it is quite plausible that the first person marker has a similar origin; especially because we can reconstruct this particular aspect marker, *-k, for the proto-language.

The first person plural suffix -unk may come from mk which is derived from *mV_k (mk is historically attested). This is therefore a suffix that we would expect in the objective conjugation, not in the subjective one. It is possible that this suffix was introduced to subjective conjugation after the j had been generalized into the objective conjugation for all plural persons. Thus j had become a marker of the objective conjugation. There is another explanation for n in -unk. There are still relic forms in Hungarian of third person singular forms with n, e.g., hiszen '(he) believes'; megyen '(he) goes'; vagyon '(he) is'. Etymologically, they have been interpreted as nominal forms. The third person plural forms in subjective conjugation -nak (and its vowel harmony variants) are derived from this form with the addition of the plural marker -k. The first person plural could therefore be derived from this same nominal form (Rédei 1966).

As it was indicated above (in section 2.3.5), the descriptions in the handbooks suggest that Proto-Finno-Ugric had two types of verbal forms, one with the 'person marker' (which I interpret as a deictic element), and the other without any morphological person marker (or definiteness). In the analysis which I propose in this paper, the focused constituent, which could be, e.g., the object NP, was marked by a deictic particle. The reflexes of that focus marking can be seen in the accusative morphemes in the modern languages (cf. section 3.1.1.). The focused element could, however, be the whole VP constituent in which case the deictic particle followed the verb (the word order being SOV). This is the origin of subject agreement markers in all FU languages. In most Finno-Ugric languages, the deictic clitics, which became suffixes and were reinterpreted as person markers, were generalized to all finite verb forms, but in those languages that show the objective/subjective conjugations, only the verb forms in the objective (definite) conjugation have reflexes of the deictic particles.

The marking of the definite object, the inflection of the verb in the objective conjugation, and the word order of the sentences seem to have interesting interrelationships. This topic cannot be discussed in the framework of this paper, but I would like to refer to some correspondences in present-day Hungarian and Ostyak (both of which have objective/subjective conjugations). The examples are greatly simplified for the purposes of the present discussion.

Hungarian can have two basic word orders, either SOV (archaic) or SVO (innovative). The SOV order occurs in sentences which contain an object without an article, whereas the sentences with definite (or indefinite) articles are SVO. Examples:

- | | |
|--------------------------------------|---|
| SOV order and subjective conjugation | <u>A fíu levelet ír</u>
the boy letter writes
'the boy is doing letter-writing' |
| SVO order and objective conjugation | <u>A fíu írja a levelet</u>
the boy writes the letter
'the boy is writing the letter' |

(Hetzron 1975)

Bése et al (1970) define the semantic differences between the above sentences as follows: in the sentence A fíu levelet ír the reference to the context or to the communicative situation is irrelevant. In this type of structure, the object has to be closely dependent on the predicate (i.e., semantically). In the second type A fíu írja a levelet, the object is related to the context or the situation. The identification refers to the noun a levelet, which is known from the context. The verb phrase contains a verb in objective conjugation and an object with a definite article.¹⁴

There is a third type of structure in Hungarian:

A fiu ír egy levelet
the boy writes a letter

where the verb is in the subjective conjugation, the object has an indefinite article, and the word order is SVO. This would be translated 'the boy is writing a (specific) letter. The object is specific, but is not given or known from the situation, so the subjective conjugation is used. These Hungarian examples show that the verb, or the object NP, or both, can be marked if there is some reference to the context or situation,' i.e., when it was definite, the deictic clitics were used to mark this feature. Hungarian had three ways of marking the definiteness of the object: by marking it in the verbal form, or in the object NP by -t, or the chronologically youngest way of marking, by the definite article. The object marker -t came to be generalized to all direct objects, not only to the definite ones, and it lost its function as a definiteness marker.¹⁵

According to the handbooks, there is a general rule in Hungarian that definite or objective conjugation always occurs if the sentence has a definite object. Personal pronouns are considered inherently definite, and therefore the verb should be in the objective conjugation if there is a personal pronoun object in the sentence. If the object is a third person pronoun (e.g., 'I saw him'), the verb is in the objective conjugation as expected, but if the object is a first or second person pronoun (e.g., 'he sees me'), the verb is in the subjective conjugation. According to Comrie, this is the situation in Hungarian, Vogul, and some dialects of Ostyak (Comrie 1975a). It is possible that these examples reflect an earlier situation when the situational context, the definiteness, was established by deictic particles (which gave rise to personal pronouns among other things), and the marking of definiteness can be omitted in the verb form because it would be redundant.

3.1.3. T-plurals

It has been firmly established that the Proto-Finno-Ugric language has a way of marking plurals by a *-k which was suffixed to the person markers/deictic particles (at least at the reconstructed stage). There is, however, disagreement among scholars as to whether a plural morpheme *-t can be reconstructed for the proto-language. This marker occurs in many of the daughter languages as a noun plural marker, as well as in third person plural subject agreement markers. The third person plural verbs showing this marker were originally nominal forms. Hungarian and the Permian languages (Zyrian and Votyak) do not show a reflex of this plural marker. Hungarian marks all plurals with -k; Zyrian marks noun plurals with -jas, Votyak with -jos. Cheremis shows the -t a plural in the third person subject agreement marker, but the noun plural morpheme -vlak is clearly an innovation.

There are, however, restrictions on the occurrence of the -t- plural even in those languages where it exists. Finnish, for example, has the -t- plural marker only in the nominative plural (which is identical to the accusative plural); other case forms have other plural markers (-i/-j). Décsy (1965:158) rejects the idea of a PFU plural marker *t on these grounds, and Collinder (1975:128) reconstructs a nominative plural marker *-t, and a plural marker *-i for the oblique cases (i.e., a situation which prevails in modern Finnish; e.g., talo-t 'houses' (nom. pl.) vs. talo-i-ssa 'in the houses' (inessive pl.)). It is interesting to note that no -t- plural marker occurs in cases where a possessive suffix is attached to the noun, for example, in Finnish talo 'house'; talo-t 'houses', but talo-ni 'my house or my houses'.¹⁶

Ravila (quoted in Hakulinen 1957:60) has suggested that in PFU, the plurality was marked only in the predicate, but not in subjects. As evidence, he cites constructions of the type: *lintu (singular) lentävä-t (plural) 'birds fly', where the subject is without plural marking, but the predicate has the plural marker -t. The noun plural marking developed, according to this theory, as a congruence phenomenon.

The evidence of the modern languages suggests that the PFU language did not have a noun plural marker per se. There was a way of marking plurality by *-k. Comparative reconstruction forces us to reconstruct the *-k morpheme as a suffix attached to the person markers/deictic particles. It does not, however, allow us to go further back in the history to determine the origin of this element, i.e., whether it was a numeral or a morpheme meaning 'a group of' (see p.109).

It has been argued in the previous chapters that deictic particles were used in the Proto-Finno-Ugric language to indicate the location/role of the participants in the speech situation. These particles came to be used as definiteness markers to indicate the focus of the utterance. In section 3.1.1, I argued that the definite, focused object came to be marked by the deictic particles. In different daughter languages, different deictic particles became generalized for specific functions after having lost their original deictic meaning. Present-day Hungarian, for example, has -t as a general marker of direct objects; many of the other languages have -t as a plural marker. If we analyze the plural morpheme -t in the modern FU languages as a reflex of the PFU deictic particle *tV, we can account for some problematic aspects of its distribution.

I suggest that the plural marker -t arose from the marking of the focused element in the sentence by the deictic particles. In the course of history, *tV came to be interpreted as the plural marker in those daughter languages which show it now. Whether this development took place entirely in their separate histories or started in the proto-language is difficult to determine. At least two specific facts about the use of the t-plural in Finnish seem to support this hypothesis. In Finnish, there is a distinction between 'total' and 'partial' objects. This distinction is a very complicated issue, but

one of the characteristics of the 'total' object is that it is definite. In the singular, the accusative case is used to mark the definite, 'total' object, but in the plural, the marker is -t (i.e., the nominative plural marker). It has been established that the accusative case (i.e., direct object marking) in FU languages arose in the marking of definite objects, not of all direct objects (e.g., Wickman 1955; section 3.1.1 in this paper). Why would the definite object in the plural be marked with the nominative marker, and not with an accusative case marker, i.e., why is there no accusative case marker in the plural? However, if the -t-plural marker arose from the marking of definiteness with the deictic particle *tV, and if the accusative marker also has the same origin (reinterpretation of another deictic particle *mV), the use of t in the plural to mark the direct, definite object can be explained. Originally both *mV and *tV were used for the same function, to define the location of the object under discussion, with respect to the participants in the speech situation. Later they were reinterpreted as different grammatical morphemes, *mV as an accusative marker, *tV as a plural marker. The above analysis would also explain why no -t-plural marker occurs in cases where a possessive suffix is attached to the noun. If the possessive suffix also has the same origin, being derived from a deictic particle and marking definiteness, it was unnecessary to add another "definiteness" marker, -t. Therefore we have talo-t 'houses', but talo-ni 'my houses' (the latter without a plural marker).

This analysis of the origin of the -t-plural accounts for its absence in some languages, and it also accounts for the Finnish data, where the nominative plurals which have the plural marker -t, function as definite objects, and for the fact that this plural marker is omitted in cases where the possessive suffixes occur.

3.1.4. Miscellaneous uses of the PFU deictic particles in definiteness marking

In addition to object markers, person markers in objective conjugation, and -t-plurals which can be interpreted as reflexes of deictic particles, which all share the feature 'definiteness', there are other grammatical morphemes in FU languages which seem to be derived from the same proto-elements and which have the function of marking definiteness. Such reflexes occur in Zyrian, Votyak and Mordva.

According to the handbooks (e.g., Décsy 1965), Zyrian has a 'definite declension,' i.e., a way of adding certain suffixes to noun roots to make them definite. The suffixes used are -yd and -ys, which can be derived from the deictic elements *tV and *sV (the -y-element is probably a transition glide which has been reinterpreted as part of the suffix). Décsy maintains that the forms with the suffix -ys are 'emphatically neutral' whereas the forms with -yd indicate something familiar, something personal. Examples used to illustrate the point are: mortlön 'with a person' (root mort and a case suffix -lön); mort-ys-lön 'with the person' (neutral); mort-yd-lön 'with the (nice) person'. Décsy also reports that in Votyak, adjectives can

have a suffix -ez ~ -yz (< *sV) attached to them to emphasize or intensify the semantic value of that adjective, for example, vyl 'new'; vyl'ez 'really new'.

The use of -yd in Zyrian and -ez ~ -yz in Votyak may represent an example of the usage of deictic particles which Lakoff has called 'emotional deixis' (Lakoff 1974). The use of deictic particles is linked to the speaker's emotional involvement. They are often used for vividness. Unfortunately the handbooks do not provide more examples of this special usage of deictic particles (which the authors analyze as possessive suffixes in modern languages).

Mordva shows another use of person markers/deictic particles, which in my opinion illustrates the deictic origin of person markers. Mordva is one of the FU languages which lack a copula. It has one-word sentences where the person of the subject is marked by special suffixes attached to the nouns, adjectives, adverbs or numerals which form the predicate. The grammar books call this a 'predicative declension'. Décsy (1965:192) gives an example of such a paradigm:

sazor 'sister'

- sg. 1. sazora-n 'I am (somebody's) sister'
2. sazora-t 'you are (somebody's) sister'
3. sazor

- pl. 1. sazor-tan(o)
2. sazor-tad(o)
3. sazor-t

The morphological markers used in the 'predicative declension' are derived from PFU *mV and *tV; the third person singular has a zero marker. The interesting fact about these forms is that e.g. sazora-n does not mean only 'I am (somebody's) sister', but depending on the context it can also mean 'this sister' or 'the sister here'.¹⁷

A characteristic common to both the 'definite declension' of Zyrian and Votyak, and the 'predicative declension' of Mordva is that the morphemes which comprise these paradigms are derived from elements which presumably functioned as deictic particles in the proto-language, and that they define the constituent to which they are attached in terms of its relationship to the speaker.

3.2. Later developments in the marking of definiteness

I have argued in the previous sections that in the history of the FU languages, the deictic particles had several functions, one of them being to mark definiteness, specifically to mark the emphasized, focused constituent of the sentence. There is, however, another group of morphemes that are used to mark definiteness, and which have also developed from deictic elements. This group includes the definite articles in Hungarian and Mordva, as well as certain topic-marking

clitics in Finnish. All these represent a late development in the respective languages, and all the morphemes are identical to or can be derived from demonstrative pronouns (or personal pronouns).

3.2.1. Definite articles

Vennemann (1975:298) discusses the development of definite articles from demonstratives, and maintains that definiteness is closely related to topicality. Verb-last languages do not have articles. According to him, articles develop through a non-deictic, anaphoric use of demonstratives in TVX (Topic - Verb - Verb Complement) languages.

In the Finno-Ugric language family, only Hungarian has developed genuine definite articles. The development has often been assumed to represent an Indo-European influence on Hungarian, but if one compares the grammatical systems of other FU languages, the development of articles in Hungarian can be seen to be based at least in part on internal changes in Hungarian, especially the word-order change from SOV to SVO. The use of the Finnish clitic se with or without the demonstrative se (which is discussed in the next section) might be an early stage in the development of a definite article. Whether Hungarian went through a similar stage cannot be determined from historical records. On the other hand, the use of se in colloquial Finnish as a definite article could also be a result of Germanic, especially Swedish, influence. The development of definite articles out of demonstrative pronouns is not a very drastic historical change in either case, because demonstrative pronouns provide a source for definite articles universally. The following discussion only demonstrates that the definite articles in FU languages have a similar history.

The definite article in Hungarian has two morphophonemic alternants: a (before a consonant) and az (before a vowel). Collinder and Károly explain it to have developed from the demonstrative pronoun az 'that' (the z comes from PFU *tV) (Collinder 1960; Károly 1972). The definite article developed as late as the 12th century. In the early stages, the demonstrative pronouns were used together with the articles, for example, azt az embert 'the man' (actually 'that the man', acc. case). The handbooks do not indicate whether this type of double marking occurred in all positions in the sentence. Definite objects were traditionally marked in all FU languages. It would be interesting to determine whether a correlation exists between definite objects and double marking with the definite article. A further analysis of the earliest Hungarian texts might shed light on this aspect.

Mordva has a suffixed enclitic article ('definite declension') which has four shapes: s' (in the nominative singular); t' (in the genitive singular); n't' (in the other singular cases); n'e (in the plural). Décsy (1965:192) derives them from various demonstrative

pronouns: s' from s'e 'this'; t' from t'e; n'e from the demonstrative pronoun n'e 'these'. The sources do not indicate whether the clitic articles represent similar reflexes of the proto-elements as to those of the present demonstrative pronouns, or if they are in fact such late developments that they are directly derivable from the demonstrative pronouns.

Other FU languages do not have articles, but according to grammar books, they can use possessive suffixes (especially 3rd sg. and 2nd sg.) as articles. Such use of possessive suffixes is characteristic of Zyrian, Votyak, Cheremis, and Vogul and Ostyak (Collinder 1960). Collinder claims that in the above-mentioned languages, possessive suffixes have two functions, one of marking possessors, and the other of functioning as a definite article. He claims that the context makes it clear which usage is intended. However, no clear examples of this distinction are given in the handbooks. It is possible that the marking of definiteness through possessive suffixes (even after they have lost their deictic meaning) indicates the process through which FU languages eventually develop genuine definite articles.

Finnish has no articles, and neither are possessive suffixes used for this function. In colloquial language, the third person pronoun se can be used in a similar function, more rarely the other two demonstratives (corresponding to 'this' and 'that'), i.e., they can lose their deictic meaning. Because these morphemes still function as demonstrative pronouns, only the context and the stress will indicate whether (or not) they have deictic meaning in any particular situation.

3.2.2. Finnish clitics

It was mentioned above that the morpheme se (deictic particle for 'audience deixis'/third person) can be used in colloquial speech as a definite article, although this use is not yet accepted in the standard language. It is very interesting that the same element can appear either before a noun or after a noun, or simultaneously in both positions. Se poika can mean 'that boy' or 'the boy'; poika se osti auton 'the boy bought a car' or 'it was the boy who bought the car'; se poika se on suosittu 'that boy is (really) popular', or 'it is that boy who is (really) popular'. The enclitic se is used to topicalize the preceding constituent.

There are at least three unaccented morphemes in Finnish which have developed from the third person deictic particles: se, sitä, and -han ~ -hän (sitä is the partitive case form of se). They all indicate something that is clear, self-evident, well-known in the particular situation. Se and sitä are free morphemes, -han ~ -hän is a bound morpheme. They all follow the first major constituent in the sentence. Examples illustrating the use of these clitics: sinä se olet onnen poika 'it is you who is the lucky guy'; poikaa sitä viedään vihille 'it is the boy who is brought to the altar'; kauniina

aurinkoisena päivänä sitä ollaan iloisia 'it is on a beautiful, sunny day that people are happy'; tuleehan se kesäkin joskus 'it is a fact that summer comes some time' (order of elements in this sentence: tulee ('comes', 3rd sg.) -han (clitic) se (another clitic) kesä ('summer') -kin ('also', clitic) joskus ('sometime')); teilläähän sitä on rahaa 'it is you who have money'. In the last example, both sitä and rahaa are in the same case, the partitive, which might indicate that they belong together, i.e., that sitä modifies rahaa. If, however, we analyze other sentences, e.g., me sitä ollaan juotu yhdessä monet oluet 'it is us (or we) who have drunk many beers together', we see that sitä topicalizes me 'we', and there is no partitive case noun in the sentence.

The above-mentioned clitics represent a late reflex of the PFU deictic elements. It would be interesting to investigate whether other daughter languages show similar developments whereby former deictic elements come to be used as topic markers. Further study is needed to clarify the history as well as the synchronic analysis of these topic markers.¹⁸

3.3. Demonstrative pronouns

I have claimed above that a great variety of forms in present FU languages are derived from three deictic particles whose broad functions in the proto-language were reinterpreted in several ways. In this section, I will briefly consider some features of the demonstrative pronoun system in FU languages in order to identify other reflexes of these deictic elements in the modern demonstrative pronoun systems.

Finnish has a three-way distinction in demonstratives: tämä 'this'; tu 'that'; se 'that yonder' (corresponding to the three-way distinction in personal pronouns). It is difficult to determine from the information the handbooks present whether the three-way distinction occurs in other daughter languages, because one and the same term can be translated as 'this', 'that', or 'it'. Without first-hand knowledge of the languages in question, it is difficult to ascertain the exact function of any demonstrative in the language. It is made more difficult by the fact that the literature I have consulted is written in German or English, which have only a two-way distinction, this vs. that, since the translations may be affected by the language used in the description. Vértes (1967) presents extensive data on demonstratives in Ostyak dialects, and several dialects exhibit a three-way distinction, similar to Finnish (e.g., Scherkaly dialect tam 'this' ("dieser"); tom 'that' ("jener"); sí 'that' ("der")).

Traditionally *t- has been reconstructed as an initial consonant segment for the singular demonstratives, *n- for the plurals. Besides Finnish (tämä/nämä 'this/these'; tu/nuo 'that/those'; se/ne 'that/those, "3rd person"'), other languages show t/n correspondences, for example Erza Mordva t'e/n'e 'this/these'; Cheremis has tə or təõe/nəne 'this/these' (Tauli 1966:245). The occurrence of -n in plural

demonstratives has caused some scholars to reconstruct an *-n as a plural marker for the proto-language (e.g., Szinnyei 1922). More recent studies have shed new light on this issue, however, and scholars tend to agree that *n- represents another deictic element in the proto-language, and not a plural marker (e.g., Tauli 1966; Collinder 1965; Hajdú 1975). The Permian-Finnic languages have just simply generalized the deictic element *n- with the plural meaning.

Hajdú (1975) points out the striking similarity between the two proto-segments for the demonstratives, *t- and *n- on one hand, and the reconstructed locative suffixes *-na ~ -nä and *-tt (*-tV) on the other hand. He supports a theory according to which locative suffixes developed from demonstratives. The same locative suffix can be found in the personal pronouns, for example mi-nä 'I' in Finnish.

Vértes (1967) has also commented on the reflexes of PFU *t- and *n-. She claims that the locative suffixes developed from demonstratives, but furthermore, she emphasizes that in various languages *t and *n developed different functions, i.e., they were reinterpreted in different ways. In 'western' languages, i.e., mostly the Permian-Finnic group (and Hungarian), *n became part of a locative suffix, and *t came to be interpreted as the second person marker (cf. the earlier discussion), but in the Ob-Ugric languages (and Samoyed) *n became the second person marker, and *t a locative suffix. These facts have led Collinder to assume that the proto-language had two morphemes (which he calls personal pronouns) which referred to the addressee, *tV and *nV (Collinder 1965). Whether the deictic elements which contained the initial *t- and *n- both referred to the addressee, or whether they were semantically differentiated in the proto-language, is difficult to determine. The occurrence of an n- reflex in so many forms of person markers and demonstratives (cf. the discussion about the role of the n- affix and the -na ~ -nä suffix in personal pronouns) suggests that the proto-language had a deictic element *n. It has not been determined whether this element was derived from *mV as Castrén suggested (relating to the speaker), or whether it had some other deictic meaning. In any case, there does not seem to be any reason to reconstruct *n- as the plural marker in demonstratives, or to assume that *n was a plural morpheme in any case.

I have so far discussed only the initial segments of the demonstratives. The demonstrative particles have been reconstructed with the shape CV in the proto-language. The vowel quality has been morphologized to mark semantic distinctions. Finno-Ugric languages show a general tendency in the system of demonstratives to have a front vowel in demonstratives which indicate proximity to the speaker (i.e., *te/ti; *ne/ni) and a back vowel in those demonstratives which indicate a further distance (i.e., *ta/to, *na/no) (Hajdú 1975). Examples that illustrate this are Finnish tämä 'this' vs. tuo 'that' (< *too); Mordva te- 'this' vs. to- 'that'. This semantic value of vowel quality is not absolute, however. The examples in grammars indicate that many languages have front vowels for demonstratives which denote distance from the speaker, and back vowels in those that indicate

proximity, e.g., Zyrian and Votyak ta 'this', Zyrian ti 'that'. The reconstruction of front and back vowels for the proto-language (as markers of semantic distinctions) does not seem to be justified. The present-day demonstratives may rather represent language-specific developments.

It is not inconceivable that the proto-language could have distinguished more locative oppositions with respect to the speaker and the addressee than the basic three-way distinctions. Some modern languages show such a system. Lappish shows five terms, the Samoyed languages have even more. They mark the location of the objects which are closer to the speaker than to the addressee; or objects which are in the same location as the speaker or the addressee. For example, Norwegian Lapp had dāt 'this here'; dāt 'that, this' (with a weakened demonstrative meaning); diet 'that (nearer to the person addressed than to the speaker)'; duot 'that one over there'; dōt 'that one far away over there' (Tauli 1966:141). A Northern Häme dialect of Finnish makes a similar distinction tämä 'this, i.e., where I am'; toa 'between you and me'; tuo 'that, i.e., near you'; se 'that, i.e., further away, not necessarily in sight'. However, the individual languages do not allow us to reconstruct specific entries for the proto-language.

There are two basic shapes of demonstratives to be found in FU languages: CVC(V) and CV, for example, Finnish tämä 'this' and tuo 'that'. The formation of the CVC(V) type resembles the formation of personal pronouns. A suffix is added to the demonstrative 'stem', *tV, *nV, or *sV, but the origin of this suffix is not quite clear, because the various languages show different patterns.

Finnish has CVCV shape only in tämä/nämä 'this/these' in the nominative case; in other case forms, the stem is tä-/nä- (e.g., tä-ssä 'in this, here'; nä-i-llä 'on these').¹⁹ Both tuo 'that' and se 'that, 3rd person' have only the CV shape. It is possible that tuo, se and other CV forms are derived from the unstressed forms of the demonstratives, i.e., they represent the reflexes of the original deictic particles. The sequence -mä in tämä/nämä looks too much like the deictic particle *mV to be a pure coincidence. If we look at other FU languages, we find stressed demonstratives in which the second CV sequence resembles a deictic particle, e.g., Ostyak has additional 'pronominal suffixes' attached to the demonstrative stem t-, such as -mV, -tV and -jV- tam(i) 'this'; teje 'this'; tit 'this' (Vertes 1967; all of these forms would not occur in the same dialect). Mordva has tona 'that' (cf. mon 'I'; ton 'you'); Cheremis has tudo or tučō 'that'; Mordva also has sec'e 'that, 3rd person'. As these examples indicate, the formation of demonstrative pronouns directly parallels that of the personal pronouns. The main pattern seems to be CVCV, where the first CV consists of the deictic element and the second CV can be analyzed either as another deictic element or a locative suffix. If, as Hajdú argued, locative suffixes are ultimately derived from demonstratives, or deictic particles, we might be able to suggest that the stressed forms of demonstratives had parallel developments in the daughter languages. The variety of reflexes in the modern languages is a result of language-specific reinterpretations of the deictic particles.

4. Conclusion

The literature on FU languages contains sporadic comments on the 'similarity' between person markers and certain grammatical morphemes, but any explanation for this similarity has been lacking. I have shown in this paper that we can reconstruct the elements **mV*, **tV*, and **sV* for the proto-language, but if we want to account for all the instances where these elements have played a role in the history of the Finno-Ugric languages, we have to reconsider their function in the proto-language. They were neither person markers, nor personal pronouns, nor possessive suffixes in the proto-language, but general deictic particles which could serve several functions at an early stage. They referred to the role of the speaker and the addressee in the speech context: **mV* referred to anything connected with the speaker or in the proximity of the speaker ('speaker deixis'), **tV* referred to anything connected with the addressee or his location ('addressee deixis'), and **sV* to anything that was not connected with the speaker's or the addressee's location ('audience deixis').

The reanalysis of person markers as deictic particles in early FU is related to the discourse notions of focus and topic. At an early stage, the deictic particles were used, among other things, to mark the focused constituents in the sentence. In the course of the history of individual daughter languages the deictic clitics were reinterpreted as various inflectional affixes. Therefore the same proto-element can occur in a variety of functions in the extent languages.

Moreover, this reanalysis of the reconstruction of person markers also explicates the controversial reconstructions of the accusative morpheme **-m*, and the plural morpheme **-t*. These morphemes were originally associated with definiteness marking. The deictic particles marked definiteness by indicating the role/location of the object in the speech situation, i.e., with respect to the speaker and/or the addressee.

The primary goal of this paper is to show that the reconstruction of grammatical morphemes has to take a larger context into consideration. If one applies strict comparative method in the establishment of e.g., an accusative morpheme for the proto-language, one misses the generalization that can be found relating this morpheme to other grammatical morphemes which marked definiteness at an earlier stage. In the reconstruction of morphemes for different grammatical categories one has to take into consideration the fact that many categories may be internally related, i.e., that the proto-language may have had only one category which is reflected in a variety of contexts as a result of reinterpretations of the reconstructed morphemes in question.

Decisions about relative plausibility play a role in any choice between alternative reconstructions. The traditional reconstructions of the accusative morpheme **-m*, the plural morpheme **-t*, as well as the reconstruction of person markers as either personal pronouns or possessive suffixes, did not take the full function of these elements

into consideration. The explanations were not plausible within the framework of sound historical methodology.

The evidence for the reconstruction of deictic particles for the proto-language comes from the roles these elements play in the extant languages, as markers of person, as well as from the obvious formal relationship which obtains between person markers on the one hand, and other morphemes with less obvious semantic relations to personal affixes on the other. The issues raised in this paper demand an even more thorough analysis. Some closely related problems have not even been considered, e.g., the relationship of the ordering of the morphemes discussed here with respect to case suffixes. This might offer insights concerning issues of chronology in the reinterpretation of deictic particles, and concerning the development of the FU case systems, for it seems to be clear that deictic particles also play a role in the development of a number of case markers in the Finno-Ugric languages.

Footnotes

*This is a revised version of my 1977 M.A. thesis. I wish to express my sincere thanks to my advisor, Professor Robert Jeffers, who introduced me to the interesting aspects of historical linguistics, inspired me in his lectures, read several versions of this paper, and showed immense patience during all that time. I would also like to thank the other members of my committee, Professors Arnold Zwicky and Olga Garnica, as well as Professor Ilse Lehiste, for their support and kind comments. But most of all, thanks to the whole faculty of the Ohio State University Linguistics Department for the knowledge they labored so long and so hard to impart to me. I am also grateful to my husband Joe, who agreed to make our home in Columbus, Ohio, against his better judgment.

Kiitos teille kaikille!

1. As the family tree of the Uralic languages indicates (see Appendix I), Proto-Uralic includes Samoyed languages; Proto-Finno-Ugric refers to the reconstructed stage of all other Uralic languages, except Samoyed. This paper deals with Proto-Finno-Ugric, but occasional references are made to Proto-Uralic.

2. The bibliography includes all the material used for this paper. It is not exhaustive on the topic, and I admit that Finno-Ugrists may have reason to object to the exclusion of potentially important evidence. This paper is a starting point for a more extensive study. A more thorough exposure to the relevant languages is needed to re-evaluate the statements in the handbooks, and to confirm the translations of the deictic elements in the literature. Finnish examples are based on my personal knowledge of the language.

3. The terms 'focus' and 'topic' are used in this paper in an informal manner. The usage of these terms corresponds to Chafe's terms 'contrastive' and 'given.' "...They...have to do with the

speaker's assessment of how the addressee is able to process what he is saying against the background of a particular context" (Chafe 1976:27). Given information ('topic' in this paper) is that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance. The focus of the contrast ('focus' in this paper) is the knowledge which is selected by the speaker from the set of possible candidates the addressee might have in mind. The focused elements are indicated in handbooks as 'emphasized elements.' As the terms are used in this paper, one sentence may have more than one focused element.

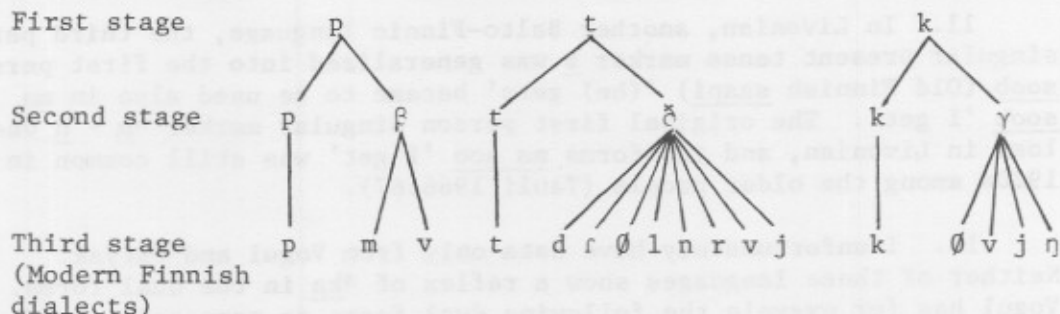
4. Finno-Ugrists use a very narrow phonetic transcription which is to a certain extent language-specific, i.e., different symbols are used to transcribe the same sounds in different languages. The handbook authors use their own simplifications in order to standardize the language-specific transcription systems. Examples in this paper have been simplified in those cases where the phonetic values of individual sounds are not under discussion. Finnish and Hungarian examples are given in their orthographic forms. The symbol \hat{a} here indicates "an etymologically short a", and \hat{u} indicates "an etymologically long u" (Collinder 1960:38).

5. 'Consonant gradation' refers to alternation in duration and manner of articulation between certain consonants conditioned by the structure of the syllable in the beginning of which the sound occurs. In Finnish, only certain stops (p,t,k) participate in the gradation, but in Lappish, for example, all consonants are subject to it. Modern Standard Finnish has the following alternations:

pp~p	:	huppu/hupun	'hood' (nominative singular/genitive singular)
tt~t	:	katto/katon	'roof'
kk~k	:	kukka/kukan	'flower'
p~v	:	lupa/luvan	'permission'
p~m	:	kumpu/kummun	'hill'
t~d	:	kato/kadon	'loss'
t~l	:	kulta/kullan	'gold'
t~r	:	kerta/kerran	'time'
t~n	:	ranta/rannan	'shore'
k~∅	:	joki/joen	'river'
k~j	:	järki/järjen	'intelligence'
k~v	:	luku/luvun	'chapter'
k~ŋ	:	vanki/vangin	'prisoner'

Anttila (1972) maintains that 'consonant gradation' in Finnish can be characterized as the alternation of a voiceless stop and a voiced continuant (except of course in the cases where a geminate stop alternates with a simple stop). Anttila presents three stages of the gradation. The second stage represents the Balto-Finnic gradation. Hakulinen (1957) and Posti (1963) assume that s participated in this alternation: s ~ z.

Earliest Finnish records show that d was a spirant, written dh or d̥ (one Finnish dialect still retains [ð̥] for d̥. The variant d̥ in Standard Finnish is a late spelling pronunciation; dialects have either ∅, j, v, l, or r. *γ was written gh or g (Anttila 1972).



6. "Western languages" that have become SVO include the Balto-Finnic languages and Hungarian.

7. The literary language is moving in this direction, as indicated by Hakulinen (1960:255). In the translation of one passage in the Bible, the 1936 version had 203 cases where the first and second person pronouns had been added to the 1913 version.

8. Well-known examples are e.g., the 'all-purpose' prepositions, such as long in Neo-Melanesian. I would also like to refer to Sankoff and Brown's article The Origins of Syntax in Discourse (1976) where the authors discuss the function of ia in Tok Pisin. Ia is etymologically derived from English 'here'. It is a postposed deictic marker which has an adverb of place function to some limited extent, but which is more frequently used to modify other expressions in the place deixis. Ia has, however, another function in the language: it is used as a focus marker. It can be postposed to a Noun or Pronoun and have the function of placing focus on that element. Furthermore, ia is considered a third person singular focal pronoun, i.e., it functions 'emphatically' or demonstratively in combination with personal pronouns to focus on a pronoun. The deictic particles in Finno-Ugric languages appear in similar functions at earlier stages of the languages, as will be shown in this paper.

9. Vértés (1967:192) suggests that the Estonian use of demonstrative pronouns may be due to the second and third persons becoming similar, because the second person *ti became si through a regular sound change. The demonstrative pronoun might have been introduced for clarity's sake. There is, however, no evidence that this is what happened historically.

10. Older scholars reconstructed tense markers for the Proto-Finno-Ugric language. Itkonen (1962) reconstructs the present tense marker *k, the preterite marker *j and *s̄. Décsy reconstructs *i(j) for the past tense, and a zero morpheme for the present tense (Décsy

1965:158). Newer evidence shows, however, that the proto-language marked aspects, but not tense. The aspect markers have reflexes in the tense markers of the daughter languages, as well as in the person markers in the subjective conjugation (e.g., Rédei 1966).

11. In Livonian, another Balto-Finnic language, the third person singular present tense marker b was generalized into the first person: soob (Old Finnish saapi) '(he) gets' became to be used also in ma soob 'I get'. The original first person singular marker *m > n was lost in Livonian, and the forms ma soo 'I get' was still common in 1920s among the older people (Tauli 1966:67).

12. I unfortunately have data only from Vogul and Ostyak. Neither of these languages show a reflex of *ka in the dual forms. Vogul has for example the following dual forms as possessive suffixes: 1. -m; 2. -n; 3. -ten, and as personal pronouns: 1. meen; 2. neen; 3. teen. Ostyak shows very similar reflexes.

13. The grammar books often define the use of the objective conjugation as follows: the objective conjugation is used with transitive verbs when the direct object is 1) a proper name; 2) a noun which has as its modifier a definite article, a possessive suffix, a demonstrative pronoun, or any other pronoun which indicates a specific object or has an all-inclusive meaning; 3) a personal pronoun in the third person singular or plural, a reflexive, reciprocal, possessive, or demonstrative pronoun or one of the indefinite and interrogative pronouns which indicate a specific object or have an all-inclusive meaning (Steinitz 1950).

14. Steinitz (1950) also indicates that the use of the objective conjugation in Ostyak is not obligatory. As an exception he mentions the case where the definite object immediately precedes the verb. With the SOV word-order, the definite object not being separated from the verb, the subjective conjugation is used. This seems to correspond to the Hungarian usage.

15. "According to the most probable hypothesis the use of -t as the mark of the object developed in the separate life of Hungarian from a determining element. At the beginning it only showed the definiteness of the object-word. Later, when the objective conjugation developed, the marking of definiteness shifted from the object-word to the verbal form, the element -t only had an objective function and as such it also spread to the indefinite object. The objective ending -t can be found in our early records: 1200: adamut, archangelt... Its use spreads in the course of the whole history of Hungarian and the scope of the old original object without an ending constantly decreased." (Károly (1972: 99)).

16. The marking of noun plurals is more restricted in the FU languages than in English; e.g., after numerals and plural modifiers, such as many, the nouns occur in the singular.

17. According to Brugmann's description of the system of demonstratives in Armenian, Mordva seems to have a remarkably analogous system. Armenian has demonstratives which are attached to nouns, verbs, and pronouns. According to Brugmann, these demonstratives function as "personal articles," e.g., ter-s 'the gentleman here' or 'this gentleman' or 'I, the gentleman'; ter-d 'the gentleman there' or 'you, the gentleman' (Brugmann 1904).

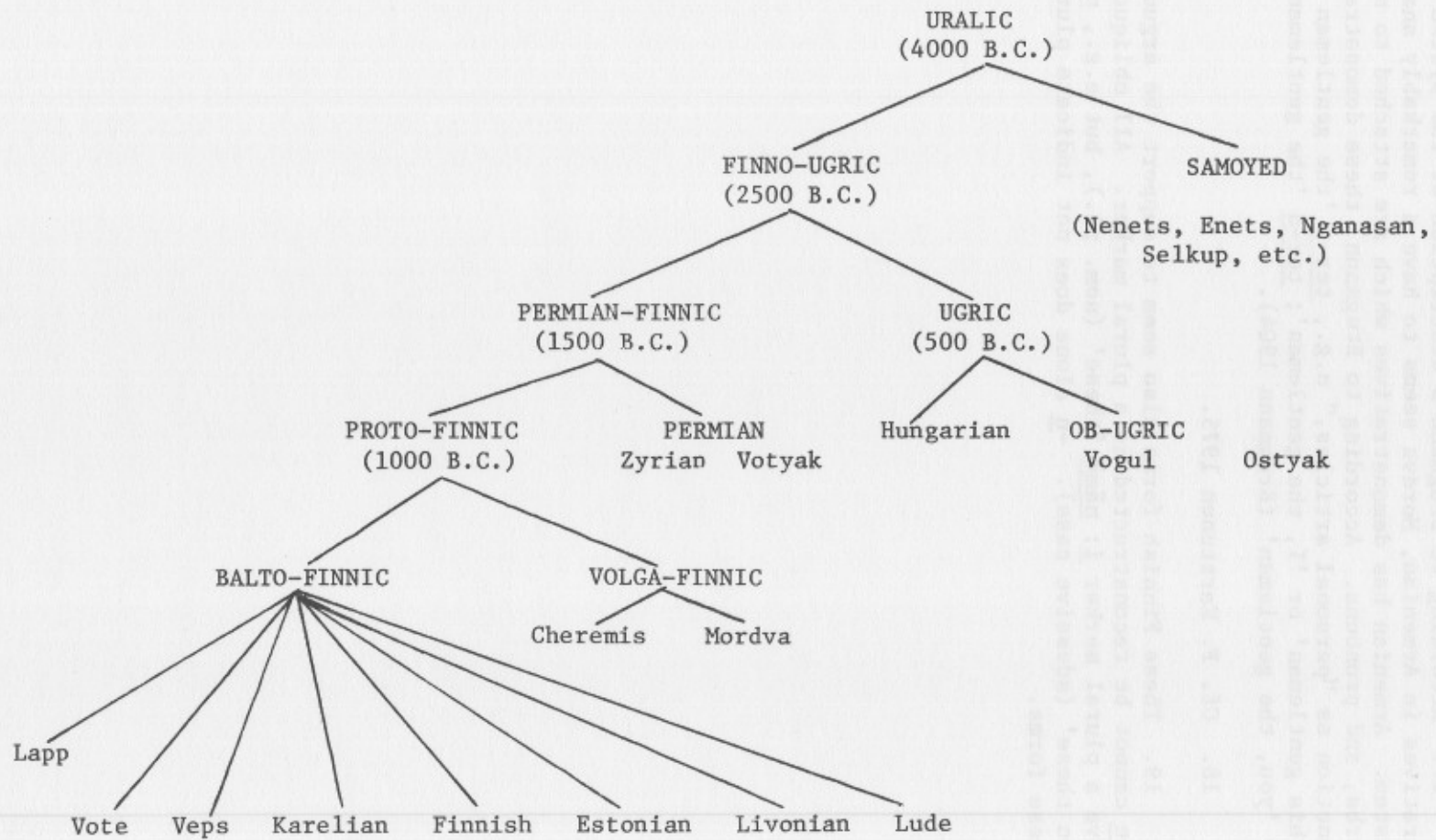
18. Cf. F. Karttunen 1975.

19. These Finnish forms also seem to support the argument that *-n cannot be reconstructed as a plural marker. All oblique case-forms have a plural marker i: nämä 'these' (nom. pl.), but e.g., nä-i-llä 'on these' (adessive case). -n alone does not indicate plurality in these forms.



APPENDIX I

Finno-Ugric Family Tree



APPENDIX II

Tables 1-3

Note on the Tables: The formal variations of the suffixes are first of all connected with the rules of vowel harmony. The explication of the morphophonemic alternations in individual languages is beyond the scope of this paper. Furthermore, no handbook gives full paradigms of the person markers in all nine languages discussed in this paper. The data presented in the tables are gathered from different sources. The possible discrepancies will not, however, affect the basic argumentation.

Language	Person	1st Person		2nd Person		3rd Person	
		sg	pl	sg	pl	sg	pl
Arabic	1st sg	an	ana	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum
Armenian	1st sg	an	an	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum
Azerbaijani	1st sg	an	an	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum
Chechen	1st sg	an	an	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum
Dagestani	1st sg	an	an	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum
Georgian	1st sg	an	an	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum
Ingush	1st sg	an	an	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum
Kabardian	1st sg	an	an	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum
Ossetian	1st sg	an	an	ant	antun	hu	hum
	2nd sg	ant	antun	ant	antun	hu	hum

TABLE 1. PERSON MARKERS IN THE LANGUAGES

TABLE 1. PERSON MARKERS: 1st PERSON

Language		Singular			Plural		
		px	Vb suffix	Pers. pron.	px	Vb. suffix	Pers. pron.
Balto-Finnic	FINNISH	ni	n	mi-nä	mme	mme	me
	ESTONIAN	(ni)	n	mi-na~ma	-	me	meie~me
Volga Finnic	MORDVA	m	n~n'	mon~mon'	mok/nok	nek/nok/nuk	m'in'
	CHEREMIS	m	m	mən'e mə(j) min' etc.	na~nä	na~nä	me/mä
Permian	ZYRIAN	m	m~∅	me	nym	mö	mi~mije
	VOTYAK	m~∅	m~∅	mon/min	my	mɛ~my	mi
	HUNGARIAN	m	m	én	nk	unk/juk etc.	mi
Ob-Ugric	VOGUL	m	m	ɛm, am, ɔm, etc.	w	we	man~män etc.
	OSTYAK	m	em/əm	ma(n)	(e)w	eu/ou ew	mon~men

TABLE 2. PERSON MARKERS: 2nd PERSON

Language		Singular			Plural		
		px	Vb suffix	Pers. pron.	px	Vb suffix	Pers. pron.
Balto-Finnic	FINNISH	si	t	si-nä	nne	tte	te
	ESTONIAN	(si)	d	sina-sa	-	te	teie~te
Volga Finnic	MORDVA	t	t~t'	ton~ton'	nk, ŋk	de~do	tin'
	CHEREMIS	t~d	t	tən/təje	da~dä ta~tä	ta~tä da~dä	ta/te
Permian	ZYRIAN	t~d	n	te	nyd	nyd	ti/tije
	VOTYAK	d	d	tin~ton	dy	d(y)	ti
	HUNGARIAN	d	d	te	tek~tok	(ja)tok~tek	ti
Ob-Ugric	VOGUL	n	n	nag nei, etc.	(ə)n	ne	nan~nän, etc.
	OSTYAK	(e)n	(e)n/(ə)n	nan/noŋ etc.	(ə)n	(ə)n/ten /ələn	nin/nan, etc.

TABLE 3. PERSON MARKERS: 3rd PERSON

Language		Singular			Plural		
		px	Vb. suffix	Pers. pron.	px	Vb. suffix	Pers. pron.
Balto-Finnic	FINNISH	nsa~nsä	V [+long]	hän/se	nsa~nsä	vat~vät	he/ne
	ESTONIAN	(sa)	b	tema-ta	-	vad	nemad~nad
Volga Finnic	MORDVA	zo	i/ś	son~son´	st, sk	t'/s't'	sin'~śín
	CHEREMIS	žo~žə etc.	žo etc.	tudo/tuǒo	što~štə	t/št	tudəβlak
Permian	ZYRIAN	s	∅ (s)	se, si (je)	nys	ny(s)	na(je)
	VOTYAK	s~z	∅/z	so	zy	∅/z	sojos
	HUNGARIAN	(j)e etc.	(j)a	ö	(j)ük etc.	ják	ök
Ob-Ugric	VOGUL	t/∅	∅	tau etc.	n/l	t	tan~tän
	OSTYAK	(ə)t/l	te/li/∅	tuw luw joy lou teu	(e)t	əl ət et	ti(γ) li luw etc.

Bibliography

- Anttila, Raimo. 1972. An introduction to historical and comparative linguistics. New York: The Macmillan Company.
- _____. 1973. Internal reconstruction and Finno-Ugric (Finnish). Current trends in linguistics, Vol. 11, ed. by Thomas A. Sebeok, 317-353. The Hague and Paris, Mouton.
- Bárczi, G. 1963. Zum Sprachgeschehen der urungarischen Zeit. Congressus Internationalis Fenno-Ugristarum, ed. by G. Ortutay, 27-47. Budapest: Akadémiai Kiadó.
- Benkő, Loránd, and Samu Imre, eds. 1972. The Hungarian language. The Hague: Mouton.
- Bése, L., L. Dezső, and J. Gulya. 1970. On the syntactic typology of the Uralic and Altaic languages. Theoretical problems of typology and the northern Eurasian languages, ed. by L. Dezső, and P. Hajdú, 113-28. Amsterdam: Grüner.
- Brugmann, Karl. 1904. Die Demonstrativpronomina der indogermanischen Sprachen. Abhandlungen der philologisch-historischen Klasse der königlich-sächsischen Gesellschaft der Wissenschaften XXII. Leipzig: B. G. Teubner.
- Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. Subject and topic, ed. by Charles N. Li, 25-55. New York: Academic Press.
- Collinder, Björn. 1957. Survey of the Uralic languages. Grammatical sketches and commented texts with English translations. Stockholm: Almqvist & Wiksell.
- _____. 1960. Comparative grammar of the Uralic languages. Stockholm: Almqvist & Wiksell.
- _____. 1965. An introduction to the Uralic languages. Berkeley and Los Angeles: University of California Press.
- Comrie, Bernard. 1975a. Subjects and direct objects in Uralic languages. To appear in Etudes Finno-Ougriennes 12.
- _____. 1975b. The antiergative: Finland's answer to Basque. CLS 11. 112-121. Chicago: Chicago Linguistic Society.
- Décsy, Gyula. 1965. Einführung in die finnisch-ugrische Sprachwissenschaft. Wiesbaden: Otto Harrassowitz.
- Fillmore, Charles. 1975. Santa Cruz lectures on deixis (1971). Bloomington: Indiana University Linguistics Club.

- Forchheimer, Paul. 1953. The category of person in language. Berlin: Walter de Gruyter & Co.
- Hajdú, Peter. 1960. The form of the object in Forest Yurak. *Acta Linguistica Academiae Scientiarum Hungaricae*, 10, 95-113.
- _____. 1972. The origins of Hungarian. *The Hungarian Language*, ed. by Loránd Benkő and Samu Imre, 15-48. The Hague: Mouton.
- _____. 1975. *Finno-Ugrian Languages and Peoples*. London: Andre Deutsch.
- Hakulinen, Auli, and Lauri Karttunen. 1973. Missing persons: on generic sentences in Finnish. CLA 9.157-171. Chicago: Chicago Linguistic Society.
- Hakulinen, Lauri. 1957, 1960. *Handbuch der finnischen Sprache 1-2*. Wiesbaden: Otto Harrassowitz.
- Hetzron, Robert. 1975. The presentative movement or why the ideal word order is V.S.O.P. *Word order and word order change*, ed. by Charles N. Li, 346-388. Austin: University of Texas Press.
- Itkonen, Erkki. 1962. Die Laut- und Formenstruktur der finnisch-ugrischen Grundsprache. *Ural-altaische Jahrbücher* 34, 187-210.
- Károly, Sándor. 1972. The grammatical system of Hungarian. *The Hungarian Language*, ed. by L. Benkő and Samu Imre, 85-170. The Hague: Mouton.
- Karttunen, Frances. 1975. Functional constraints in Finnish syntax. *Papers from the parasession on functionalism*, 232-243. Chicago: Chicago Linguistic Society.
- Lakoff, Robin. 1974. *Remarks on This and That*. CLS 10.345-356. Chicago: Chicago Linguistic Society.
- Lyons, John. 1971. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Mark, Julius. 1925. Die Possessivsuffixe in den uralischen Sprachen. I Hälfte. *Suomalais-ugrilaisen seuran toimituksia* 54, 1-279.
- Moravcsik, Edith A. 1969. Determination. *Working Papers on Language Universals* 1. Stanford University, 63-130.
- Posti, Lauri. 1953. From Pre-Finnic to Late Proto-Finnic. *Finno-Ugrische Forschungen* 31, 1-91.
- _____. 1963. Fragen der ostseefinnischen Verbalflexion. *Congressus Internationalis Fenno-Ugristarum*, ed. by G. Ortutay, 216-225. Budapest: Akadémiai Kiadó.

- Raun, Alo, and Andrus Saareste. 1965. Introduction to Estonian linguistics. Wiesbaden: Otto Harrassowitz.
- Rédei, K. 1966. Die Entstehung der objektiven Konjugation im Ungarischen. Acta Linguistica Academiae Scientiarum Hungaricae 16, 111-133.
- Ruoppila, Veikko. 1967. Suomen kielen historiallista äänneoppia. Helsingin yliopisto ms.
- Sankoff, Gillian, and Penelope Brown. 1976. The origins of syntax in discourse. A case study of Tok Pisin relatives. Language 52, 631-666.
- Sebeok, Thomas A. and Frances J. Ingemann. 1961. An Eastern Cheresis Manual. Bloomington: Indiana University Publications.
- Serebrennikov, Boris A. 1973. Strukturmerkmale der frühuralischen Grundsprache. Ural-altaische Jahrbücher 45, 65-79.
- Steinitz, Wolfgang. 1950. Ostjakische Grammatik und Chrestomathie. Leipzig, Otto Harrassowitz.
- Szinnyei, Josef. 1922. Finnisch-ugrische Sprachwissenschaft. Berlin and Leipzig: Walter de Gruyter & Co.
- Tauli, Valter. 1966. Structural tendencies in Uralic languages. The Hague: Mouton.
- Vennemann, Theo. 1975. An explanation of drift. Word order and word order change. ed. by Charles N. Li, 269-305. Austin: University of Texas Press.
- Vértes, Edith. 1967. Die ostjakischen Pronomina. The Hague: Mouton.
- Wickman, Bo. 1955. The form of the object in the Uralic languages. Uppsala: Almqvist & Wiksell.

The Elimination of Ergative Patterns of Case-Marking
and Verbal Agreement in Modern Indic Languages

Gregory T. Stump

Introduction.

As is well known, many of the modern Indic languages are partially ergative, showing accusative patterns of case-marking and verbal agreement in nonpast tenses, but ergative patterns in some or all past tenses. This partial ergativity is not at all stable in these languages, however; what I wish to show in the present paper, in fact, is that a large array of factors is contributing to the elimination of partial ergativity in the modern Indic languages. The forces leading to the decay of ergativity are diverse in nature; and any one of these may exert a profound influence on the syntactic development of one language but remain ineffectual in another.

Before discussing this erosion of partial ergativity in Modern Indic, I would like to review the history of what the Indian grammarians call the prayōgas ('constructions') of a past tense verb with its subject and direct object arguments; the decay of Indic ergativity is, I believe, best envisioned as the effect of analogical developments on or within the system of prayōgas.

There are three prayōgas in early Modern Indic. The first of these is the kartariprayōga, or 'active construction' of intransitive verbs. In the kartariprayōga, the verb agrees (in number and gender) with its subject, which is in the nominative case--thus, in Vernacular Hindōstānī:

(1) kartariprayōga:

'aurat chalī. mard chalā.
woman (nom.) went (fem. sg.) man (nom.) went (masc. sg.)

The karmaniprayōga is the 'passive construction' of transitive verbs: the verb agrees in number and gender with its object, which is in the nominative case, while the subject is in the ergative case:

(2) karmaniprayōga:

'aurat-nē ghōṛī māri.
woman erg. mare (nom.) struck (fem. sg.)

'aurat-nē ghōṛā mārā.
woman erg. horse (nom.) struck (masc. sg.)

Finally, the bhāveprayōga is the 'impersonal construction', which is historically used only with intransitive verbs. In the bhāveprayōga, the verb is singular and neuter (or masculine, in those languages having lost the neuter gender), and the subject is ergative, as in Vernacular Hindōstānī:

(3) bhāveprayōga:

'aurat-nē chalā. us-nē chalā.
woman erg. went (masc. sg.) he erg. went (masc. sg.)

(Examples from Grierson (1916: IX.I.51-52))

1. The history of the prayōgas.

These prayōgas are, in some form or another, as old as attested Indic. In Vedic and especially in epic Sanskrit, there was a tendency to use the past passive participle in -ta (with or without the copula) in place of finite preterit verb forms (Whitney 1889:362, Bloch 1906). This past passive participle or verbal adjective could be derived from any verb, whether transitive or intransitive; in the latter case, the participle was less passive in meaning than merely preterital-ukta 'spoken', but gata 'gone' (see Whitney 1889:340). Thus, past passive participial sentences could stand as active intransitive sentences and as passive transitive sentences--like any other adjective, this participle agreed with its subject in number and gender in such constructions:

(4) rāmo gataḥ (asti)
Rama (nom.) gone (masc. sg.) is

(5) rāmeṇa pustakam paṭhitam (asti)
Rama (instr.) book (nom.) read (neut. sg.) is

Some few transitive verbs could also be used actively:

(6) devadatta odanam prabhuktaḥ
Devadatta (nom.) porridge (acc.) enjoyed (masc. sg.)
(asti)
is

Fairly late on in the history of Sanskrit, an impersonal construction rose to prominence with the past passive participle of an intransitive verb in the neuter singular and the subject in the instrumental case:

(7) rāmeṇa gataṁ (asti)
Rama (instr.) gone (neut. sg.) is

It's likely that this impersonal construction resulted from an extension of the passive construction exemplified by (5) to intransitive verbs (Renou 1930:498; Bloch 1906:58-9). Perhaps such transitive verbs as prabhuj (see (6)) provided for this analogical extension--

- (8) devadatta (nom.) odanaṁ (acc.) prabhuktaḥ (masc. sg.):
devadattenaudanaṁ prabhuktaṁ (neut. sg.):
(instr.) (nom.)
rāmo (nom.) gataḥ (masc. sg.): X
X = rāmeṇa (instr.) gataṁ (neut. sg.)

In any event, the historical basis of the three prayōgas is clearly reflected in the Sanskrit participial constructions exemplified in (4), (5), and (7). It is no more than reflected, however, since the modern Indic languages aren't directly descended from the classical language, but from its sister dialects; nevertheless, since the germ of the prayōgas is attested even in Vedic (from which the modern Indic languages, as well as Sanskrit, do ultimately descend), we can rest assured that the Sanskrit reflection is an accurate one.

The emergence of the three prayōgas in Middle Indic is also clear, even if many of our conclusions regarding this development must be drawn from texts whose language is an artificial abstraction from spoken Prakrits. In the earlier Prakrits, such as Pali and Jaina Prakrit, the equivalents of constructions (4), (5), and (7) were still treated as participial, but, since the Old Indic preterit tenses were starting to disappear--the imperfect and the aorist had fallen together, and the perfect had virtually vanished (Beames 1879:8-20; Bloch 1965:228-9; Grierson 1916:IX.I.50-51; Hoernle 1880:217; Sen 1960:143)--the reliance on participial constructions in preterit contexts was snowballing (Bloch 1965:234). The classical Prakrits such as Mahārāṣṭrī and Śaurasenī, regularly expressed the past tense participially (Beames 1879:23); and by late Middle Indic, the Apabhraṁśa dialects had retained no other means of expressing it (Beames 1879:26-27; Tagare 1948:282,316-19; Sen 1960:164). Thus, by the end of the Middle Indic period, the descendant of the Old Indic past passive participle had become functionally integrated into the verbal system--that is, it had come to provide the basis for a number of preterit conjugations in late Middle Indic (these conjugations are referred to as participial tenses, whether they are periphrastic or synthetic, in Modern Indic). As a consequence, the three prayōgas had become established as the means of organizing sentences in the participial tenses; the Indic languages had become partially ergative. This late Middle Indic ergativity may be schematized as follows:

(9)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal Inflection</u>
	kartariprayōga	nominative	--	number, gender of subject
	karmaṇiprayōga	ergative (former instrumental)	nominative	number, gender of object
	bhāveprayōga	ergative	--	neuter singular

It is at this stage, during the transition from Middle Indic to Modern Indic, that diverse forces began to erode this partial ergativity, in spite of a few conservative tendencies.

2. Conservative and eliminative tendencies in Modern Indic.

I would now like to survey both the conservative and the eliminative tendencies according as their effect is to reinforce or eliminate ergative characteristics of object and subject case-marking, and of verbal agreement.

2.1. The transitive impersonal construction.

Early on in their modern development, nearly every Indic language begins using transitive verbs with explicit objects in a construction clearly derivative of the bhāveprayōga (Chatterji 1926:897). In this secondary construction, the subject is ergative, the verb impersonal (neuter or masculine singular), and the object in the dative or accusative (hereafter, oblique) case; in many languages, this construction may only be used when its direct object is definite (and in some cases, animate). Thus, we find in Hindōstānī (examples have in some instances been altered to eliminate major orthographic inconsistencies):

(10) 'aurat-nē ghōṛē-kō mārā.
woman erg. horse obl. struck (masc. sg.)

'aurat-nē ghōṛī-kō mārā.
woman erg. mare obl. struck (masc. sg.)

(Grierson 1916:IX.I.52)

in Marāthī:

(11) ghōḍyā-lā mī sōḍilē.
horse obl. I (erg.) loosed (neut. sg.)
(masc.)

pōthī-lā mī vācilē.
book obl. I (erg.) read (neut. sg.)
(fem.)

(Hoernle 1880:327)

in Kumaunī, a Central Pahārī language:

- (12) maī-le wī - kaṇi m̄aro.
I erg. she obl. struck (masc. sg.)

(Grierson 1916:IX.IV.147)

and in East Panjābī:

- (13) óne dūjīaṇ kitābaṇ nuṇ meṇ te rəkkhya
he (erg.) other books obl. table on put (masc. sg.)

(Shackle 1972:82)

Again, this construction is the rule in early Modern Indic (although it has since become obscure in the more innovative languages).

The analogical creation of this impersonal transitive construction is apparently based on a pattern established in nonparticipial (i.e. accusative) tenses; for example, in Vernacular Hindōstānī, the impersonal transitive sentence (14) fulfils the analogical proportion 15 : 16 :: 17 : x.

- (14) us-ne is ciṭṭhī-ko likhā.
she erg. this letter obl. wrote (masc. sg.)
(fem.)

- (15) voh bol rəhī hōy.
she (nom.) is speaking (3rd sg. fem.)

- (16) voh is ciṭṭhī-ko likh rəhī hōy.
she (nom.) this letter obl. is writing (3rd sg. fem.)

- (17) us-ne bolā.
she erg. spoke (masc. sg.)

(Cf. Harley 1944:32-33)

This newly-created construction is clearly eliminative of an ergative characteristic of direct objects: it allows direct objects in ergative contexts to be case-marked exactly as they are in accusative contexts.

2.2. Conservative trends.

Despite this first very general blow dealt to Modern Indic ergativity, several of the languages have, in their modern development, shown tendencies to retain ergative features of case-marking and verbal agreement. I shall survey these tendencies as they are manifested in Hindī, Gujārātī, and Marāthī.

karmaṇipr.: krpā kelī tumhī.
pity (nom.) done (3rd sg. fem.) you (erg.)
(fem.)

bhāvepr.: arjunē mhanitalē.
Arjuna (erg.) said (3rd sg. neut.)

trans. imp.: tyā-nē rāmās mārīlē.
he erg. Ram (obl.) struck (3rd sg. neut.)

(Bloch (1914:260-61))

Thus, the extended use of personal endings reinforces the ergativity of verbal agreement in Marāṭhī in its most conservative usage. Furthermore, an innovative construction found in contemporary usage results from a levelling of the transitive impersonal construction similar to the Gujarātī levelling just discussed, with the exception that in Marāṭhī, the formerly impersonal verb comes to agree with its object not only in number and gender, but also in person (although the object remains oblique, as in Gujarātī). Thus, colloquial Marāṭhī allows all three of the following constructions;

(20) karmaṇipr.: tyānē āplā mulgā
he (erg.) own (nom.) son (nom.)

śālēt pāṭhaviḷā.
school (loc.) sent (3rd sg. masc.)

trans. imp.: tyānē āplyā mulās śālēt pāṭhaviḷē.
(erg.) (obl.) (obl.) (loc.) (3rd sg. neut.)

trans. ex-imp.: tyānē āplyā mulās śālēt pāṭhaviḷā.
(erg.) (obl.) (obl.) (loc.) (3rd sg. masc.)

(Bloch (1914:262))

This construction is standard in the western Marāṭhī dialects Konkan (Grierson 1905:VII.67) and Koṅkaṇī (Grierson 1905:VII.170, Katre 1966:169) and is apparently spreading eastward. Here again, the tendency seems to be towards the reinforcement of ergativity in verbal inflection.

To summarize what has been seen in this section: there are evidently some tendencies to conserve partial ergativity in a few Modern Indic languages. I have discussed a tendency to maintain subjects in ergative contexts in the ergative case, via elimination of the bhāveprayōga (dialectally in Hindī); a tendency to reinforce ergative patterns of verbal agreement through the use of personal endings in the participial tenses (as in Marāṭhī); and a tendency for all transitive verbs to agree with their objects in ergative contexts, at

the expense of the impersonal transitive construction (as in Gujarātī and Marāṭhī).

2.3. Eliminative trends.

I shall now proceed to a consideration of Modern Indic tendencies toward the elimination of partial ergativity. I have already mentioned one such trend, namely the analogical introduction of the impersonal transitive construction, whose effect is to allow direct objects in ergative contexts to be oblique rather than nominative. I shall survey further tendencies of this nature as they occur in Marāṭhī, Nepali, Lahndā, Eastern Māgadhaṇ, and Maithilī.

Surprisingly, many eliminative tendencies are to be found in colloquial Marāṭhī, despite the suggestions of conservativeness discussed in section 2.2.2. First, nonthird person subjects of transitive verbs in participial tenses are often nominative in idiomatic Marāṭhī. When this happens, the verb (which, as always, agrees with its object (which remains nominative) in person, number, and gender) is marked for the number and person of the subject. Thus, in current speech:

(21) tũ kām ke-lē-s.
 thou (nom.) work (nom.) have done (3rd sg. neut;
 (neut.) 2nd sg.)

tũ pothyā lihi-lyā-s.
 thou (nom.) books (nom.) have written (3rd pl. fem. ;
 (fem.) 2nd sg.)

(Bloch (1914:262))

In the Konkani dialect, this agreement of a transitive verb with its subject has been further extended to the third person (Bloch 1914:262). This analogical development based on transitive constructions in the accusative tenses evidently suppresses the distinction between transitive and intransitive subjects and verbal agreement, and therefore contributes in two respects to the elimination of ergativity in Marāṭhī. Furthermore, it gives rise to another idiom, still more radically affecting Marāṭhī ergativity (although limited to a specific--if rather large--set of verbs (see Bloch 1914:263)). In this construction, a participial tense transitive verb agrees in person, number, and gender with its subject, which is nominative, as is its object, with which, however, the verb no longer agrees in any way. Ergativity is thus levelled in favor of accusativity in every respect besides the case of the direct object:

(22) mī tujhī goṣṭ visarlō.
 I (nom.) your story (nom. sg. fem.) have forgotten
 (1st sg. masc.)

(It is also significant that Tibetan never shows verb-object agreement (Grierson 1916:IX.IV.26).) The upshot of this development in colloquial Nepali is that not only is ergativity no longer inherent in verbal inflection, but is no longer held distinct from accusativity in the case-marking of either the intransitive or the transitive subject--that is, the ergative/accusative distinction once maintained in the inflection of (transitive) subjects has become levelled in favor of a transitive/intransitive distinction. Furthermore, the case-marking of the object can no longer be thought to keep ergative constructions distinct from accusative ones, since, both in historically ergative contexts and in accusative contexts, the direct object may be either nominative or oblique (although animate nouns must apparently be oblique--Grierson 1916:IX.IV.25):

(26) mēyle yaslāi phalphūl diē.
I (erg.) him (obl.) fruit (nom.) gave (1st sg.)

(Clark (1977:32))

mēyle tyasko chorālāi kuṭeko chu.
I (erg.) his son (obl.) beaten have

(Grierson (1916:IX.IV.98))

sitale rāmlāi cineko cha.
Sita (erg.) Ram (obl.) has recognized

(Southworth (1967:21))

(27) nāñile tyo ghaḍī phāllā.
baby (erg.) that (nom.) clock (nom.) will knock down

(Clark (1977:226))

inlāi kasari mārḍā-hun.
these (obl.) easily is killing (3rd sg. honorific)

(Grierson (1916:IX.IV.38))

Those constructions exemplified in (26) are in historically ergative contexts; those in (27), in accusative contexts. This confusion of nominative and oblique forms may be in part the result of an analogical extension of the object case-form of the since-levelled karmaniprayōga to historically accusative contexts: as with the levelling of the transitive subject cases to the ergative, Tibetan influence has probably contributed to the confusion (Grierson 1916:IX.IV.24). Thus, if literary Nepali can be said to have retained some vestiges of partial ergativity, the colloquial language certainly cannot.

Both Marāṭhī and Nepali tend toward the elimination of ergativity in verbal inflection; both languages do so by means of an

extension of verbal endings from the accusative tenses to the participial tenses. Interestingly, several other Modern Indic languages have also weakened or eliminated ergative verbal inflection, but have done so by a different strategy.

Lahndā (Western Panjābī), for example, employs such a strategy. Lahndā and Sindhī are unique among Modern Indic languages in their use of pronominal suffixes (from Old Indic enclitic pronouns--Chatterji (1926:970-71)) on both nouns and verbs. In Lahndā, there are two sets of suffixes, one nominative, the other referring to any case (including the nominative) (Grierson 1919:VIII.I.260-61). These endings may be used as or in agreement with any subject or object noun phrase (and double-suffixing sometimes occurs--Grierson (1919:VIII.I.271)). Now, this pronominal suffixation reinforces Lahndā ergativity to the extent that it is used to mark intransitive subjects and transitive objects identically:

(28) (mā̃) jāteu-m.
I (nom.) knew (masc. sg.; 1st sg.)

us (mā̃) mārea-m.
he (erg.) I (nom.) struck (masc. sg.; 1st sg.)

(Grierson (1919:VIII.I.270))

But this suffixation weakens the ergativity of verbal agreement in that it also allows intransitive subjects and transitive subjects to be identically expressed; compare (28) and (29).

(29) (mā̃) usnū mārea-m.
I (erg.) he (obl.) struck (masc. sg.; 1st sg.)

(mā̃) gā dīṭṭhī-m.
I (erg.) cow (nom.) saw (fem. sg.; 1st sg.)
(fem.)

(Grierson (1919:VIII.I.270))

This neutralization of the ergative/accusative distinction with regard to the pronominal suffixation of the Lahndā verb is all the more significant given that transitive subjects in Lahndā often drop their ergative postposition (Chatterji 1926:970) and consequently appear to be oblique in case.

Thus, Lahndā pronominal suffixation sometimes obscures the formal distinction between transitive and intransitive subjects (Sindhī is similar in this respect). The Māgadhan languages show a similar development, but one whose effect has been the virtual elimination of ergativity from this subgroup (Chatterji 1926:971).

The Eastern Māgadhan languages Bengali, Assamese, and Oṛiyā have substantially similar developments. In all three, the patterns of case-marking and verbal agreement of the accusative tenses have become the established patterns of case-marking in historically ergative contexts: subjects are uniformly nominative (although Bengali and Assamese preserve a trace of an ergative suffix in special nominative forms used only as subjects of transitive verbs-- e.g. Bengali santānē 'son' is used as a transitive subject, while santān is used as an intransitive subject); direct objects are uniformly oblique (dative if definite (and in most cases animate), accusative otherwise--Chatterji (1926:897)); and pronominal clitics of recent origin (and therefore not cognate with the pronominal suffixes of Sindhī and Lahndā) have evolved into personal endings by which verbs uniformly agree in person (and number, regularly in Oṛiyā, but irregularly in Bengali, where plural endings are used generally in nonthird persons) with their subjects (although in some dialects of Bengali and Assamese, the third person singular inflection differs according as the verb being inflected is transitive or intransitive--(Grierson 1903:V.I.13, fn 1). It follows that, in these languages, the bhāveprayōga has lost all distinctness from the kartariprayōga:

- (30) Bengali: āmi gelām. (pl.) (Grierson (1903:V.I.384))
Assamese: may gāiṣilō. (Grierson (1903:V.I.444))
Oṛiyā: mū gali. (sg.) (Grierson (1903:V.II.448))
I (nom.) went (1st person)

The karmaniprayōga has become fully accusative in case-marking and verbal agreement patterns:

- (31) Bengali: ghōrā āmi chōṛilām.
horṣe (obl.) I (nom.) loosed (1st person pl.)
(Hoernle (1880:326))
Assamese: xi nasār xabad xunile.
he (nom.) of dancing sound (obl.) heard
(3rd pers.)
(Grierson (1903:V.I.407))
Oṛiyā: sē bājāra śabda ṣuṇilā.
he (nom.) of music sound (obl.) heard (3rd sg.)
(Grierson (1903:V.II.387))

Similarly for the secondary transitive impersonal construction, of which constructions with definite objects are the modern remnant:

- (32) Bengali: ghōṛā-kē āmi chōṛilām.
horse obl. I (nom.) loosed (1st pl.)
'I loosed the horse'

(Hoernle (1880:326))

- Assamese: tār pitek-ak may khobālō.
his son obl. I (nom.) beat (1st person)

(Grierson (1903:V.I.445))

- Oṛiyā: mū tā pua-ku mārili.
I (nom.) his son obl. beat (1st sg.)

(Cf. Grierson (1903:V.II.447, 449))

The Central Māgadhan language Maithilī is somewhat more conservative; but it has come to use an unusual variety of accusative verbal agreement. In modern Maithilī, the kartariprayōga has retained its most general characteristics unchanged since early Maithilī (although number agreement has been lost--Jhā (1958:288-90)):

- (33) ham^a gel^a (chi).
I (nom.) gone (masc.) am/are

- o gelⁱ (achⁱ).
she (nom.) gone (fem.) is/are

(Jhā (1958:542))

The bhāveprayōga, karmaniprayōga, and transitive impersonal construction, on the other hand, are all but levelled in favor of accusative constructions. First, during the modern development of Maithilī, the impersonal transitive construction has supplanted the karmaniprayōga (Jhā 1958:543); this development has had the effect of eliminating the only participial tense construction in Maithilī in which verbs agree with their objects and in which a direct object may be nominative. Subsequently, the impersonal constructions have become accusative: first, personal suffixes of recent origin (from optional pronominal clitics of late development--Jhā 1958:479) are extended from the accusative to the participial tenses (an extension whose recent completion is reflected in a neat age-gradation among present-day Maithilī speakers--Jhā 1958:472, 508); secondly, the ergative case of the subject in these constructions is replaced by the nominative. The result of these developments (whose analogical basis is, no doubt, the established patterns of agreement in the accusative tenses) is that the impersonal constructions have become fully accusative, the only trace of their former impersonality being found in a periphrastic participial tense, the so-called present perfect instantaneous (Jhā 1958:526). Thus, modern Maithilī shows the following intransitive usages:

(34) ham^a gelahū.
I (nom.) went (1st pers.)
(Jhā (1958:472))

ham^a haslahū achⁱ.
I (nom.) laughed (1st pers.) is (3rd pers.)
(Jhā (1958:543))

The suffixation of transitive verbs marks agreement with the subject, as with intransitive verbs; to a transitive verb so marked, however, personal suffixes may further be added in agreement with the direct object (or other oblique objects)--

(35) ham^a tora beṭa-ke { dekhā -i-au.
I (nom.) your son obl. saw (1st pers.; 2nd pers.)
dekhā -i-ai.
saw (1st pers.; 3rd pers.)
(Jhā (1958:473))

ham^a khael-i-ai achⁱ.
I (nom.) ate (1st pers.; 3rd pers.) is (3rd pers.)
(Jhā (1958:543))

Thus, as the result of a historical suppletion of the karmaniprayōga by the impersonal transitive construction, of the loss of the ergative case, and of the introduction of an accusative scheme of verbal agreement into the participial tenses, Maithilī has become a fully accusative language.

To summarize what has been seen in this section: I have examined a number of tendencies eliminative of Modern Indic partial ergativity. These include the total suppletion of the ergative case by the nominative (as in Marāṭhī, Bengali, Assamese, Oriyā, and Maithilī); a confusion of the ergative and oblique cases (as in Lahndā); the transforming of the ergative case into a variant of the nominative case to be used with subjects of transitive verbs (as in Nepali); a concurrence of the nominative and oblique cases in the direct object in both ergative and accusative contexts (as in Nepali); the total suppletion of the nominative case by the oblique in direct object position (as in Bengali, Assamese, Oriyā, and Mathilī); the use of pronominal suffixes on the verb allowing intransitive subjects and transitive subjects to be identically marked (as in Lahndā); the use of personal endings on the verb by which agreement with the subject is expressed, whether to the exclusion of agreement with the object (as in Nepali, Eastern Māgadhan, and sometimes Marāṭhī) or not (as in Maithilī and, generally, Marāṭhī); and the de-personalization of transitive verbs in ergative contexts (historically in Maithilī).

3. Partial categorization of Modern Indic languages.

The conservative and eliminative tendencies discussed in the preceding sections are widely attested in the Modern Indic languages. To summarize these tendencies once again:

Conservative of ergative characteristics

in subject case-marking

(A) the elimination of the bhāveprayōga;

in verbal agreement

(B) reinforcement by the secondary use of ergatively-patterned personal endings;

(C) object-agreement in the transitive impersonal construction;

Eliminative of ergative characteristics

in object case-marking

(D) the introduction of the transitive impersonal construction;

(E) the concurrence of nominative and oblique cases in both (historically) ergative and accusative contexts;

(F) the uniform use of the oblique case for direct objects;

in subject case-marking

(G) the uniform use of the nominative case for subjects;

(H) the use of the (historical) ergative as a nominative of transitive subjects;

(I) the confusion of the ergative and oblique cases;

in verbal agreement

(J) the use of pronominal suffixes to mark subject-agreement uniformly;

(K) the use of personal endings to consistently mark agreement with the subject, whether to the exclusion of object-agreement or not;

(L) the de-personalization of the karmaniprayōga.

I would now like to undertake a classification of the following Modern Indic languages according to their manifestation of any of the above tendencies (hereafter A-L):

Assamese	Panjābī
Bengali	Rājasthānī
Bhīlī	a. Mārwarī
Bihārī	b. Mālvi
a. Bhōjpurī	Sindhī
b. Maithilī	West Hindī
East Hindī	a. Bundēlī
Gujarātī	b. Braj Bhākhā
Khāndēśī	c. Kanaujī
Lahndā	d. Vernacular Hindōstānī
Marāthī	e. Dakhinī Hindōstānī
a. Konkan	
b. Konkanī	
Oriyā	
Pahārī	
West (Jaunsārī)	
Central	
a. Garhwālī	
b. Kumāunī	
East (Nepali)	

(In Appendices I and II, I have indicated the genealogy and geographical distribution of these languages and dialects.)

The Hindi dialects of Bundēlī (Grierson 1916:IX.I.94), Braj Bhākhā (Grierson 1916:IX.I.78), Vernacular Hindōstānī, and Kanaujī (Grierson 1916:IX.I.84), as well as most dialects of Rājasthānī (e.g. Mārwarī, Mālvi--Grierson 1908:IX.II.28, 58), are the Modern Indic languages most conservative of the early Modern Indic system of participial tense prayōgas:

(i) <u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
kartariprayōga	nominative	--	number, gender of subject
karmaṇiprayōga	ergative	nominative	number, gender of object
bhāveprayōga	ergative	--	masculine singular
transitive impersonal construction	ergative	oblique	masculine singular

Only one eliminative characteristic is in evidence in this system, namely (D), the introduction of the transitive impersonal construction.

Standard Hindōstānī, Eastern Panjābī, and the Pahārī dialect of Kumāunī employ a similar system, the difference being the conservative loss (A) of the bhāveprayōga in the latter languages:

(ii)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
	kartariprayōga	nominative	--	number, gender of subject
	karmaṇiprayōga	ergative	nominative	number, gender of object
	transitive impersonal construction	ergative	oblique	masculine singular

It should be noted that, due to the sporadic dropping of the ergative postposition in Eastern Panjābī, this language can be thought to exhibit (I) the confusion of the ergative and oblique cases, as an irregular innovation.

Gujarātī, the Pahārī dialects Jaunsārī and Garhwālī, Bhilī, and Khāndēśī show a similar scheme of participial tense constructions, the difference being the additional incidence in these languages of the conservative development (C), object-agreement in the transitive impersonal construction:

(iii)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
	kartariprayōga	nominative	--	number, gender of subject
	karmaṇiprayōga	ergative	nominative	number, gender of object
	transitive impersonal construction	ergative	oblique	number, gender of object

This system is also irregularly employed in Rājasthānī.

The western languages Sindhī and Lahndā show a similar system, which, however, on the one hand lacks the conservative development (C), and on the other hand incorporates the additional innovations of (J) employing pronominal suffixes on the verb by which subject-agreement may be uniformly marked, and of (I) confusing the ergative and oblique cases (which are, in fact, syncretized everywhere except in the pronominal suffixes in Sindhī):

(iv)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
	kartariprayōga	nominative	--	number, gender of subject; Suffix: person, number of subject
	karmaṇiprayōga	ergative (~ oblique)	nominative	number, gender of object; Suffix: person, number of subject and/or object

(iv) (continued)

<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
transitive impersonal construction	ergative (~ oblique)	oblique	masculine singular; Suffix: person, number of subject and/or object

It should be noted that the innovation of (A) eliminating the bhāve-prayōga hasn't entirely eliminated this construction from Lahndā (cf. Smirnov (1975:112)); further, it should be observed that in Sindhī, different sets of pronominal suffixes are used to mark agreement with intransitive subjects on the one hand and transitive ones on the other. Both languages allow agreement with direct or other oblique objects to be similarly marked.

As was seen above, conservative Marāṭhī retains the early Modern Indic system of participial tense constructions, reinforcing its ergative agreement patterns by an extension of personal endings to the participial tenses:

<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
kartariprayōga	nominative	--	person, number, gender of subject
karmaniprayōga	ergative	nominative	person, number, gender of object
bhāveprayōga	ergative	--	3rd singular neuter
transitive impersonal construction	ergative	oblique	3rd singular neuter

Thus, we find only the eliminative tendency (D) and the conservative tendency (B). But recall that idiomatic Marāṭhī appears to be levelling this scheme through a series of (incomplete) innovations--in addition to (C) the conservative marking of object-agreement on verbs in the transitive impersonal construction, colloquial Marāṭhī also shows the innovative tendencies (G) to use the nominative uniformly as the subject case, and (K) to mark verbs to agree with transitive as well as intransitive subjects (to the exclusion of any object-agreement, for some verbs). The result of these two innovations has been the weakening of contemporary Marāṭhī ergativity, especially in the nonthird persons--that is, the restriction of the karmani and bhāveprayōgas as well as the transitive impersonal construction in favor of accusative constructions. Thus, despite the resemblance of the conservative Marāṭhī schema (v) of prayōgas to that of such conservative languages as Vernacular Hindōstānī (i), modern Marāṭhī is apparently drifting towards a radically reduced schema of participial tense constructions devoid of ergativity:

(vi)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
	kartariprayōga	nominative	--	person, number, gender of subject
	transitive	nominative	{nominative} oblique	person, number, gender of subject

Dakhinī Hindōstānī shows a similar system, although verbs show no personal agreement in this language (Grierson (1916:IX.I.62)).

Literary Nepali has apparently arrived at a very similar stage of development, the differences in Nepali being that transitive subjects remain ergative in participial tense constructions (i.e. innovation (G) is lacking) and that the impersonal constructions are retained as an 'impersonal honorific conjugation' (i.e. the conservative tendencies (A) and (C) aren't fully in evidence) (Grierson (1916:IX.IV.41-43)):

(vii)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
	kartariprayōga	nominative	--	person, number, gender of subject
	transitive	ergative	{nominative} oblique	person, number, gender of subject
	bhāveprayōga	nominative	--	3rd singular masculine
	transitive impersonal construction	ergative	{nominative} oblique	3rd singular masculine

It will be recalled that colloquial Nepali gives evidence of the innovations of (H) treating the ergative case as a version of the nominative to be used with transitive subjects of verbs of any tense and of (E) using either the nominative or the oblique case to mark direct objects, regardless of whether the tense is historically ergative or accusative. The consequence of these developments is that the Nepali system of participial tense constructions appears to be shaping up as in colloquial Marāṭhī.

Maithilī has virtually attained accusativity, although a vestige of the transitive impersonal construction (D) is retained in periphrastic constructions in the participial tenses. Recall that by a suppletion of the karmaniprayōga by the transitive impersonal construction, innovations (F) (the uniform use of the oblique case for direct objects) and (L) (the use of personal endings to consistently

mark agreement with the subject) have become established in Maithilī; by two subsequent developments ((K) and (G)), both the bhāveprayōga and the transitive impersonal construction have become accusative (although both subject- and object-agreement are marked on transitive verbs). Thus, the only remnant of the impersonal constructions in modern Maithilī is the personal inflection of an auxiliary verb in a periphrastic construction. Hence, the following system of participial tense constructions occurs in Maithilī:

(viii)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
	kartariprayōga	nominative	--	gender of subject; (*person of subject)
	transitive	nominative	oblique	person of subject, (person of object); (*3rd person)
	bhāveprayōga	nominative	--	person of subject; (*3rd person)

(*Only in periphrastic constructions)

The participial tense constructions have become fully accusative in Bengali, Assamese, and Oriyā: the case-marking of direct objects (F), of subjects (G), as well as verbal agreement (K) all suggest this:

(ix)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
	kartariprayōga	nominative	--	person, number of subject
	transitive	nominative	oblique	person, number of subject

Eastern Hindī has similarly reduced its formerly ergative system of participial tense constructions (Grierson (1904:VI.5); Chatterji (1926: 971-72)), although it has retained gender agreement between verb and subject (Hoernle (1880: 217, 326)).

The situation is parallel in Bhōjpurī: case-marking and verbal agreement have become fully accusative. Erstwhile pronominal clitics have come to function as obligatory personal endings on the verb: all verbs, intransitive or transitive, agree with their subject (always nominative) in person and, less regularly, in number and gender (the former category being, in common usage, levelled in favor of the plural in the nonthird persons, the latter in favor of the masculine). The original kartariprayōga is preserved in endingless third person singular verbs agreeing with their subject in number and gender.

(x)	<u>Construction</u>	<u>Case of subject</u>	<u>Case of object</u>	<u>Verbal inflection</u>
	kartariprayōga	nominative	--	3rd person: number, gender of subject
	intransitive	nominative	--	person, number, gender of subject
	transitive	nominative	oblique	person, number, gender of subject

Clearly, the overwhelming tendency in these languages is to reorganize the inherited, ergative system of participial tense constructions as an accusative system fully parallel to that found in historically accusative contexts.

4. Analysis.

Thus, many Indic languages give evidence of a general tendency to eliminate ergative patterns of agreement in the participial tenses. I shall now briefly consider the theoretical conclusions to be drawn from this fact.

First, it should be observed that, despite trends in a few languages favoring the conservation of partial ergativity within the participial tenses, no modern Indic language has shown signs of extending ergative agreement patterns to historically accusative tenses. This fact stands in notable contrast to the widespread tendency in Modern Indic languages to level ergativity in favor of accusative patterns of inflection.

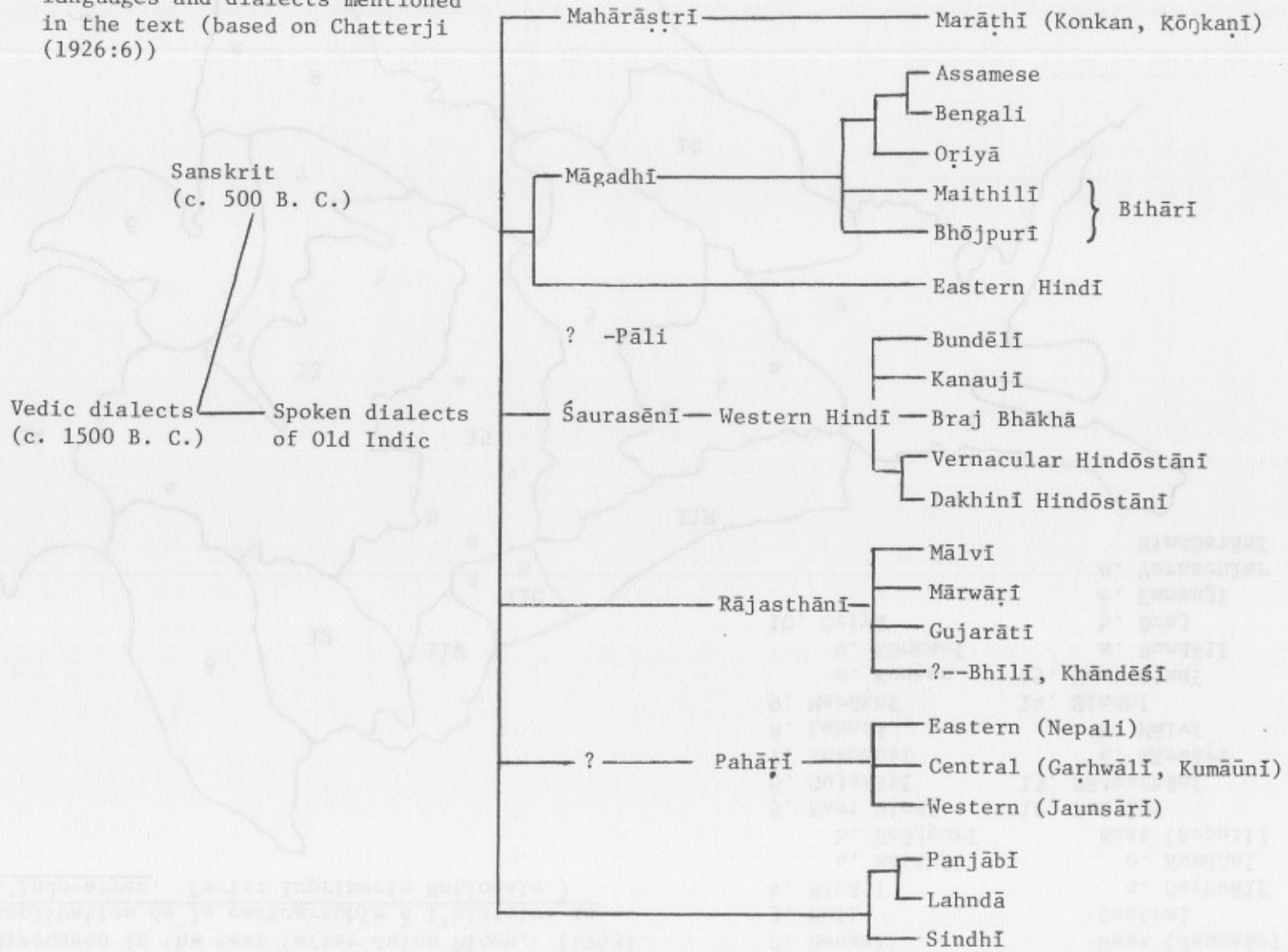
These diachronic observations bear significantly on the question of deep vs. derived ergativity in Indic. If it is indeed true that grammatical change is often motivated by a drive toward derivational transparency, then the evidence seems to suggest that ergativity is a purely derivative relational notion in the Modern Indic languages: if ergativity were, instead, a basic relational characteristic of Indic grammar, we would expect accusativity, if anything, to be levelled out, again in the interests of derivational transparency; but there is no sign of such a development in any of the languages considered. It may seem that I am begging the question of whether ergativity and accusativity might not both be able to be basic relational notions in a single language, of whether partial ergativity might not be able to be as 'deep' as deep ergativity. Such could perfectly well be the case in some language, but not, I believe, in any of the languages I have discussed here; the patterns of relational levelling in Indic are too regularly assertive of accusativity and eliminative of ergativity. Interestingly, the claim (that Indic ergativity is derived) that I am making on diachronic grounds is supported by synchronic evidence: Pandharipande and Kachru (1977) have suggested that ergative patterns of agreement as well as apparent instances of rules sensitive to ergativity

can be explained away on independent, nonrelational grounds in Hindi (in the present context, the Modern Indic language most conservative of partial ergativity).

5. Conclusion.

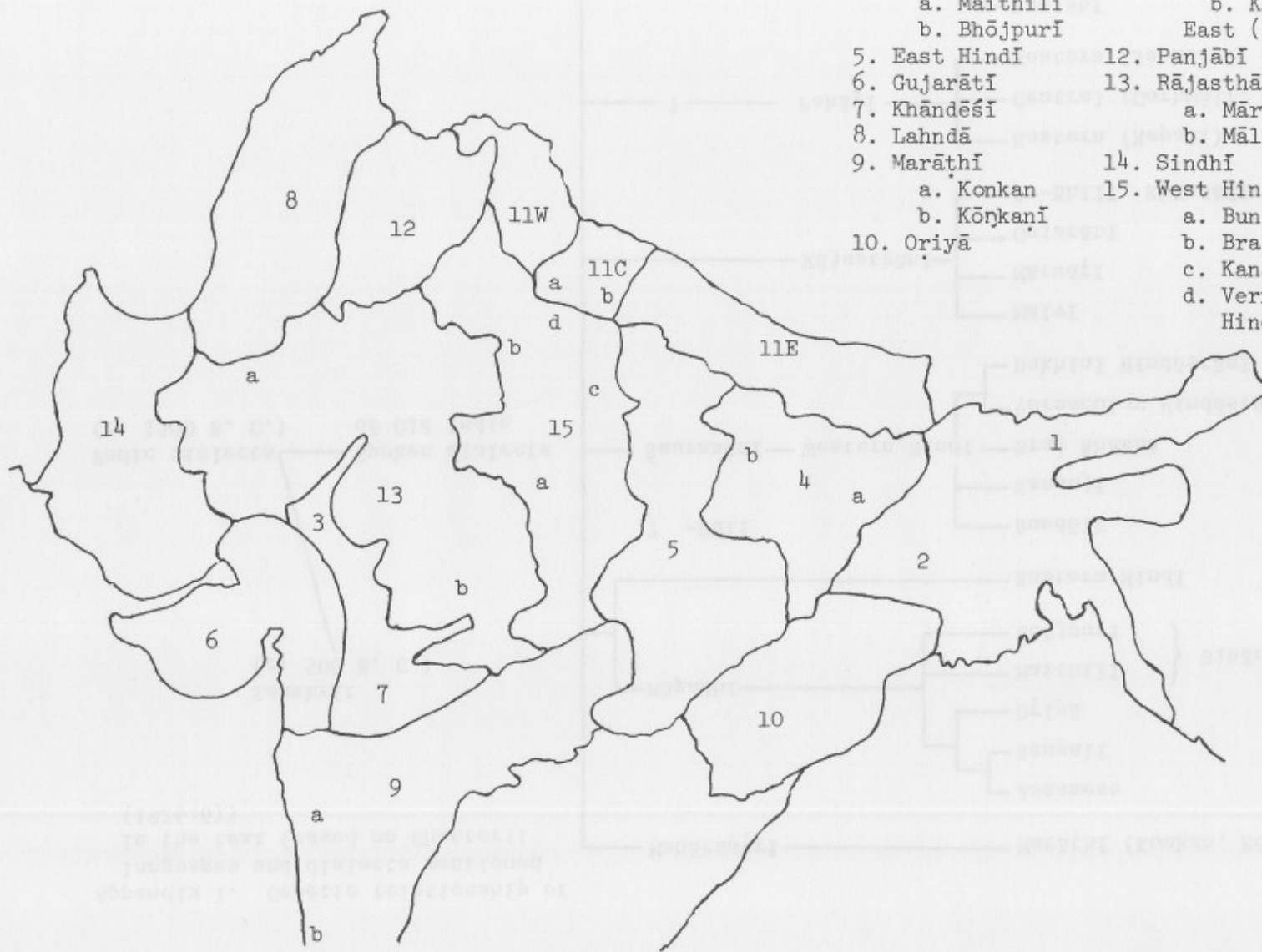
Having examined a broad range of Modern Indic languages, I have demonstrated the remarkable predominance of eliminative (as opposed to conservative) tendencies affecting ergative case-marking and verbal agreement patterns in the participial tenses in these languages. These tendencies are sufficient to be said to constitute a Modern Indic 'drift': given the assumption that Indic partial ergativity is a derivative phenomenon, it is evidently an opaque enough rearrangement of the underlying accusativity of these languages to induce its own elimination by successive generations of language learners.

Appendix I. Genetic relationship of languages and dialects mentioned in the text (based on Chatterji (1926:6))



Appendix II. Modern Indic languages and dialects discussed in the text (after Jules Bloch. (1963). Application de la cartographie à l'histoire de l'indo-aryen. Paris: Imprimerie Nationale.)

- | | |
|---------------|--------------------------|
| 1. Assamese | 11. Pahārī |
| 2. Bengali | West (Jaunsārī) |
| 3. Bhīl | Central |
| 4. Bihārī | a. Garhwālī |
| a. Maithilī | b. Kumāūnī |
| b. Bhōjpurī | East (Nepali) |
| 5. East Hindī | 12. Panjābī |
| 6. Gujarātī | 13. Rājasthānī |
| 7. Khāndēśī | a. Mārwarī |
| 8. Lahndā | b. Mālvi |
| 9. Marāthī | 14. Sindhī |
| a. Konkan | 15. West Hindī |
| b. Kōṅkanī | a. Bundēlī |
| 10. Oriyā | b. Braj |
| | c. Kanaujī |
| | d. Vernacular Hindōstānī |



Bibliography

- Beames, John. (1879). A Comparative Grammar of the Modern Aryan Languages of India, Book III. Delhi: Munshiram Manoharlal.
- Bloch, Jules. (1906). "La Phrase Nominale en Sanskrit," Mémoires de la Société de Linguistique de Paris, XIV. Paris: Honoré Champion.
- Bloch, Jules. (1914). The Formation of the Marāṭhī Language, trans. by Dev Raj Chanana. Delhi: Motilal Banarsidass.
- Bloch, Jules. (1965). Indo-Aryan from the Vedas to Modern Times, trans. by Alfred Master. Paris: Adrien-Maisonneuve.
- Chatterji, Suniti Kumar. (1926). The Origin and Development of the Bengali Language. London: George Allen and Unwin, Ltd.
- Clark, T. W. (1977). Introduction to Nepali. School of Oriental and African Studies, University of London.
- Grierson, G. A. (1903-1919). Linguistic Survey of India, vols. V.I, V.II, VI, VII, VIII.I, IX.I, IX.II, IX.III, IX.IV. Delhi: Motilal Banarsidass.
- Harley, A. H. (1944). Colloquial Hindustani. London: Routledge and Kegan Paul, Ltd.
- Hoernle, A. F. Rudolf. (1880). A Comparative Grammar of the Gauḍian (Aryo-Indian) Languages. Amsterdam: Philo Press.
- Jhā, Subhadra. (1958). The Formation of the Maithilī Language. London: Luzac and Co., Ltd.
- Katre, S. M. (1966). The Formation of Koṣkanī. Poona: Deccan College.
- Lambert, H. M. (1971). Gujarati Language Course. Cambridge University Press.
- Matthews, W. K. (1952). "The Ergative Construction in Modern Indo-Aryan," Lingua 3.391-406.
- Pandharipande, Rajeshwari and Yamuna Kachru. (1977). "Relational Grammar, Ergativity, and Hindi-Urdu." Lingua 41.217-38.
- Renou, Louis. (1930). Grammaire Sanscrite. Paris: Adrien-Maisonneuve.
- Scholberg, H. C. (1940). Concise Grammar of the Hindi Language. Oxford University Press.
- Sen, Sukumar. (1960). A Comparative Grammar of Middle Indo-Aryan. Poona: Linguistic Society of India.
- Shackle, C. (1972). Punjabi. London: Teach Yourself Books.
- Smirnov, U. A. (1975). The Lahndi Language, trans. by E. H. Tsipan. Moscow: Nauka.
- Southworth, Franklin C. (1967). Nepali Transformational Structure: A Sketch. Poona: Deccan College.
- Tagare, Ganesh Vasudev. (1948). Historical Grammar of Apabhramsa. Poona: Deccan College.
- Whitney, William Dwight. (1889). Sanskrit Grammar. Cambridge, Mass.: Harvard University Press.

