

**On Projecting Variation Back into a Proto-Language, with Particular  
Attention to Germanic Evidence**

Brian D. Joseph

The Ohio State University

**1. Introduction**

Reconstruction of aspects of a proto-language almost always entails some degree of idealization. For instance, it is generally the case that the outcome of reconstruction has a neat and tidy look to it that belies the often-made claim — used for instance by typologically inclined linguists to justify a reconstruction based on observed properties of attested language — that a proto-language is a natural human language no different in kind from any other language. If such is the case, then the absence of “loose ends” in the reconstructed language state makes it quite unlike known languages, all of which show variation.

Linguists have been aware for some time of this uneasy relationship between awareness of synchronic variation and the prerequisites of historical reconstruction, and the topic is a matter of some discussion in the literature. In this regard, the instructive remarks in a recent textbook on reconstruction, Fox (1995: 51), are worth quoting at length:

[the] insights [of Labov, Milroy, and others concerning the role of variation in language change] are potentially of great importance for linguistic reconstruction; in practical terms, it has proved difficult to incorporate them into the methodology of reconstruction, and indeed even to reconcile them with the

assumptions upon which this methodology is based. In the first place, the Comparative Method assumes — indeed depends on — the uniformity of reconstructed languages, which is at odds with the inherent variability assumed by many sociolinguists.

Fox (1995: 52) resolves this dilemma by asserting that “comparative reconstruction and sociolinguistic work on language change are concerned with rather different phases of the process of change”, the former more with the “*mechanisms* of change” and the latter more with the “interpreting its *results*”. He later notes (p. 140) that “we are *not* entitled ... to mistake our idealizations [of a uniform proto-language] for reality”.

This largely programmatic and synthetic paper is intended as a contribution to the discussion about the ways in which there can be a recognition of variation in reconstruction, both as a realistic goal of the reconstruction process, i.e. a goal that can be achieved, and as a goal of realistic reconstruction, i.e. reconstruction that mirrors what a natural human language is like (akin to Fox’s reference to “reality”). That is, some idealization of a proto-language speech community might certainly be a methodological requirement, as emphasized by Dyen 1969, but so also should there be, of necessity, a role for variation in the process.

Several pertinent general questions can be raised in this regard:

- When is it valid to invoke proto-language variation as a solution to a problem in comparative reconstruction?
- What is meant by “variation” in this context?
- Are some types of “variation” more amenable than others to being projected back into a proto-language?

In the sections that follow, these questions, and related ones, are addressed, and though there are no startling conclusions coming out of these programmatic remarks, there is some value to reviewing long-standing issues in linguistic reconstruction pertaining to variation. Moreover, the discussion is illustrated with numerous examples, many drawn from Germanic, and these examples can be viewed as offering fresh new perspectives on old, but still highly relevant and definitely intriguing, methodological matters.

## **2. Methodological Preliminaries and some (Relatively) Easy Case Studies involving Phonology**

Since the focus here is necessarily on methodology, it is fair to start by reviewing something that is well known, but which bears repeating, about methods. The best method for reconstruction and for reaching some conclusions about a proto-language is, as suggested by Fox's remarks above, the Comparative Method. In a very real sense, the Comparative Method is always confronted with a "crisis" — i.e., a situation in need of being resolved — whenever there is a certain type of variation in the data, namely a mismatch between two (or more) correspondence sets for ostensibly the same element in different languages. While it is of course tricky to determine how to apply the notion "ostensibly the same element", most often, such cases can be dealt with easily, if the analyst simply draws on the usual armamentarium that historical linguists have at their disposal.

For instance, as an example of a straightforward case, consider the following correspondence sets — Greek **p** (e.g. **patér** 'father') : Germanic **f** (e.g. **father**) vs. Greek **p** (e.g. **spora** 'sowing') : Germanic **p** (e.g. **sprawl**). Several ways of

accounting for these correspondence sets and the variation they suggest might be contemplated if one is inclined to think in terms of proto-language variation, although a straightforward, nonvariationist account ultimately is to be preferred. In each case in which variation is posited for the proto-language in order to explain differences between related languages, one has to think in terms of one language generalizing one variant and another language generalizing another. For instance, one might posit a variable proto-language \*p that had a set of unconditioned (i.e., freely) varying realizations in the proto-language, e.g. between aspirated and unaspirated pronunciations, somewhat like the variation between released versus checked final stops in contemporary American English; one of those variants, it would be claimed, developed into Germanic **p**, and the other into Germanic **f** (whereas both yielded Greek **p**). Alternatively, the different results in the different languages might represent sociolinguistically conditioned proto-language variation, where the relevant factors were related to the speech act (who is talking to whom, about what, under what circumstances, with what style, etc.); such a situation would be somewhat like prenasalized versus pure oral voiced stop variation in Modern Greek, where, as discussed in Arvaniti & Joseph 2000, the pure oral realizations are found more with younger speakers and in casual speech.

However, neither of these steps is really called for, and indeed, there is no independent evidence to suggest the need to resort to them, for the mismatch in these correspondence sets can be resolved by attending to purely linguistic environments. That is, the **p:f** set occurs prevocally and **p:p** occurs only after **s**. Thus, one can treat the variant correspondences as phonologically conditioned outcomes of the same proto-language element (say, \*p), where the different

To appear in *Variation and Reconstruction*, ed. by T. Cravens. Amsterdam: John Benjamins Publishing Co. (2005)

environments determine the different outcomes.<sup>1</sup> This differentiation presumably took place on the way from Proto-Indo-European to Germanic, but it cannot be ruled out that the proto-language had conditioned allophones, the difference between which was neutralized in Greek. The important point, though, is that purely linguistic factors are all that need to be summoned forth in this case, and that sort of variation is so ordinary, normal, and expected in natural language as to not require comment. Positing such linguistically determined variation is exactly what the Comparative Method is best at, one could say.

A slightly trickier, yet classic, example is the case of Greek **#h** / **#zd** versus **#y**- in the rest of Indo-European, e.g. **hus**-(**míné**)‘fight’ / **zdúgon** ‘yoke’ versus Sanskrit **yudh-** / **yugam**; here the “ostensible same element” is a phonological unit, as was the case with the **p/p** and **p/f** sets, yet are the same sorts of solutions possible? There are no obvious conditioning factors (e.g. both correspondences occur in initial position before **u**), and positing two different distinct **y**-like elements in proto-language is not very satisfying, inasmuch only Greek would give evidence of this difference; thus, variation may be the solution, but it need not be proto-language variation — rather, the variation may have been within Greek, perhaps sociolinguistically conditioned, with the **zd** outcome found in rural jargon, as suggested by Hock 1991 — and thus not in the proto-language.<sup>2</sup>

Even more difficult yet is the following case from contemporary American English. As first mentioned by Labov 1984, and more recently by Janda, Joseph, & Jacobs 1994, and thereafter by Shapiro 1995 and Lawrence 2000, with considerable discussion in Janda & Joseph 2001ab, there is speaker-to-speaker variation in realization of clusters involving sibilants and stops, e.g. **spray** pronounced as [sprej]

or [sprej].<sup>3</sup> Here, the “ostensible same element” to be determined for reconstructive purposes is the lexical item; that is, [sprej] and [sprej] have to be recognized as being instantiations of the same lexeme but that type of variation poses a problem for the purposes of reconstruction. In particular, to use the (standard) Comparative Method in such a case, a choice has to be made to elevate one variant into the role of the primary element for external comparison; without information about which variant is innovative, the choice would be purely arbitrary, but necessary nonetheless for the method.<sup>4</sup> If the two variants are treated as equals, then one might need to consider positing two distinct elements in the proto-language, even though the ostensible “sameness” of [sprej] and [sprej] renders that step unviable. Finally, in such a case, treating the variability as a reflection of proto-language variation (re)surfacing in English is potentially tricky, since, as with the Greek **h/zd** developments discussed above, it seems to take an English-internal development and project it back into the proto-language. This particular “solution” is revisited below, but first some variation at other levels of analysis deserves to be examined.

### **3. Similar Cases involving Morphology and the Lexicon**

The cases of variation examined so far have all dealt with phonology, but similar points could be made concerning different types of variation discussed above with examples involving words and their components, i.e. morphology and the lexicon.

For instance, Classical Latin made use of the Proto-Indo-European so-called “primary” (i.e., present-oriented) endings in both past tense forms and present tense forms, as in **dat** ‘he gives’ alongside of **dedit** ‘he gave’.<sup>5</sup> This situation contrasts with what is found, for instance, in Sanskrit, where there is a 3SG ending **-t**

(roughly marking past) and another 3SG ending **-ti** (roughly marking present). As an artificial restriction of the data, in order to highlight the essence of the methodology, one can concentrate just on these facts, without adding in any additional information from, e.g., other languages or other stages of the language (see footnote 5).<sup>6</sup> Then there are two morphological correspondence sets to take seriously as a basis for reconstruction, 3SG present **t** : **ti** and 3SG past **t** : **t**, and several possibilities emerge.

First, we could reconstruct an invariant 3SG ending **\*-t**, and posit that Sanskrit innovated a distinction between present and past through the creation of **\*-ti** (presumably from suffixation of a presential marker **\*-i**). While this may be a viable account from the restricted vantage point first adopted here, it turns out not to be appropriate once additional evidence is taken into account: the facts from Greek, Hittite, and Old Irish verbal endings, for instance, among others, make it clear that the **\*-t** versus **\*-ti** distinction is to be reconstructed for Proto-Indo-European and is not simply a Sanskrit creation.<sup>7</sup>

Alternatively, one might take the need to reconstruct **\*-t** and **\*-ti** for Proto-Indo-European and posit variation between the two for the proto-language. Of relevance here is the fact that in Vedic Sanskrit, one finds fluctuation in the so-called “subjunctive” mood (which functions as a future tense) between, e.g. 3SG **-t** and **-ti**, as in **bhavát** / **bhaváti** ‘he will be’. That is, the Vedic fluctuation might be taken as a relic of the proto-language state of affairs. This means, however, that essentially unmotivated variation is what is being reconstructed, or else we just have to make guesses (though they may well be educated guesses deserving of the label “hypothesis”) about what the distinction between **\*-t** and **\*-ti** could have been

To appear in *Variation and Reconstruction*, ed. by T. Cravens. Amsterdam: John Benjamins Publishing Co. (2005)

signalling. That is tantamount to giving up in reconstruction (or even in the analysis of any language from a synchronic standpoint).

Thus, the best solution, as far as the standards of the Comparative Method are concerned, is to assume that Latin neutralized a Proto-Indo-European distinction, most likely one of present versus past, that is preserved in Sanskrit indicative forms. Moreover, this solution is actually indicated by Old Latin, as noted in footnote 5, where the 3SG past **-d** of, e.g., **deded** ‘he gave’ is the regular outcome of Proto-Indo-European **\*-t**, and contrasts with present **-t**, the regular outcome of Proto-Indo-European **\*-ti**. The Classical Latin situation is thus the result of analogical extension of 3SG present **-t** into the original territory of 3SG past **-d**. Admittedly, this leaves the fluctuation in the Vedic Sanskrit subjunctive unexplained, but it could be considered to be a Vedic-internal issue, much as the English **spr-/spr-** fluctuation would appear to be, or possibly also the Greek **h-/zd-** contrast.

Still, it is the best that the method can do, and what this example shows for morphology, as for the phonological cases discussed above, is that the notion of proto-language variation can in principle, if adopted, undermine the whole comparative reconstruction enterprise. One can end up with loose ends in the proto-language that cry out for some explanation.

Thus positing variation for the proto-language in and of itself is not necessarily a desired — or desirable — end, but it can drive one to come up with suitable explanations. A case in point is the classic conundrum of Proto-Indo-European **\*m** versus **\*bh** in dual and plural oblique cases (as in the Sanskrit INST.PL **-bhis** versus Balto-Slavic **-mis**). This has traditionally been “accounted



To appear in *Variation and Reconstruction*, ed. by T. Cravens. Amsterdam: John Benjamins Publishing Co. (2005)

for” by reference to variation within Proto-Indo-European that is unmotivated and unexplained, except possibly to refer to a geographic dialect split within the proto-language (e.g., with a “northern Indo-European” consisting of Balto-Slavic and Germanic having the \*-m- variety). Still, there could be straightforward solutions. For instance, as Hock (1991: 590) suggests, a solution might be found by positing a (pre-)Proto-Indo-European sound change of **bh** → **m** / \_\_\_\_...**m** (cf. the final **m**-element in the Sanskrit dual INST.DAT.ABL **-bhyám**) and generalization of the newly created variant with initial \*-m-. Alternatively, the \*m/bh variation may reflect two distinct elements in the proto-language which had different (but related oblique) functions but which were “redistributed” and generalized in individual languages (so Beekes 1995, where the **-m-** case is dative while the **-bh-** case is instrumental). Either way, then, this could simply be a case like **p/f** above or the 3SG endings, in which a single proto-language element has split or else an original distinction has been redistributed in the individual languages.

Yet another dimension to variation is lexical government, as with Slavic 1SG **-m** vs. **-∅**. This distinction is found all over Slavic in varying degrees (see Janda 1996), e.g. as the normal ending in Slovak, Slovene, Macedonian, and (with two exceptions) Serbo-Croatian; as an ending restricted to a relatively small class of verbs in Bulgarian and Russian; etc. In this case, it is clear that there were two endings in Proto-Slavic (roughly, \*-m vs. \*-ō) and the basis for their distribution was essentially lexical in nature; in Old Church Slavonic, for instance, the \*-m ending occurs with just 4 verbs. Here the mismatches across the Slavic languages with regard to the occurrence of **-m** and **-∅** and the absence of any (obvious) conditioning, as well as the evidence of the oldest attested Slavic language, lead to

the view that this case represents lexically governed proto-language variation. It may be that from an Indo-European perspective, the Proto-Slavic situation is a development out of two endings that originally had a rational basis for their distribution, but the likelihood of variation in a proto-language — here, Proto-Slavic — cannot be denied. Proto-language variation therefore perhaps ought to be taken as a serious possibility, though crucially, such variation in this case is along a different kind of dimension from the previous examples considered heretofore.

#### **4. Some Problems and Consequences for Reconstruction**

Cautionary notes must be sounded, however, since once we begin to tolerate such lexically-governed variation, abuse of the method becomes all too easy. In particular, one has to wonder if there is any problematic case that could not simply be resolved by positing that it was an idiosyncratic fact about the proto-language lexical item involved. For instance, an irregular sound correspondence where one language has an unexpected segment in one word, e.g. the initial **p** in Germanic \*paq- ‘path, road’, where the correspondence with **p** in Greek **patós** ‘path’, among other relevant forms in other languages, has long been recognized as problematic, since Germanic \*f would be expected (as seen above). While the Germanic form is usually taken to be a borrowing (e.g. from Iranian, and note that the vowels do not match up well either), why could one not simply say that this lexical item in Proto-Indo-European fluctuated between initial \*b and initial \*p, and Germanic happened to generalize the \*b- form. In a sense, many Indo-Europeanists have tacitly accepted such “solutions” quite readily, positing formal “root extensions” that have no clear function (e.g. \*yew-/yew-g- ‘join’) and only mechanically “account for”

To appear in *Variation and Reconstruction*, ed. by T. Cravens. Amsterdam: John Benjamins Publishing Co. (2005)

variant root forms in the various languages (e.g. Sanskrit **yu-/yu-j-**, Greek **zdug-**, Latin **iug-** etc.).<sup>8</sup>

Without a motivation for the different forms, via some reasonable hypothesis for a value or function that they might have had, one is really left with no suitable hypothesis and thus no real explanation or accounting for the facts. All one has is unmotivated proto-language variation.

Thus, to a significant degree, the most relevant question is that of when to invoke variation, and further, whether there are some sorts of cases where it is more appropriate to do so and more importantly where there is some external justification for doing so.

## **5. Towards Solutions — Cases where Proto-Variation makes Sense**

As a first step in that direction, it is useful to consider two admittedly highly speculative cases from (West) Germanic. In particular, there are recurrent but seemingly parallel vowel developments to be found in different branches of West Germanic: **ú** > **aw** in the passage from Old English to Modern English and also from Old High German to Modern High German, as seen in **house/Haus**, both from earlier \*hús- and **í** > **ai** also in both the English and the German branches, as seen in **white/weiß**, both from earlier \*hwít. Are the parallel developments just a coincidence, or perhaps governed by universals of vowel changes (e.g. those suggested by Labov 1994)? Or, could they reflect the outcome, via generalization, of proto-language variation in the realization of the vowels, i.e. Proto-West-Germanic **ú/aw** and **í/ai**, with each language happening to generalize the same variant? This is not a hypothesis that can be proven in any strict sense, and one

To appear in *Variation and Reconstruction*, ed. by T. Cravens. Amsterdam: John Benjamins Publishing Co. (2005)

would ideally want to specify what conditioned the variation in the proto-language; moreover, it must be admitted that the chronology here is difficult, since the emergence of **aw/ai** seems to be somewhat “late” within each branch (e.g. post-Middle English, in the case of English). One would have to assume that the variation entered each language from the proto-language but was then “submerged” sociolinguistically for a long period of time, i.e. present but not particularly salient. This would be akin to what , Trudgill et al. (2000: 136), in their discussion of changes in various southern hemisphere varieties of English, describe as situations in which “changes that are already in progress in the common source may be continued after separation”. While not an impossible scenario, there is really only the parallelism of the changes and the relatedness of the languages to even license thinking of positing proto-(West)-Germanic variation that is only manifested later in the different branches.

However, one could similarly look to the retraction of initial **s** to **ʃ** before stops, as discussed above, as in **spr-** > **ʃpr-**, **str-** > **ʃtr-** and so on, for this is also found not only in contemporary English but also dialectally (though widespread) in German. Both changes seem to have **str** clusters as their focal point — so Shapiro 1995, though Labov 1984 suggests that the retraction is nearly categorical in **str** clusters for some younger African-American speakers of English — but are also found with other stops as well.<sup>9</sup> Thus they seem to reflect some commonality of development, and one could posit, as with the vowel changes above, that the **s/ʃ** variation was there already in Proto-West-Germanic, and that it “surfaced” in German at a relatively early point, and in English only somewhat later. Here though the chronology is even more problematic than in the previous case, as this change

To appear in *Variation and Reconstruction*, ed. by T. Cravens. Amsterdam: John Benjamins Publishing Co. (2005)

seems to be quite recent in English, to judge from how and when it has been discussed in the literature, though Lawrence 2000 suggests it may date at least to the 1800s, and certainly its distribution, occurring not only in American English but elsewhere as well (so Lawrence 2000) and found among American speakers of all ages now, is consistent with it being an older feature than its recent notice in the literature might suggest. Thus while suggestive, the proto-language variation account again requires that the variation in question was “submerged” for some time and (re-)emerged as salient at a later point. Still, this assumption may not be too problematic, since this may well be the sort of low-level variation that might never make it into written texts.

Nonetheless, it is interesting to put these two cases, and the three changes they entail, together, for one has to wonder if there is not perhaps a cumulative effect here. Each case above suggesting Proto-West-Germanic variation may be weak on its own. However, taken collectively, the fact that there are three such changes (and maybe others, once one starts to look for them) might well constitute enough of a critical mass to make one think that the standard methods are missing something potentially important about the proto-language in not allowing access to the (potential) reality of proto-language variation.

A case that is even stronger, in that reasonable speculation on the causes of variation can be ventured, is the following. In Old English **sceal/sculon** ‘shall/should’, the sounds represented by orthographic < sc > have given, as expected, [ s ] in most dialects of Modern English. Some northern dialects, however, have [ s ], e.g. **sal / suld** (from 13th century at least). The corresponding verb in German also shows an “irregular” development for the initial element, viz.

Modern German **sollen** (with [z-]). The *Oxford English Dictionary* (s.v.) has the following to say about this:

It is remarkable that a similar form, with **s** irregularly representing OTeut[onic] **sk**, existed as a dialectal variant in OHG (**sal**, **sol**, **sulun**) and OFris[ian] (**sal**, **sel**), and ousted the regular form in Ger[man] (**soll**, **sollen**) and Dutch (**zal**, **zon**). Some scholars regard the **s** form as representing an OTeut. variant, originating from the euphonic dropping of **k** in inflexional forms like the subjunctive \*sklí-. It seems more probable that it was independently developed in the different dialects at an early period, while the **sk**- retained its original pronunciation; in stressless position, the **k** might naturally be dropped, and the simplified initial afterwards extended by analogy to the stressed use.

While the “euphony” of dropping [k] in the cited inflexional forms might be open to question, the invoking of prosody through the reference to stresslessness is significant. Once it is admitted that prosody could have played a role (and note German **Schuld** ‘guilt’ from the same root but as a noun presumably always prosodically strong and never with [#z-]!) and that the generalization of an unstressed form over a stressed form might have occurred, both of which are eminently reasonable assumptions, one has to ask if it is necessary to say these events occurred independently in each of the “different dialects”. Why not treat this as a case of proto-language variation — maybe even as a proto-language fast-speech vs. careful speech phenomenon, or a stylistically conditioned phenomenon in which sentence prosody might well be expected to have played a role? All the elements

needed for such an account are in place, and the fact of a parallel irregularity in English dialects and in German is enough to make this a plausible situation in which to look to proto-language variation.<sup>10</sup>

A problem here, though, one that holds as well for the other phonological cases discussed above is that if the changes in question are “natural” changes, then there is always the chance, as the OED says, of taking the changes to be independent in each language. Thus, what is really needed would be developments that have a highly particularized aspect to them — morphology is often good for that, but as seen above, there can be lots of ways in which morphology can be variable (especially along a lexically-governed dimension, as with Slavic 1SG **-m**).

Must phonology be ruled out altogether, just because of this potential problem with naturalness? The answer is no, and to see that, consider the following as a type of case where in principle one might be able pin things down (with additional work). Initial clusters of **h** with a sonorant **R** (i.e., **r**, **l**, or nasal) in earlier English gave later English **R-** (e.g. **hlúd** → **loud**, **hræfn** → **raven**, **hnitū** → **nit**, etc.), though there is still fluctuation with **hw-** / **w-** (as in **which**) and (on a more limited basis) **hy-** / **y-** (as in **human**). A parallel change occurred between early Old High German and later German, e.g. **hlút** → **laut** ‘loud’, **hréo** → **roh** ‘raw’, and in fact, began within the Old High German period, in the 9th century (Braune §153), so that **lút** occurs in later OHG). At first this seems just like the vowel and **sTr-** cases, but there is something additional that might be exactly the idiosyncratic clue needed to decide definitively if this is a case of proto-language variation (or not).

In particular, in both English and German, there is evidence of variation between #hR- and the innovative #R. In Old High German, as Braune (1967:





attention to. The level of (unnatural) particularity would presumably be great enough that a historical explanation for the parallel variability would need to be sought, and that explanation would be Proto-West-Germanic variation.

Statistics on Old High German still need to be done, and more on Old English would be a useful and important addition.<sup>11</sup> It must be admitted that the early returns, so to speak, are not promising. In particular, **hw** → **w** seems to have taken place quite early in Old High German but is still going on in present-day English. Nonetheless, since the degree of convergence on the details of the variability of **h**-loss in early stages of English and German is what will make or break the claim of proto-language variability underlying these processes, the more that is known about the variation in each language, the better the validity of the claim can be assessed.

Thus, this example, as part of a programmatic statement about the methodology of positing variability in a proto-language, provides a clear case of the type of investigation that could lead to promising results in determining proto-language variation; as with the **shall/sollen** case, all the ingredients are there, waiting only to be fully tested.

## **5. Some Concluding Remarks**

By way of conclusion, two points should be emphasized. First, it should not have escaped notice that while proto-language variation can be fatal to the goals of comparative reconstruction, the discussion here makes it clear that there are indeed conditions under which variability can happily feed into the Comparative Method. In a sense, in both the **shall/sollen** case and the **h**-loss case, highly particular

configurations of facts are present that can be compared across related languages. They thus present cases that are reminiscent of what Meillet (1967/1925]: 41) referred to as the “singular facts ... which are most suited to establish a “common language”, i.e. the sorts of facts that comparative reconstruction is best based on.<sup>12</sup> Meillet may not have had proto-language variability in mind with this statement but rather various grammatical features that are inherently arbitrary (“shared aberrancies”, such as singular/plural full-grade/zero-grade ablaut in the verb ‘to be’ in various Indo-European languages). However, his basic idea is applicable here, and thus can lead to more realistic reconstructions that are methodologically unassailable, unlike some of the possible solutions advanced but rejected in preceding sections.

Second, going along with the goal of realistic reconstructions is the goal of holding to realistic views of language change, and in this domain too, the discussion here can be illustrative. In particular, while the notion of “drift” in language change has often been given a somewhat “mystical” interpretation, as a sort of linguistic “invisible hand”, it can be given some substance through a recognition of proto-language variability and its aftermath. That is, if variation in a proto-language is inherited into individual languages as variability, lying dormant — or rather, “submerged” sociolinguistically — and waiting to bubble up to the surface under different social conditions, then it stands to reason that two related languages could show parallel developments that make an overt appearance late in their respective traditions. There would then be nothing mystical about drift; rather, it would be rationalized from a sociolinguistic standpoint.<sup>13</sup>

## REFERENCES

- Arvaniti, Amalia & Brian D. Joseph. 2000. "Variation in voiced stop prenasalization in Greek". *Glossologia, A Greek Journal for General and Historical Linguistics* 11-12.131-166 (Preliminary version in *Historical Linguistics: Ohio State University Working Papers in Linguistics* 52.203-233 (1999).)
- Beekes, Robert S. P. 1995. *Comparative Indo-European Linguistics. An Introduction*. Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Braune, Wilhelm. 1967. *Althochdeutsche Grammatik (fortgeführt von Karl Helm. 12. Auflage bearbeitet von Walther Mitzka)*. Tübingen: Max Niemeyer Verlag.
- Campbell, Lyle. 1997. *American Indian Languages. The Historical Linguistics of Native America. (Oxford Studies in Anthropological Linguistics, 4)*. New York/Oxford: Oxford University Press.
- Dyen, Isidore. 1969. "Reconstruction, the Comparative Method, and the Proto-Language Uniformity Assumption". *Language* 45.499-518.
- Fox, Anthony. 1995. *Linguistic Reconstruction. An Introduction to Theory and Method*. Oxford: Oxford University Press.
- Hock, Hans Henrich. 1991. *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter (2nd edition).
- Janda, Laura. 1996. *Back from the Brink*. Munich: Lincom Europa.
- Janda, Richard D. & Brian D. Joseph. 2001a. "Sound Change: Phonetics, phonology, sociology, or all of the above?". Paper presented at Annual Meeting of the Linguistic Society of America, Washington, DC, January 6, 2001.

To appear in *Variation and Reconstruction*, ed. by T. Cravens. Amsterdam: John Benjamins Publishing Co. (2005)

Janda, Richard D. & Brian D. Joseph. 2001b. "Reconsidering the Canons of Sound Change: Towards a Big Bang Theory". Paper presented at International Conference on Historical Linguistics, La Trobe University, Australia, August 15, 2001.

Janda, Richard D., Brian D. Joseph, & Neil G. Jacobs. 1994. "Systematic Hyperforeignisms as Maximally External Evidence for Linguistic Rules". In S. Lima, R. Corrigan, & G. Iverson, eds. *The Reality of Linguistic Rules*. Amsterdam/Philadelphia: John Benjamins Publishing Co., pp. 67-92.

Joseph, Brian D. 1992. "Diachronic Explanation: Putting Speakers Back into the Picture". In G. Davis & G. Iverson, eds., *Explanation in Historical Linguistics (Current Issues in Linguistic Theory, 84)*. Amsterdam: John Benjamins Publishing Co., pp. 123-144.

Joseph, Brian D. & Richard D. Janda. (Forthcoming). "Variability in Sound Change in English, German, and Proto-West-Germanic".

Labov, William. 1984. "Field Methods of the Project on Linguistic Change and Variation". In J. Baugh & J. Sherzer, eds., *Language in Use: Readings in Sociolinguistics*. Englewood Cliffs, NJ: Prentice-Hall, pp. 28-53.

Keiser, Steven Hartman. 2001a. "Demystifying drfit: Explaining linguistic change across speech islands". Paper presented at Annual Meeting of the Linguistic Society of America, Washington, DC, January 6, 2001.

Keiser, Steven Hartman. 2001b. *Language Contact and Dialect Contact across Speech Islands: The Emergence of a 'Midwestern' Variety of Pennsylvania German*. Ph.D. Dissertation, The Ohio State University.

To appear in *Variation and Reconstruction*, ed. by T. Cravens. Amsterdam: John Benjamins Publishing Co. (2005)

Lawrence, Wayne. 2000. “/str/ → /ftr/: Assimilation at a Distance?” *American Speech* 75.82-87.

Meillet, Antoine. 1967 [1925]. *The Comparative Method in Historical Linguistics* (English translation of *La méthode comparative en linguistique historique* (1925)). Paris: Champion.

Peters, Martin. 1980. *Untersuchungen zur Vertretung der indogremanischen Laryngale im Griechischen unter besonderer Berücksichtigung der Verbal morphologie*. University of Vienna Dissertation.

Shapiro, Michael. 1995. “A Case of Distant Assimilation: /str/ → /ftr/”. *American Speech* 70.1.101-107.

Toon, Thomas. 1976. “The Actuation and Implementation of an Old English Sound Change”. In *The Third LACUS Forum 1976*. Columbia, SC: Hornbeam Press, pp. 614-622.

Toon, Thomas. 1991. “The Sociopolitics of Literacy: New Methods in Old English Dialectology”. In P. Eckert, ed. *New Ways of Analyzing Sound Change (Quantitative Analyses of Linguistic Structure 5)*. San Diego: Academic Press.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis, Margaret Maclagan. 2000. “The Role of Drift in the Formation of Native-Speaker Southern Hemisphere Englishes: Some New Zealand Evidence”. *Diachronica* 17.1.111-138.

---

<sup>1</sup> This observation, by the way, removes from consideration one other possible account, namely one in which two types of \*p-like sounds are posited as distinctive phonological units (i.e. as separate proto-language phonemes).

<sup>2</sup> Although Hock's rural jargon account is reasonable, being based in large part on the nature of the lexical items in which the **zd** outcome is found (e.g. animal names, brewing terminology, etc.), there is an equally plausible solution that sees the origin of the different Greek outcomes in different Proto-Indo-European starting points. That is, Peters 1980 has argued quite compellingly that Greek **#h-** in these words results from an initial cluster of a laryngeal consonant with \*y (thus \*Hy-) whereas **#zd-** is the outcome of bare initial \*y. The significant point here is that if proto-language variation is involved at all, it is only linguistically determined variation, and if variation is involved at all, it is only in Greek, not in Proto-Indo-European.

<sup>3</sup> This development is reminiscent of the (South) German realization [spr-] (e.g. in **Sprache** 'language'), a fact that is discussed at greater length below in section 4.

<sup>4</sup> Given the match between the South German [spr-] (etc.) and the English variant, one might be inclined, if the only facts known to the comparativist were English spr- and spr- corresponding to German [spr-], to view the alveolar s- as the innovation and the retracted [s] pronunciation as the basis for comparison. Additional information (e.g. about the dialect distribution of the retracted pronunciations in each language) makes it clear that this is almost certainly the wrong choice, but see below for further discussion.

<sup>5</sup> It is necessary to focus on Classical Latin here, since Old Latin shows not **-t** as the 3SG past ending but rather **-d**, as in **deded** 'he gave', discussed somewhat below.

<sup>6</sup> Of course, artificially restricting the data is exactly what is done in carrying out Internal Reconstruction, where the only data allowed comes from a single synchronic stage of a language.

<sup>7</sup> This is of course the more or less standard view of the relationship between the Proto-Indo-European ending \*-t and the ending \*-ti, as far as the emergence of \*-ti within the proto-language is concerned, and thus reflects a hypothesis about pre-Proto-Indo-European (arrived at essentially by Internal Reconstruction on the reconstructed proto-language).

<sup>8</sup> The classical problem of Proto-Indo-European “s-mobile” is a similar case in point, where Indo-Europeanists routinely deal with “formulas” such as \*(s)meuk- ‘wipe, release’ (Latin **mucus**, Greek **smussetai** ‘wipes the nose’, etc.), with an optional and unexplained presence or absence of an initial \*s.

<sup>9</sup> Janda & Joseph 2001a, 2001b also note instances where the -r- precedes the st-cluster, as in **thunderstorm** (pronounced by some as [...str...]) as well as other contexts, and suggest that the change is spreading, in part through generalization of the linguistic contexts involved.

<sup>10</sup> The development of \*sk to s might even tie in with the retraction of s in clusters previously discussed, since a likely first step of \*sk on the way to s would have been a retraction of the sibilant, to \*sk̥. Thus this might constitute evidence for a very early part of the s-retraction process and might help to place it temporally.

<sup>11</sup> Both of these are planned as part of Joseph & Janda (Forthcoming).

<sup>12</sup> See the illuminating discussion in Campbell (1997: 212ff.) on Meillet’s views.

<sup>13</sup> In Joseph 1992, I suggest that another way of making sense of drift is to see it as the result of substantive universals of natural language. This view does not conflict with the sociolinguistically based interpretation offered here; rather each represents a different way in which related languages can come to exhibit certain similar

aspects of development. For an extended discussion of a few case-studies of such a sociolinguistically based notion of drift, see Trudgill et al. 2000, already cited above), and, for general discussion of the ways in which “drift” has been defined and characterized by various historical linguists, see Keiser 2001ab and references there.