

“Internal Reconstruction”

by

Brian D. Joseph, The Ohio State University

To appear in **A Companion to Historical Linguistics**
(Continuum Publishers)

edited by Silvia Luraghi and Vit Bubenik

[See chapters on **comparative method, morphological change,**
typology]

1. Introduction to the Method

Imagine that you enter a classroom and see that the desks and chairs are all in a different place from when you last saw them -- what sort of surmises might you reasonably make as to the causes of the disarray, and more particularly, what would guide you in those surmises? Many possibilities are imaginable: the movement of the furniture could be the work of aliens; it could be the result of a windstorm; it could be that the chairs staged a rebellion against the desks that had been oppressing them; or, another teacher may have rearranged the furniture in the classroom in order to offer a setting for a movie or to stage a play or simply to promote discussion in his/her class.

All of these are possible scenarios that allow for an explanation of the history behind the particular observed synchronic state of affairs encountered in the classroom. However, not all of them are equally plausible, and in fact, some of these can be ruled out rather easily. We know on independent grounds that chairs are simply not capable of holding the propositional attitudes or carrying out the actions necessary for staging a rebellion, and that visits by extra-terrestrials are highly unlikely (and if such creatures did visit, why would they pull a prank like changing around the furniture?). And, while a windstorm could wreak havoc in a room if windows were left open or were blown out, that is not a likely event, and in any case, an absence of broken glass would allow one to eliminate that possibility. This means that the best hypothesis remaining is the one that explains the alterations as the result of human actions sometime before your entry into the classroom.

This exercise is a matter of trying to deduce the historical events that led to a synchronic state, and the reasoning was guided by our sense of what events are likely and unlikely to have created the observed synchronic state of affairs.

This type of reasoning is found in all walks of life. We see a puddle or wet pavement in the morning and can hypothesize that it rained overnight even without directly experiencing the rain. We see a friend's hair in disarray and guess that he had lost his comb. And so on. In each case we are attempting to reconstruct some aspect of the past that is not directly observable but which is inferable from the outcome and what we know about how such outcomes generally arise.

This same reconstructive method can be applied to language, so that the causal

historical underpinnings to a particular configuration of facts in a language can be guessed at, or to use more scientific-sounding terminology, hypothesized about, with the most reasonable hypotheses being those that are supported by what is known about language and about language history in general (just as the most reasonable hypothesis in the classroom example did not involve aliens or animate furniture). Thus **language typology** (see Hock, this volume) informs this method, by giving a sense of what can be expected for a given language state. In historical linguistics, this method has a special name: **internal reconstruction**, so-called because it is a reconstructive technique that relies entirely on observed evidence from a single stage of a language, and thus is “internal” in that there is no “external” comparison to related languages (as there is in the **comparative method** (see Hewson, this volume)). In a sense, the designation “internal” is not completely justified, since by drawing on known properties of language and language change, considerations external to the language stage in question are brought into play; nonetheless, the method is “internal” as far as the source of the data one works with is concerned (again, unlike comparative reconstruction).

2. The Method of Internal Reconstruction Exemplified

The classic application of this method involves drawing inferences about the historical sources of morphophonemic alternations (that is, alternations in the phonemic shape of morphemes).

For example, the nominative singular of the Ancient Greek word for ‘honey’ is *méli*, and the genitive singular is *mélitos*. Other facts about Ancient Greek noun inflection, e.g. a comparison of nominative *poimén* ‘shepherd’ with genitive *poiménos* ‘of a shepherd’ or of nominative *óar* ‘wife’ with genitive *óaros* ‘of a wife’, demonstrate that the genitive singular ending is clearly *-os* with other nouns, and in particular consonant-stem nouns. The best synchronic analysis of ‘honey’, therefore, segments it as *méli-Ø* / *mélit-os*, so that there is allomorphy in the stem between *méli-* and *mélit-*. How did such allomorphy arise? Knowing that paradigms generally start out as perfectly regular, with no allomorphy at the outset, and that languages often lose consonants at the ends of words, it is reasonable to suppose that prior to the attested Ancient Greek stage with the nominative *méli*, there was a stage in which the nominative was **melit*. The asterisk, as

with comparative reconstruction, indicates that this form is hypothetical, not directly attested but inferable. Moreover, to get from this posited **melit* to the attested *méli*, a sound change of final *t*-deletion (perhaps to be viewed as a more general final stop deletion) must be posited as well. Internal reconstruction in this case thus resolves the synchronic *méli/mélit*- allomorphy into an earlier unity, with a single stem form **melit*-, and recognizes as well a sound change that gave rise to the later allomorphy.

Moreover, this account generalizes to other similar alternations in Greek, e.g. neuter present participle nominative singular *lúon* ‘loosening’ / genitive singular *lúont-os*, for which an earlier nominative form **luont* can be reconstructed, guided by a recognition that sound changes, such as the final stop deletion posited for *méli*, typically affect a wide range of forms. Knowledge of what can happen to sounds is thus brought to bear here on the analysis of *méli/mélitos*, just as knowledge of likely forces moving furniture around was brought to bear on the reasoning in the classroom example.

As another example, consider the two words for ‘sleep’ in Latin: *somnus* and *sōpor* (differentiated as ‘sleep’ versus ‘deep sleep’, respectively). Given other nouns in *-nu-* and *-or-* in Latin (e.g. *signum* ‘sign’, *lignum* ‘wood’, *calor* ‘heat’, *tumor* ‘swelling’), a reasonable synchronic analysis would segment these nouns as *som-nu-* and *sōp-or-*, respectively, thus yielding root allomorphy in consonantism of these derivationally related forms, *som-* versus *sōp-*. The difference in the final consonants in these forms can be resolved by noting that *som-* occurs before a nasal, and that regressive assimilation of a stop to a following nasal is common cross-linguistically. Thus, *somnus* can be internally reconstructed as **sop-nu-*, and a sound change of *p > m / __n* can be posited. As in the Greek case, this account generalizes to other alternations of a labial stop and a nasal, as with *dap-* in *dap-s* ‘sacrificial meal’ and its root cognate *dam-* in *dam-num* ‘loss’.

This Latin case allows for a generalization in a somewhat different direction that the Greek did not. That is, there are isolated forms in Latin, words without any apparent relatives within Latin itself, that have the same *-mn-* sequence as in *somnus*. The generality implicit in the positing of a sound change turning a labial stop into *m* before an *n* means that even for a word like *amnis* ‘river’, with no related forms sharing its root element and thus nothing that can give a clue that *am-* had ever been anything other than *am-*, one nonetheless can speculate that in a prior but unattested stage of Latin, this word

may have been **ap-nis*. The significance of this hypothesis becomes clear in the next section.

Morphophonemic alternations offer a direct basis for the historical inferences that we call internal reconstruction, but as the case with *amnis* shows, certain configurations of facts allow for internal reconstruction even when there are no overt clues in the form of alternations. Sometimes, gaps in patterns are enough to allow for internally derived historical hypotheses. For instance, Old Irish has a stop system with voiced stops *b d g* and voiceless stops *t k*, thus with a gap at the labial point for the voiceless set as compared to the voiced set. It is reasonable to infer from that distributional fact that there may once have been a *p* in pre-Irish and that a sound change eliminating *p* from the phonemic inventory of the language may have occurred.

Similarly, other sorts of synchronic irregularities -- thinking of morphophonemic alternations and gaps in patterns as a type of “irregularity” in that they constitute nonuniformity in the system where uniformity might otherwise be expected -- provide a basis for the formation of historical hypotheses. For instance, within Latin, one irregularity about combinations of prepositions and the objects they govern is that whereas the order of elements is usually Preposition + Object, both with noun objects and pronoun objects, as in *cum Marcō* ‘with Marcus’ or *ad eōs* ‘to them’, when the preposition is *cum* ‘with’ and the object is a pronoun, the pronoun goes first and the preposition is enclitic to it, e.g. *mēcum* ‘me-with’ (i.e. ‘with me’). This invites the inference that at an early stage of Latin, prepositions more generally were enclitic and thus that *mecum* reflects an archaic usage that, for whatever reason, had not fallen in line with the regularizing that other preposition-plus-object combinations underwent.

3. Confirming the Results of Internal Reconstruction

The method of internal reconstruction thus allows for the generation of hypotheses, of greater or lesser plausibility, about an earlier linguistic state of affairs. Some of these hypotheses can be readily ruled out, but once that is done, how might one determine if the best remaining hypothesis is accurate? The answer lies in the other historical linguistic reconstructive method, the comparative method, and thus in bringing external evidence from other languages to bear on the internally arrived at hypotheses.

That is, in the case of the Greek word for ‘honey’, the evidence of Hittite *milit-* ‘honey’ and Gothic *miliþ* ‘honey’ shows that a reconstruction of the oldest form of this word in the Greek branch with a final *-t-* in the stem is well warranted. Similarly, cognates to the Latin forms for ‘sleep’, such as Greek *húpnos* ‘sleep’ and Sanskrit *svapna-* ‘sleep’, point to the validity of reconstructing the pre-Latin form of ‘sleep’ as **sop-no*, and cognates with *p* in various related languages but *Ø* in Irish, such as Latin *pater* ‘father’, compared with Old Irish *athir*, show that the positing of a prehistorical **p* in early stages of the Celtic branch of Indo-European is a reasonable step to take. And, in the case of the Latin *amnis*, external comparisons such as Hittite *hap-* ‘river’, confirm the speculative hypothesis of a pre-form **ap-nis* that was arrived at by extending the internal reconstruction of *somnus* to an isolated form.

The most famous example of the confirmation of internal reconstruction via comparative evidence is the case of Ferdinand de Saussure’s “coefficients sonantiques” (later somewhat erroneously referred to by others as “laryngeals”). These are a class of consonants that Saussure 1879 posited for a stage prior to Proto-Indo-European, working just from the evidence of reconstructed patterns for vowel alternations for the proto-language. For instance, he assimilated the pattern of *ē* alternating with *a* to that of *er* alternating with syllabic *r*, by hypothesizing that there was a consonant that had the property of lengthening a preceding vowel and surfacing as a syllabic element when the vowel was eliminated (for morphological reasons, an “ablaut” state of the root referred to in Indo-European linguistics as the “zero-grade”). This was a purely internally arrived at reconstruction but it received support over 40 years later when Jerzy Kuryłowicz (see Kuryłowicz 1927) demonstrated that certain consonants in Hittite, usually transcribed as *h*, appeared in exactly the positions that de Saussure predicted for his “coefficients sonantiques”. This discovery not only confirmed de Saussure’s hypothesis, paving the way for the development of what is now called “laryngeal theory” for the Indo-European phonological system, but also validated the methodology of internal reconstruction.

4. Limitations of Internal Reconstruction

For all the fact that internal reconstruction has been shown to be a powerful means of shedding light on the prehistory of linguistic states that might otherwise not be amenable

to any further historical speculation, it has its limitations as a method.

For one thing, not all synchronic alternations have arisen by the relatively “clean” path that forms like Greek *méli* show. For instance, the alternation seen in the Greek noun for ‘name’, with a nominative *ónoma* and a genitive *onómatos*, lends itself to the same sort of analysis as that given for *méli*, so that one might reconstruct the nominative as **onomat* and segment the genitive as *onómat-os*. That is a perfectly reasonable internal reconstruction, but the comparative evidence in this case is disconfirmatory: cognate forms in other languages show no sign of a *-t-* in this stem at all, neither in the nominative (cf. Sanskrit *nāma*, Latin *nōmen*, Hittite *laman*) nor in the genitive (cf. Sanskrit *nāmn-as*, Latin *nōmin-is*, Hittite *lamn-as*). The *-t-* presumably entered the Greek paradigm in some way other than being an inherited part of the stem, quite possibly being added to the genitive due to influence from adverbial forms in *-tos* (e.g. *ektós* ‘except’) or else analogically based on genitives of *-t-*stems; that is, there is no evidence for a prehistoric stage of Greek with a nominative **onomat*, even though that is the form that internal reconstruction would lead one to.

And, in the case of reconstructing a **p* for some pre-stage of Irish, the comparative evidence shows that a Proto-Indo-European **p* was lost on the way to Irish, and more accurately on the way to Celtic. Yet, the absence of a consonantal reflex of **p* in all of the Celtic languages points to the conclusion that the loss of **p* was a very early step in the development of the entire Celtic branch of Indo-European, and thus not as recent a phenomenon as a hypothesis based just on Irish evidence alone would suggest.

More generally, hypothesizing loss based on absence is a risky proposition; the fact that English lacks uvular consonants, for instance, does not mean the language once had them and lost them; it might well be the case that it simply has never had them.

Despite such limitations, internal reconstruction is useful in historical investigation, and, indeed, is widely considered to be among the standard methods used in historical linguistics; not surprisingly, therefore, it is included in handbook-style surveys of the field (see, e.g., Kuryłowicz 1973, or Ringe 2003) and in nearly all textbooks (Anttila 1972/1989 being a notable example where the method is given particular prominence) and specialized treatments of reconstruction methodology (e.g. Fox 1995: Ch. 7).

5. Concluding Remarks

In a very real sense, internal reconstruction can be thought of as a hypothesis-generating methodology, and to the extent that there are no constraints on hypotheses that may be entertained, engaging in internal reconstruction is a license to be creative and propose possible scenarios, i.e. historical hypotheses, that are constrained only by the plausibility offered by what is known about language and language change in general.

Still, especially given the interplay noted in section 4 between internal reconstruction and comparative data, one can wonder if internal reconstruction is really needed. That is, if one worked solely with the comparative method and compared Latin *somnum* with Greek *hupnos* or Sanskrit *svapna-*, it would be trivial to explain the $m : p : p$ correspondence set by reference to its occurring in the context of a following nasal. Similarly, a direct comparison of Greek *méli* with Hittite *milit* would readily lead one to suppose that the Greek nominative had once been **melit* and that a final-stop-deletion sound change had occurred. Thus it may well be that internal reconstruction rarely tells us something we could not know by other means.

There are, however, some circumstances in which internal reconstruction offers the only avenue for historical inferences. In particular, in cases in which there is no other potentially corroborating data available, internal reconstruction is the only method available. Such a case arises when one is dealing with a proto-language, reconstructed by comparative evidence; in order to push the temporal frame for the proto-language back even further than the comparative method allows for, applying the reasoning of internal reconstruction can offer some basis for surmises about the prehistory of the proto-language. The hypothesis of the 19th century Indo-Europeanist August Schleicher (see Schleicher 1871: 13, discussed also in Pedersen 1959: 270) that the nominative singular of the word for ‘mother’ (Latin *mater*, Greek *mātér*, Sanskrit *mātā*, etc.) in Proto-Indo-European was not **mātē(r)* with a lengthened ablaut grade in the final syllable but rather **mātar-s*, a view reached independently but cast in a more modern form by Oswald Szemerényi (see Szemerényi 1956; 1996: 116), is essentially a form of internal reconstruction on the proto-language, deriving the final **-ēr* from an earlier, “pre-Proto-Indo-European”, stage involving the more widely distributed nominative ending **-s* and the long vowel from a compensatory lengthening with the loss of that **-s*. Moreover,

there are language families for which comparative data from a range of languages is not easy to come by where internal reconstruction can help to get one started (see Campbell & Grondona 2007, for example) as well as instances where such data is lacking altogether, as in case of the language isolates (e.g. the American Indian language Zuni); in the latter situation, all reconstruction can only be internal, drawing just on data from that one language. In such a case, dialect variation could in principle offer some comparative basis for reconstruction, but in a technical sense, all the data would be coming from a single language, and thus would be “internal”.

Thus, even if not always providing novel hypotheses or reconstructions that would not be possible otherwise, internal reconstruction does have something to offer the historical linguist and is an important and valuable part of the historical linguist’s toolkit.

References

- Anttila, Raimo (1972/1989), An introduction to historical and comparative linguistics. New York: Macmillan (2nd edn., Introduction to historical and comparative linguistics. Amsterdam: John Benjamins, 1989).
- Campbell, Lyle and Verónica Grondona (2007), ‘Internal reconstruction in Chulupí (Nivaculé)’. Diachronica, 24, 1-29.
- Fox, Anthony (1995), Linguistic reconstruction. Oxford: Oxford University Press.
- Kuryłowicz, Jerzy (1927), ‘*a* indoeuropéen et *h* Hittite’, in A. Gawronski and T. Lehr Splawinski (eds), Symbolae grammaticae in honorem Ioannis Rozwadowski. Krakow: Gebethner and Wolff, pp. i.95-104.
- Kuryłowicz, Jerzy (1973), ‘Internal reconstruction’, in T. Sebeok (ed), Current trends in linguistics 11. The Hague: Mouton, pp. 63-92.
- Pedersen, Holger (1959), The discovery of language. Linguistic science in the nineteenth century (translated by John W. Spargo). Bloomington: Indiana University Press.
- Ringe, Donald (2003), ‘Internal reconstruction’, in B. D. Joseph and R. D. Janda (eds),

Handbook of historical linguistics. Oxford: Blackwell Publishing, pp. 244-61.

Saussure, Ferdinand de (1879), Mémoire sur le système primitif des voyelles dans les langues indo-européennes. Leipzig: Teubner.

Schleicher, August (1871), Compendium der vergleichenden Grammatik der indogermanischen sprachen. Kurzer Abriss einer Laut- und Formenlere der indogermanischen Ursprache, des Altindischen, Alteranischen, Altgriechischen, Altitalischen, Altkeltischen, Altslawischen, Litauischen und Altdeutschen. Weimar: Hermann Böhlau.

Szemerényi, Oswald (1956), 'XXXX'. Zeitschrift für vergleichende Sprachwissenschaft, 73, 190-96.

Szemerényi, Oswald (1996), Introduction to Indo-European Linguistics (4th edn.). Oxford: Oxford University Press.