

Unnecessary asterisks and realism in reconstruction: underspecified is still real

Patrick Honeybone

1 Introduction

One of the classic questions in historical linguistics is: how real are reconstructions?¹ In this article, I consider this question as a historical phonologist – as someone who is equally interested in theoretical phonology and in phonological change. It seems surprisingly rare to consider the question from this perspective. This is a shame because novel avenues open up for understanding and answering it if we do – indeed, I think that we cannot hope to properly understand the question, or come to an answer to it, unless we consider the phonological entities that are used in historical phonology in the light of the conceptual frameworks that have been developed to analyse and understand phonological objects in the heat of phonological theory. This is really just an application of uniformitarianism (broadly definable as: what we find in the past should be fundamentally the same as what we find in the present). It will lead us, however, to what may be a surprising conclusion: that the entities that are reconstructed in historical linguistics (which are conventionally marked with an asterisk) have fundamentally *exactly* the same ontological status as the entities that are considered in theoretical phonology. One corollary of this will be that the asterisk is unnecessary (and, even, a tautology). This is *not*, however, intended to imply that historical linguistics is *wrong* and theoretical phonology is *right*. The analytical methods and concepts of both are compelling and crucial. Nor do I mean – just because asterisks are unnecessary – that historical linguists should necessarily stop using them (it doesn't *hurt*, after all), but I do think that historical linguists should take seriously the issues that I discuss. The paper is intended as an essay into the philosophy of historical phonology, which I think, on occasion, is worth considering.

In what follows, I first consider some basics of linguistic reconstruction (in section 2), focusing on the reconstruction of phonology and the use of the 'reconstructed' asterisk. I then address, in section 3, the issues that are classically considered concerning the ontology of reconstructed objects. In section 4, I discuss what happens if we think about this through the lens of phonological theory. The ideas expressed there are applied, in section 5, to two case studies – one is from Old English, the other is the renowned case of Proto-Indo-European stops. Section 6 concludes. My key arguments are made in section 4. They are: (i) that historical phonological reconstruction is *cognitive* reconstruction; (ii) that in fact *all* analysis of phonological states from the past is cognitive reconstruction; and (iii) that there is no *qualitative* difference between the objects that are assumed in cognitive reconstruction and the objects that are assumed in the phonological analysis of contemporary languages.

It is worth being explicit that I focus here on the reconstruction of *phonology*, not vocabulary (which is what 'reconstruct' often refers to). It is equally important for historical linguists to reconstruct both words and linguistic systems, but I am not interested here in which segments occurred in any particular word of any previous state of any language. It is not a phonological fact that the PIE word for 'nose' contains an /n/ (at its start), but it *is* that PIE had the segment /n/ in its phonological system, and that it could occur in word-initial position.

¹ I thank the editors of this volume and Pavel Iosad for comments of a draft of this article, which improved it and helped me clarify my thinking. They should not, however, necessarily be thought to agree with it, of course.

I am interested in the systems of segments in past languages states: which features were they made up of? how could they combine? These are phonological questions, and if reconstructed languages are real languages (see section 3), we should be able to ask them.

2 Phonological reconstruction

It is not unusual to come across claims like this: “[o]ne of the greatest achievements in the history of linguistics was the establishment, beginning in the eighteenth century, of the comparative method, which made possible the reconstruction of prehistoric languages” (Beadle 1989:173), or even claims that the comparative method is “one of the greatest achievements of the human sciences” (Longobardi & Guardiano 2009:1681). These claims are not wrong: the comparative method is a remarkable tool, and the reconstruction of past linguistic states for which we have no direct records that it has enabled (along with other evidence) is extraordinary. Hale (2007:225) is both representative and right to say that the comparative method “has established its usefulness in empirical research with great regularity” – it works again and again to produce compelling results. What is more, it makes predictions which have been shown to be met,² such as Bloomfield’s (1925) prediction of a distinct segment cluster in the Proto-Central Algonquian word for *red* purely from his insistence on exceptionless correspondences, which was shown to be correct (in Bloomfield 1928) through the discovery that a distinct cluster of this type was preserved in Swampy Cree (see Campbell 1996).

There is no space here to set out the comparative method’s principles of correspondence-selection and of weighing-up-majority-forms-and-phonological-typology/directionality. They are often discussed in textbooks and handbooks on historical linguistics (e.g., Fox 2015, Weiss 2015), and it is right that they are defended, along with comparative reconstruction as a whole, when they are not properly applied (see Fellner & Hill 2019 for a recent defence). It is important to recognise, however, that the comparative method is not the *only* kind of evidence that we have for the phonology of unwritten stages of languages: internal reconstruction can give crucial insights into underlying forms, and has long been used hand-in-hand with the comparative method when reconstructing past phonological states (Fox 2015 cites Brugmann 1876 and Saussure 1879 as early examples); we can consider the behaviour of loanwords into and out of the language that is being reconstructed; we can cautiously consider how results from typology and theory can guide a reconstruction; we might even be able to use indirect evidence for a language (from comments describing it if they happen to exist). It is expected that we use all available types of evidence when we engage in phonological reconstruction. Janda & Joseph (2003:37) rightly recommend “informational maximalism” in historical linguistics – that is, “the utilization of all reasonable means to extend our knowledge of what might have been going on in the past”. The comparative method is thus not the only source of evidence that we have for phonological reconstruction, but, nonetheless: it surely *is* an extraordinary achievement that we are able to reconstruct the phonology of past states of languages for which we have no written evidence.

Or is it? Are we *really* reconstructing past phonological states? I address this in section 3. Before that, I consider the use of the asterisk, given that I am going to take aim at it later.

² I assume, along with, for example, Hale (2007, 2012, 2015), Weiss (2015), and Fellner & Hill (2019), that standard phonological changes (of the type that Honeybone 2016 calls ‘N-changes’) are exceptionlessness (for good phonological reasons), and that this is an inherent aspect of the comparative method.

2.1 Asterisks in phonological reconstruction

Reconstructed objects are standardly indicated with an asterisk in historical linguistics, as in (1), taken from the discussion of how “PGmc word-final *-z has been lost throughout WGmc when the preceding syllable nucleus was unstressed” in Ringe & Taylor (2014:43, omitting some evidence for the Proto-Germanic form). The initial syllable was stressed in Proto-West Germanic, so the PGmc final -z is in the environment where it was lost.³

(1) PGmc *gastiz ‘guest’ nom. sg. > PWGmc *gasti > OE *giest*, OHG *gast*

There are two types of forms given in (1). The first, (*gastiz and *gasti) from Proto-Germanic and Proto-West Germanic, are marked with an asterisk, and the second (*giest* and *gast*) from Old English and Old High German are not. These forms are in a section on ‘Proto-West Germanic sound changes’ and are all intended to give information about the (pre- and post-change) phonological states involved.

There is not much in-depth discussion of this use of the asterisk in historical phonology because the general impression is that it plays an obviously necessary role – to differentiate between the discussion of forms which have one status and forms which have another.⁴ Trask’s (1996:204) detailed textbook, for example, writes that an “asterisk marks a sound that is not directly recorded, but which linguists have concluded was probably the original sound” and does not say any more about it. Bynon’s (1977:47) classic textbook simply says that asterisks are used with a form “in order to indicate that it is not actually attested but merely postulated”, and McMahon (1994:6) writes that a proto-form appears “with a preceding asterisk, to indicate that there is no direct evidence for it; we have no texts and no speakers”. There is no equivalent description of forms *without* an asterisk. By implication, though, we can assume that they are phonological forms which *have* been directly recorded, which *are* actually attested, and for which there *is* direct evidence.

It is true that forms like *giest* and *gast* in (1) are representations of written forms, but they are of interest to historical phonologists because they function as quasi-phonemic transcriptions. Once we know how to phonologically decode the OE and OHG orthographic forms in (1) (for example, that the dot above the <g> in *giest* transcribes [j], and that most other letters transcribe the phonological segments that they are conventionally used to represent – for example <s> = [s]⁵) then this is how they are interpreted. It is not (of course) the marks on parchment (or any other kind of written mark) that are the subject matter for historical phonologists, but the phonological forms that they represent. To take something like *gasti > *giest* literally (if the pre-change state is a reconstructed phonological object and the post-change state is a set of letters) would be ontologically incoherent: phonological objects do

³ I do not single this book out because I think it bad – quite the opposite. Like other work that I use for exemplification below (from Clackson, Minkova and Lass, for example), I refer to it simply to provide examples of conventions or terms that everyone uses, which I think we could reflect on.

⁴ The story of how linguists came to use the asterisk in this way is an interesting one, as is the way in which this morphed into the current alternative linguistic use of the asterisk to indicate ungrammatical forms. I lack the space to trace this here. Koerner (1976) shows the development of the former usage, from cases in the 1840s to indicate unattested but morphologically predictable forms, to the precise usage of interest here (marking reconstructed forms) starting in the 1850s, and made famous in Schleicher (1861). Graffi (2002) considers the development of latter usage (from the former), showing its early use in the 1920s and 1930s to indicate ‘non-existent’ or ‘impossible’ forms, and its establishment as the conventional way to indicate ungrammaticality in the 1950s and 1960s.

⁵ In line with, for example, Hale (2007), I see phonological objects in square brackets as the ‘end of phonology’, qualitatively the same as underlying phonological objects (occurring before transduction to a gestural score), although it seems likely that my key points here would hold in models with different assumptions about this, too.

not change into letters. This is not news to anyone, and practice of the sort in (1) makes perfect sense when we know how to interpret it: that both forms on either side of the ‘>’ are relevant because they both represent phonological forms.

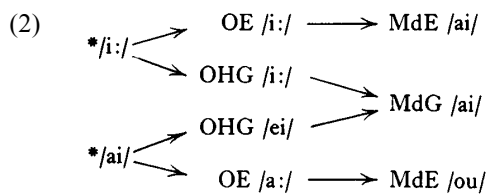
The point is even clearer in material like that in Table 1, from Clackson (2007:37). This gives reconstructed forms for stops at three of the places of articulation that Proto-Indo-European (PIE) had and comparative evidence for them from 11 languages (OCS = Old Church Slavonic, Lith. = Lithuanian, Arm. = Armenian, and Toch. = Tocharian). Clackson also gives forms for palatals and labio-velars, which are omitted here for simplicity.

PIE	Hittite	Sanskrit	Avestan	Greek	Latin	Gothic	OCS	Lith.	Old Irish	Arm.	Toch.
*p	p	p	pf	p	p	fb	p	p	∅	hw ∅	p
*b	bp	b	bβ	b	b	p	b	b	b	p	p
*b ^h	bp	bh	bβ	ph	fb	b	b	b	b	b	p
*t	t	t	tθ	t	t	þd	t	t	t	t’y	tc
*d	dt	d	dδ	d	d	t	d	d	d	t	ts ś
*d ^h	dt	dh h	dδ	th	fdb	d	d	d	d	d	tc
*k	k	kc	kc	k	c	hg	gžz	k	c	k’	k ś
*g	gk	gj	gj	g	g	k	gz	g	g	k	k ś
*g ^h	gk	gh h	gj	kh	hg	g	kč	g	g	g	k ś

Table 1: PIE segmental reconstruction with comparative evidence, from Clackson (2007)

Whereas the forms in (1) are words (which are comprised of phonological segments, of course), those in Table 1 are unambiguously simply segments. Clackson (2007) gives clear guidance about how to interpret the orthographic forms from attested languages phonologically (so, for example, Armenian <t’> = [t^h] and <y> = [j]), and the PIE forms can be recognised to comprise a segmental system with stops at all places of articulation showing a three-way laryngeal contrast of the type /*t, *d, *d^h/ (to exemplify using coronals). Relevant for my purposes is that, again, only certain forms in Table 1 have asterisks – those of PIE. The other forms are unasterisked, because, the reasoning goes, they have been recorded.

The material in (2), taken from Bynon (1977:48), shows things even clearer still. Bynon is summarising her discussion of the development of certain vowels from Proto-Germanic (the forms given in the first column) to Modern German and English.



The forms */i:/ and */ai/ are Proto-Germanic, ‘OHG’ and ‘MdG’ indicate two stages in the history of German (Old High German and Modern German), and ‘OE’ (Old English) and ‘MdE’ (Modern English) indicate two stages in the history of English. Here things are made very clear through the abandonment of orthographic forms for attested languages and the direct use of overtly phonological symbols (shown using slanted brackets). Proto-Germanic

forms are marked with an asterisk, we can assume, because Proto-Germanic is a reconstructed language, and the OE, Mde, OHG and MdG forms are not asterisked because they are from languages that are attested through written documents (and/or living speakers). The unasterisked forms for OE and OHG in (2) are more ‘obviously’ phonological than those in (1) and Table 1, but the unasterisked forms in all three have fundamentally the same status: they represent ‘attested’ phonological entities from the past.

All of this implies that there are two fundamentally different kinds of thing in historical phonology: ‘reconstructed forms’ and ‘attested forms’. That is one of the points that this paper engages with. As will become clear, I think it is conceptually problematic.

3 Realism in reconstruction

Overviews of reconstruction discuss both ‘realist’ and ‘formulist’ (also spelt ‘formalist’) interpretations of reconstructed objects. A formalist (or ‘instrumentalist’) approach considers reconstructed segments to be quasi-fictional cover-labels – place-holder summaries of correspondence sets. Thus, the PIE stop system reconstructed in the previous section (for coronals) as /*t, *d, *d^h/ could just as well be represented as /1, 2, 3/ or /☪, ☩, ☪/, because /*t/, for example, is just a way of referring to the collection of attested segments in the fourth row of Table 1.

This formalist philosophy of phonological reconstruction was widespread in the early and mid 20th century – Meillet (1903-1937; cited by many, including Lass 1997), was clear on this, arguing that reconstruction methods *cannot* produce “a reconstruction of Indo-European as it was spoken: it is a defined system of correspondences between historically attested languages” (1937:47, my translation). This position leads Pulgram (1959:423) to talk about a “kind of fabricated reconstruction which we call Proto-Indo-European”.

Formulism was not incontestable during this period, however, and has since been largely replaced by a realist approach which assumes that the segments of reconstructed languages are *real* phonological objects, on the assumption that “protolanguages ... are simply languages like any others” (Lass 1997:274), just as cognitive realism has replaced instrumentalism in theoretical phonology. Hale (2007:246) illustrates the commonness of this position in contemporary work, writing (quoting a passage from Lichtenberk 1994) in favour of “a ‘realist’ conception of protolanguages – they are, indeed, ‘as real as the language around us.’”

Following this contemporary consensus, I too assume that phonological reconstructions need to be interpreted along realist lines – I see this is simply as an application of uniformitarianism.⁶ What is more, it gives us the potential reality check of typological control on reconstruction: we shouldn’t reconstruct phonologies that are impossible because we expect phonological reconstruction to produce systems which are phonologically real. As Hale (2007) continues, however (in discussion following the just-cited passage), if we place the status of phonological reconstructions under the microscope, assuming them to be ‘as real as the languages around us’, we need to be explicit about what we mean by ‘language’. For our purposes, this is the phonological aspect of language: what do we mean by ‘phonology’? Hale’s answer is the widely-held position that, in terms of their phonology, “‘the languages around us’ are grammars, i.e. computational devices in the mind/brain of individuals” (2007:246). This strikes me as exactly right. And what is more, if we combine it with a consistently realist approach to phonological reconstruction, interesting ideas emerge. I consider

⁶ Uniformitarianism is more complex than I am treating it here (Walkden 2019). I am assuming standard uniformity of law and, especially in section 4, substantive uniformitarianism of state concerning phonological structure.

what it means to take the point seriously in the following section, and then draw some analytical conclusions from it in section 5.

4 The nature of phonological objects: phonological reconstruction is cognitive reconstruction

The realist position on reconstruction assumes that the forms in (1), (2) and Table 1 are intended to represent phonological objects from the past. What *are* phonological objects and where do they exist? Hale's (2007:246) answer above (they exist in *grammars*, which exist in the *mind/brain* of speakers) is the mainstream position in theoretical phonology. As van Oostendorp (2013:274) puts it, "phonology is about (a speaker's) knowledge of sound systems". This is the basic mentalist position, which is standard in grammar-focused linguistics: that phonology (like syntax, and other structural aspects of language) is "a branch of the study of mental representation, the psychology of mind" (Hale & Reiss 2000:161). Hale & Reiss (2008) describe this as an 'I-phonology' approach, adopting Chomsky's (1986) convention of distinguishing between I-language (which is *internal* to the mind and is the subject of grammatical investigation) and E-language which is what humans produce when they use their language to communicate (and is thus *external* to the mind). This position is both contemporary and classic. Reiss (2018:439) cites Hammarberg (1976:354) saying that, "it should be perfectly obvious by now that segments do not exist outside the human mind". We can interpret this to mean that phonological objects (such as segments and syllables) only exist in the mind – they are cognitive objects.

In Honeybone (2012), I described this as leading to the conclusion that phonological reconstruction is *cognitive reconstruction*. This is no great realisation: if we adopt realism, then we need to assume that the products of reconstruction are the same kind of thing as the objects that are assumed to exist in the analysis of contemporary languages. And if phonological objects (as assumed in the analysis of contemporary languages) are *cognitive* objects then the aim of historical phonology is to work out what were the objects that existed in speakers' minds in the past: to reconstruct past cognitive states. These reconstructed objects must be subject to the same principles that constrain the objects recognised in the analysis of segmental and prosodic systems and of rule or constraint components of the phonological grammar. I do not think this should be contentious – as Lass (1997:27) puts it: "The primary constraint on a historical subject is its non-historical metasubject. Historical biology is part of biology, and hence constrained by biological knowledge and theory; historical linguistics is a branch of linguistics, constrained by non-historical linguistic knowledge and theory."

I therefore assume that the objects recognised in the phonological analysis of contemporary languages ('contemporary phonological objects', CPOs) and those assumed in the analysis of past stages of languages ('reconstructed phonological objects', RPOs) have the same ontological status. I further argue that our *degree of certainty* about CPOs and RPOs is fundamentally the same – not necessarily *exactly* the same, but qualitatively commensurable.

Some historical linguists may be cautious about this because the certainty they seek is for the reconstruction of whole words. Thus Janda & Joseph (2003:93) consider assigning a "percentually expressed level of confidence in a particular reconstruction" (rather than using an asterisk to mark RPOs), so, for example, they propose that they would mark the reconstruction of the PIE word for 'master'/'powerful one' as 'potis' as 90% certain. My reply is two-fold: firstly, that we can be quite uncontroversially certain about significant aspects of RPOs (and, in fact, these 'significant aspects' are what phonologists should be reconstructing in the

first place),⁷ and secondly, that the certainty that historical phonologists have about RPOs is (as argued above) not fundamentally different to the certainty that theoretical phonologists have about CPOs. I discuss the first point in section 4.1 and the second in 4.3. Between this, section 4.2 addresses another point which will be important to my argument.

4.1 Basic phonological objects: what is reconstructed in phonological reconstruction?

Up till now, I have assumed the phonological segment to be the relevant CPO to which RPOs should be compared. This is flawed. In the segmental phonological realm, the basic unit that phonological theory assumes to exist is the phonological *feature*.⁸ Segments exist as phonological objects, but they are units which organise features together. Phonological features are the ‘significant aspects’ (mentioned in the previous paragraph) that phonologists need to focus on in phonological reconstruction.

I need a set of (consonantal) features to apply in section 5. I shall use features from the set in Hayes (2009), which is quite traditional in that all features are fundamentally binary, but which adopts some innovations from post-*SPE* work. I use the features in (3), which I set out here (and in section 5) to make explicit which account for Place, Manner and Laryngeal specifications, and readers will be able to translate them into the feature-system that they prefer. These are not *all* the features that Hayes uses, but they are enough to characterise the segments that I discuss.⁹

(3) P:	[labial]	[coronal]	[dorsal]
M:	[continuant]	[sonorant]	[nasal]
L:	[voice]	[spread glottis]	[constricted glottis]

Assuming (3), consonants consist in the phonology of specifications for all these features unless a segment is underspecified for a particular feature (as shown by its phonological behaviour). Lexically underspecified values can be filled in during phonological computation, or, as Keating (1988) and Hale & Kisser (2007) point out, underspecification can persevere right to the end of phonology. If we think of RPOs in terms of features, we can recognise that the first segment in the PIE word for ‘nose’ (reconstructed by Pokorny 1959 and many others as ‘nas-’) was, among other things: [–labial], [+coronal], [–dorsal]. All segmental symbols (such as ‘p, m, ŋ, ʃ, x, h, ʔ’) stand for sets of feature-specifications, and, while most symbol-to-feature-specification-equivalences are conventionally agreed, not all are (for example: ‘j’ or ‘ʃ’?; are laterals [+continuant] or [–continuant]?) In principle, we need to define what the segmental symbols that we use mean in terms of their featural composition. This may all sound obvious, but the points made here will allow us to say interesting things about RPOs below.

⁷This links to the end of section 1: historical phonology needs to know how whole words are reconstructed, but its interest is in the phonological objects that the words are made up of, not any individual whole word.

⁸Differing traditions use different terminology in subsegmental discussion (‘features’ are also called ‘elements’ or ‘components’ – see, for example, Honeybone 2005), but this should not mask the fundamental agreement that subsegmental objects are the basic unit of analysis. I use the term ‘feature’ here because it is most common.

⁹In Hayes’ system, segments also have specifications for other features (like [consonantal], [anterior] and [high]), which are not necessary to make my points here – these feature-specifications could be handled in exactly the same way as those in (3).

4.2 All analysis of past phonological states is reconstruction

On the argumentation above, forms like ‘*gastiz’ in (1), as RPOs, are cognitive reconstructions. The end of section 2 recognises that it is commonly implied that there are two fundamentally different things in historical phonology: ‘reconstructed forms’ and ‘attested forms’. It seems to follow from this that ‘reconstructed forms’ are RPOs, and ‘attested forms’ are something different. This does not seem right, however. Just as those RPOs which are conventionally marked with an asterisk do not only rely on the comparative method (but are produced on the basis of all possible evidence, following informational maximalism), the forms in (1), (2) and Table 1 which do not have asterisks also represent phonological objects which are deduced on the basis of informational maximalism. If there are written records for phonological objects from the past, this is obviously important evidence for them, but orthography can fail to represent change, and indeed orthography is not intended to represent all of phonology – we need to analyse such evidence carefully (and add to it).

This is clear from how working historical phonologists talk. For example, Minkova (2003:115) works with Old English alliterative poetry as her key evidence, but also recognises that facts observed in contemporary languages are a relevant source of evidence, and that, if we project them “back to early Old English, we can reconstruct a gradient picture: the palatal allophone of /ɣ/ is phonetically extremely close to /j/, while maintaining an allophonic relation with the corresponding velar [ɣ]”. As Minkova says, she is *reconstructing* the phonology of Old English, despite the fact that there is copious written evidence for it. This is not a new point. Janda & Joseph (2013:94) make it, too:

... while many linguists limit their use of the term ‘reconstruction’ to the positing of forms and constructions for linguistic stages from which no records survive, it is actually the case that even attested stages of languages require considerable interpretation and filling-in of details – as well as more substantial aspects. Hence virtually all historical linguistic research merits the descriptor ‘reconstruction.’

So: all analysis of past phonological states is reconstruction. All of the past forms in (1), (2) and Table 1 are RPOs. There is no qualitative difference between the ontology of ‘reconstructed forms’ and ‘attested forms’ – we need to use whatever evidence is available at different stages of a language (comparison, writing, the behaviour of loanwords, *etc.*) to reconstruct the phonological forms (made up of phonological features) which existed in past speakers’ minds.

4.3 Phonological reconstruction and unnecessary asterisks

Everything is now in place to explain why I think that asterisks are unnecessary in historical phonology. We have seen that some RPOs (‘reconstructed forms’) may get marked with asterisks while others (‘attested forms’) may not, and we have seen that this difference is not a principled distinction, anyway: so-called ‘attested forms’ still need to have their phonology reconstructed. This latter point is a spring-board for the discussion in this section: I argue that what historical phonologists do when they reconstruct the phonology of RPOs is not qualitatively different to what non-historical phonologists do when they analyse the phonology of contemporary languages.

At heart, what an asterisk implies is uncertainty about a phonological form. In section 2, we saw the idea that phonological forms without an asterisk have been ‘directly recorded’ (are ‘actually attested’, are assumed on the basis of ‘direct evidence’) – they might be thought

to be certain. We can now be clear, however, given that phonological objects (POs) are cognitive objects, that *no* POs are ever recorded in writing (written forms are just attempts to represent POs), and that there can be no *direct* evidence for POs at all (on currently imaginable technology). Neither RPOs nor CPOs are directly observable or are ever attested outside of the minds of the speakers in which they exist. The only certainty that we can have about phonological forms is if the analysis of them is immensely compelling and if the predictions that the analysis makes are met.

RPOs and CPOs are all simply POs. If RPOs are ‘reconstructed’, then we might say that CPOs are ‘constructed’, but then both ‘reconstruction’ and ‘construction’ just means ‘produced through phonological analysis’. The difference between RPOs and CPOs is simply that phonologists may place emphasis on somewhat different types of evidence in the analysis involved. Both historical and non-historical phonologists use as many kinds of evidence as are available to them in analyses (both follow the maxim of informational maximalism). In terms of RPOs, where written evidence is available, it will naturally be used. Where it is not available, comparative and internal reconstruction (and whatever other forms of evidence may be available) will be used. In terms of CPOs, spoken evidence will surely be used (unless it is unavailable for some reason), but so will other types of evidence, as I discuss below.

For RPOs with orthographic evidence available, such evidence does not render other types irrelevant. Ringe & Eska (2013:282) list 13 types of evidence for the reconstruction of aspects of past phonology, insisting that “every scrap of available evidence must be found and exploited”. They are: our knowledge of what is possible and probable in phonetics; our knowledge of what is possible and probable in phonological structure; statements from native speakers of past languages; metre; rhyme and similar sound-patterns; graphemics; variation in spelling; representation of animal sounds; direct cross-linguistic evidence from written forms; the behaviour of loanwords; knowledge of what has occurred in cases of phonetic change observed in progress; knowledge of what has occurred in cases of phonemic change; traditions of pronunciation of ‘dead languages’.

For CPOs, van Oostendorp (2013) lists 18 types of evidence for the analysis of contemporary phonology. While all are potentially light-shedding, “none is without conceptual problems and furthermore there simply is *no* direct window at present to the human mind. Every piece of evidence we have is therefore potentially polluted by other information” (2013:291). They are: minimal pairs; distribution; phonological alternations; descriptive grammars; corpus frequency; typology; variation data; language change; pathologies; language acquisition; nonce words; artificial learning; language games and secret languages; poetry; phonetic measurement; psycholinguistic experimenting; formal simplicity; modelling.

There is some clear overlap between the two lists (e.g., the use of poetry, typology, theoretical concerns) and there could be more (e.g. the behaviour of loanwords, corpus frequency, alternations/internal reconstruction). Both lists are incisive, but neither is exhaustive (for example, comparative evidence is useful for RPOs even where there is written evidence). The key point is that both show that there is no single central source of evidence for POs (of either type). Those working on either RPOs or CPOs use as much evidence as is available to them: both follow the maxim of informational maximalism. It is true that there is more evidence (and more types of evidence) available for CPOs, but evidence for RPOs can be very good, and – crucially – all evidence for all POs is indirect and needs to be carefully interpreted.

Given all this, it seems clear that there is no *qualitative* difference between the objects that are assumed in cognitive reconstruction and the objects that are assumed in the phonological analysis of contemporary languages. It is not conceptually coherent to mark one kind of PO with an asterisk and not the other – the asterisk may be intended to indicate that a PO

is derived from non-direct evidence, but *all* POs are derived from non-direct evidence. In this sense, the asterisk is unnecessary for RPOs. It is a tautology to mark a PO as being based on non-direct evidence.

Asterisks aside, the basic point that I am making here is not radical: as Hale (2007:246) says, “protolanguages are only partially recoverable given available evidence and existing techniques – but, of course, this holds of synchronic mental grammars as well”. *Partial* recoverability is still important, however, and I argue in the following section that we can be as sure about *some* aspects of a reconstruction as we are about any kind of PO.

5 Case studies

This section returns to issues raised in section 4.1 and develops reasoning from Honeybone (2017), which was inspired by Lass (2017). My point here is that the recognition that asterisks are unnecessary in historical phonology can be taken further, and given empirical weight. If asterisks do indeed imply a kind of uncertainty about a phonological form (as claimed above), that uncertainty can be misplaced in reconstruction once we recognise that the objects to be reconstructed are phonological features. I show this on the basis of two case studies.

5.1 OE back fricatives

Lass (2017) discusses a number of issues relevant to the points considered here. He defends the use of the asterisk for reconstructions (Lass, pc), but adduces a range of points to show that reconstructions need to be seen as real. One of the examples that Lass (2017:158) considers is the segment which occurs (for example) as

... the initial segment of the Indo-European root meaning ‘heart’ and its fate in Germanic. The correspondence set {Latin *cord-*, Old Irish *críde*, Old English *heorte*, Old Norse *hjarta*}, using standard procedures, makes us reconstruct PIE *kerd-, and then via Grimm’s Law Proto-Germanic *xert-.

The point about this segment is that, while it is attested only as [h] in this environment in contemporary Germanic languages, we must reconstruct a diachronic pathway from PIE to English along the lines of $k > x > h$, if we respect evidence of the type that was above described on van Oostendorp’s (2013) list as ‘typology’ and could be related to ‘knowledge of what has occurred in cases of change’ from Ringe & Eska’s (2013) list.¹⁰ Lass recognises that, although ‘attested forms’ are involved, understanding OE requires phonological reconstruction, and argues that this case is interesting because the reconstruction (the segmental RPO) is uncertain:

In early Old English the orthographical *h-* may have stood for either [x] or [h]; in late Old English there is no doubt that it was [h], as it is now in those English dialects where it survives.

This seems problematic for my case. How can I argue that this RPO (the initial segment in the early OE word for ‘heart’, and also other words such as that for ‘help’, ‘high’ and ‘horn’) is the same kind of thing as the CPO which is the initial segment in contemporary

¹⁰ The [x] stage is also evidenced by the fact that the PIE stop *must* have had [x] reflexes in non-initial environments – this is shown by the fact that dorsal fricatives are still preserved in closely-related languages like German, for example, and Scots: e.g., *lauch* [lax] ‘laugh’, *dochter* [doxtər] ‘daughter’, both of which words had [k] in the relevant form of (late) pre-Gmc PIE.

English words *heart*, *help*, *high* and *horn*, which (in varieties which have an initial consonant) is unambiguously [h]? Doesn't the RPO need an asterisk to indicate the uncertainty?

My answer returns to section 4.1 and links to the range of entities which are accepted as CPOs. The discussion in that section recognises that *features* are the basic phonological unit, and with this in mind, there is no uncertainty that the early OE segmental RPO in question is, for example, [+continuant], [–sonorant] and [–nasal], because both glottal and dorsal fricatives have those specifications. On Hayes' system, both are also [–voice] and [–constricted glottis], and [–labial] and [–coronal], too. Given that [h] is [+spread glottis] and [–dorsal] and, again on Hayes' system, [x] is [–spread glottis]¹¹ and [+dorsal], however, we do not know for sure what these feature-specifications for the segment were, but this does not mean that we are uncertain about the other specifications. On this basis, the segment can be understood as in (4), where '[—]' represents an absent specification.

- (4) **P:** [–labial] [–coronal] [—]
M: [+continuant] [–sonorant] [–nasal]
L: [–voice] [—] [–constricted glottis]

We could reasonably use the symbol 'h' (although we could also use 'x') to represent the early OE PO in (4) – we just need to define what the symbol means, as we should expect, in line with the discussion at the end of section 4.1. Crucially for my argument, the RPO represented in (4) is the same as what we encounter in the phonological analysis of CPOs where underspecification is assumed. One case that is close to that seen in (4) is the analysis that Hall (1989) proposes of *ich-laut* and *ach-laut* in contemporary standard German. This case is quite well-discussed in the literature – fundamentally, there is morphophonologically predictable distribution of [ç] and [x] (and alternations involving the two), with [ç] occurring following front high vowels, as in [zi:ç] *siech* 'ill', and in certain other environments, and with [x] occurring following back vowels, as in [bu:x] *Buch* 'book', and in certain other environments. Hall (1989) argues that these dorsal fricatives are best analysed as deriving from an underlier which is neither /ç/ nor /x/, but is rather a segment which is [+continuant] and fortis and which is *not specified* for backness (because it gets these place features filled in during a derivation).¹² This can be represented as in (5), if we translate Hall's analysis into Hayes' feature system (in which [ç] is [+coronal] and [+dorsal], while [x] is [–coronal] and [+dorsal]), and follow Hall's argument that the features determining 'back' place in the segment are underlyingly non-specified.

- (5) **P:** [–labial] [—] [—]
M: [+continuant] [–sonorant] [–nasal]
L: [–voice] [–spread glottis] [–constricted glottis]

The two cases in (4) and (5) – one representing an RPO and the other a CPO – are really establishing the same level of certainty about phonological objects: the feature specifications

¹¹ There are good reasons to think that [x], at least in a language like OE (and also in German, discussed next), is in fact [+spread glottis], but I stick with Hayes' system here, for simplicity. If we were to assume the [+spread glottis] alternative, the basic point would still hold, but only the specification for [dorsal] would be absent in (4).

¹² It is not necessary to accept this analysis of German *ich-laut-ach-laut* for my arguments to hold, of course. Any case of synchronic underspecification or non-specification would do.

set out in both are firmly established, and those which are missing (‘[—]’) leave the segments underspecified.¹³ Underspecified segments are still real, however. Each of the feature-specifications in (4) can be seen as firmly reconstructed phonological objects for early OE. All the specifications in (4) and (5) are cognitive objects, all of which are equally unobservable, because they either existed in the minds of speakers of early OE or because they exist in the minds of speakers of contemporary German. Although they may be treated slightly differently if we consider entire phonological derivations (as considered in footnote 13), the first segment of early OE *heorte* and the last segment of German *siech* have exactly the same status at the relevant level of analysis on this argumentation – it seems unnecessary to mark one of them with an asterisk.

5.2 Proto-Indo-European stops

The reconstruction of PIE stops used up till now in this article (given, for example, in the first column of Table 1) is a traditional one. As is well-known, however, this reconstruction has been subject to considerable challenge. This makes the case highly relevant to our purposes: if the RPOs reconstructed for PIE stops are uncertain, how can we assume that they are ontologically the same as CPOs? It is widely agreed that there were three series of stops, but linguists have argued at length about the segments’ laryngeal specifications. The best known alternative reconstruction is the ‘Glottalic Theory’ of Gamkrelidze & Ivanov (1973), who argue that the series were not of the types /t, d, d^h/ (again exemplifying using coronals), but were rather /t^h, t^ʰ, d^h/.¹⁴ Proposals of this type have been made for several reasons, in part centring around the idea that the traditional reconstruction is typologically highly unlikely because we do not expect to find languages with a breathy voiced (‘voiced aspirated’) series, like /d^h/, if they do not also have a voiceless aspirated series, like /t^h/; ‘glottalic’ reconstructions also account better for the infrequency of the labial in this series and for certain segmental co-occurrence restrictions in roots.

Gamkrelidze & Ivanov’s proposal has found some acceptance, but has also met some rejection. Salmons (1993:72) argues that “there is some evidence to support the claim that the Glottalic Theory is becoming the standard view”, but Clackson (2007:47) is cautious, noting that “the traditional paradigm remains resolutely in place” for many scholars. We are left with a situation where the precise laryngeal specifications for the PIE stops can hardly be called certain – there are multiple proposals for their reconstruction, each with some merit.

¹³ Charles Reiss has insightfully pointed out that there are two types of rationale for the absences involved in the phonological objects that are considered here (both indicated by [—]). In (4), specifications are absent because we don’t know what they are (an epistemological issue), while in (5) there is intentional ontological absence. This is certainly true, but I think my point is that the *representations* involved are the same. The ‘[—]’ might be there in a representation for different reasons, but the representations themselves don’t know that. We, as phonologists, may grasp this distinction, but the representations don’t know if they are ‘intentionally’ or ‘unintentionally’ underspecified. This difference may be relevant (and recorded) at some point of a derivation: intentionally underspecified segments may get a feature specification before the surface, but it seems to me that this is not relevant to the status of the underspecified representation itself at the relevant point in the derivation. On this basis, and at the relevant points in the phonological analyses involved, the absences in (4) and (5) are the same: [—] simply marks that something is absent.

¹⁴ Gamkrelidze & Ivanov (1973) in fact use the symbols /th/ and /dh/ for the first and third series – the superscript aspiration used here is just a typographical issue. They describe the segments as “voiceless aspirates” and “voiced aspirates”, and, while they consider whether the aspiration should be seen as phonologically redundant, they insist that it was a “relevant feature” in the segments. It is not entirely straightforward to translate their position into the categories used in contemporary laryngeal phonological theory. I assume a position implying that the segments were underlyingly [+spread glottis].

All this in fact fits well with the line of argumentation developed in this article. We may not know the segments' laryngeal specifications, but we do know beyond doubt their specifications for manner, and the place of those listed in Table 1 is also firmly reconstructed. I set this out in Table 2, which shows (with some feature-names abbreviated) the feature-specifications which can be firmly reconstructed on the basis of agreement between the 'traditional' and 'glottalic' approaches (as in Table 1, this only includes three of the typically reconstructed series of stops in PIE, leaving out, for example, the labiovelars – including other stop series would complicate things, but would not change the fundamental point).

Traditional	Glottalic	firmly reconstructed
p	p ^h	P: [+labial] [–coronal] [–dorsal] M: [–cont] [–son] [–nasal] L: [–voice] [–] [–constr]
b	p ^ʰ	P: [+labial] [–coronal] [–dorsal] M: [–cont] [–son] [–nasal] L: [–] [–spread] [–]
b ^h	b ^h	P: [+labial] [–coronal] [–dorsal] M: [–cont] [–son] [–nasal] L: [+voice] [+spread] [–constr]
t	t ^h	P: [–labial] [+coronal] [–dorsal] M: [–cont] [–son] [–nasal] L: [–voice] [–] [–constr]
d	t ^ʰ	P: [–labial] [+coronal] [–dorsal] M: [–cont] [–son] [–nasal] L: [–] [–spread] [–]
d ^h	d ^h	P: [–labial] [+coronal] [–dorsal] M: [–cont] [–son] [–nasal] L: [+voice] [+spread] [–constr]
k	k ^h	P: [–labial] [–coronal] [+dorsal] M: [–cont] [–son] [–nasal] L: [–voice] [–] [–constr]
g	k ^ʰ	P: [–labial] [–coronal] [+dorsal] M: [–cont] [–son] [–nasal] L: [–] [–spread] [–]
g ^h	g ^h	P: [–labial] [–coronal] [+dorsal] M: [–cont] [–son] [–nasal] L: [+voice] [+spread] [–constr]

Table 2: reconstruction of PIE stops at three places of articulation

These underspecified stops will likely need laryngeal feature specifications in order to be realised (unless the underspecification perseveres, as in the cases considered by Keating 1988 and Hale & Kisko 2007), and each analyst will need decide how/whether to fill them in (and decide on a set of symbols to represent them), but the RPOs in Table 2 can be seen to have the same status (and lack of direct observability) as those in (4) and (5).

6 Conclusion

If what I have argued here holds, RPOs are as real as CPOs, even if those RPOs may be underspecified (just like CPOs). The two are deduced on the basis of partially different kinds of evidence, but both types of segmental POs have the same status: they are cognitive objects, which are composed of feature-specifications. The conclusions reached here need not change our practice as historical phonologists, but they may have consequences in some small ways. In any case, it is worthwhile reflecting on our conception of what we do and of what entities we deal with. To conclude, it is not clear to me when in the history of a language we should *stop* using an asterisk to mark phonological forms, if we use it at all: when it was first written? or first attested in large amounts? or attested for all dialects? for forms from a century ago? a decade ago? last week? I am not sure that there is a principled basis on which to decide this. Perhaps the asterisk is unnecessary.¹⁵

References

- Beade, Pedro. 1989. Falsification and falsifiability in historical linguistics. *Philosophy of the Social Sciences* 19:173–181.
- Bloomfield, Leonard. 1925. On the sound-system of Central Algonquian. *Language* 1/4:130–156.
- Bloomfield, Leonard. 1928. A note on sound change. *Language* 4/2:99–100.
- Brugmann, Karl. 1876. Nasalis Sonans in der indogermanischen Grundsprache. In *Studien zur griechischen und lateinischen Grammatik*, ed. Georg Curtius and Karl Brugmann, 9:285–338. Leipzig: Hirzel.
- Bynon, Theodora. 1977. *Historical linguistics*. Cambridge: Cambridge University Press.
- Campbell, Lyle. 1996. On sound change and challenges to regularity. In *The Comparative Method reviewed: regularity and irregularity in language change*, ed. Mark Durie and Malcolm Ross, 72–89. New York and Oxford: Oxford University Press.
- Chomsky, Noam. 1986. *Knowledge of language: its nature, origin, and use*. New York: Praeger.
- Clackson, James. 2007. *Indo-European linguistics: an introduction*. Cambridge: Cambridge University Press.
- Fellner, Hannes and Nathan Hill. 2019. Word families, allofams, and the comparative method. *Cahiers de Linguistique Asie Orientale/East Asian Languages and Linguistics* 48/2:91–124.

¹⁵ Pavel Iosad has nicely pointed out to me that, if I'm right about all this, then it would be just as consistent to use asterisks for both CPOs and RPOs as it would be not to use them for RPOs – to mark the fact that both CPOs and RPOs are unobservable – and notes that it might be a useful reminder about their unobservability if we did so. I agree. However, phonological objects are typically indicated as such in the analysis of contemporary languages in some other way (with slanted brackets, for example, or by placing them in some pre-phonetic part of a derivation), so asterisks are not *necessary* to mark POs in this way, either.

- Fox, Anthony. 2015. Phonological reconstruction. In *The Oxford handbook of historical phonology*, ed. Patrick Honeybone and Joseph Salmons, 49–71. Oxford: Oxford University Press.
- Gamkrelidze, Thomas and Vjačeslav Ivanov. 1973. Sprachtypologie und die Rekonstruktion der gemeinindogermanischen Verschlüsse. *Phonetica* 27:150–156.
- Graffi, Giorgio. 2002. The asterisk from historical to descriptive and theoretical linguistics: an historical note. *Historiographia Linguistica* 29/3:329–338.
- Hale, Mark. 2007. *Historical linguistics: theory and method*. Malden: Blackwell.
- Hale, Mark. 2012. Foundational concepts in the scientific study of sound change. In *The initiation of sound change: perception, production, and social factors*, ed. Maria-Josep Solé and Daniel Recasens, 235–246. Amsterdam and Philadelphia: John Benjamins.
- Hale, Mark. 2015. The Comparative Method: theoretical issues. In *The Routledge handbook of historical linguistics*, ed. Claire Bowerman and Bethwyn Evans, 146–160. Abingdon and New York: Routledge.
- Hale, Mark and Madelyn Kisson. 2007. The Phonetics-Phonology Interface and the Acquisition of Perseverant Underspecification. In *The Oxford handbook of linguistic interfaces*, ed. Gillian Ramchand and Charles Reiss, 81–101. Oxford: Oxford University Press.
- Hale, Mark, and Charles Reiss. 2000. Phonology as cognition. In *Phonological knowledge: conceptual and empirical issues*, ed. Noel Burton-Roberts, Philip Carr and Gerard Docherty, 161–84. Oxford: Oxford University Press.
- Hale, Mark, and Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.
- Hall, Tracy Alan. 1989. Lexical Phonology and the distribution of German [ç] and [x]. *Phonology* 6:1–17.
- Hammarberg, Robert. 1976. The metaphysics of coarticulation. *Journal of Phonetics* 4:353–363.
- Hayes, Bruce. 2009. *Introductory phonology*. Oxford: Wiley-Blackwell.
- Honeybone, Patrick. 2005. Sharing makes us stronger: process inhibition and segmental structure. In *Headhood, elements, specification and contrastivity: phonological papers in honour of John Anderson*, ed. Philip Carr, Jacques Durand and Colin Ewen, 167–192. Amsterdam and Philadelphia: John Benjamins.
- Honeybone, Patrick. 2012. History and historical linguistics: two types of cognitive reconstruction? In *Language and history, linguistics and historiography*, ed. Nils Langer, Steffan Davies & Wim Vandenbussche, 15–47. Bern: Peter Lang.
- Honeybone, Patrick. 2016. Are there impossible changes? $\theta > f$ but $f \not> \theta$. *Papers in Historical Phonology* 1:316–358.
- Honeybone, Patrick. 2017. (On realism in phonological reconstruction.) Comment on 'Reality in a soft science: the metaphonology of historical reconstruction' by Roger Lass (*Papers in Historical Phonology* 2).
- Janda, Richard and Brian Joseph. 2003. On language, change, and language change – or, of history, linguistics, and historical linguistics. In *The handbook of historical linguistics*, ed. Brian Joseph and Richard Janda, 3–180. Malden and Oxford: Blackwell.
- Keating, Patricia A. 1988. Underspecification in Phonetics. *Phonology* 5:275–292.
- Koerner, E.F.K. 1976. Zu Ursprung und Geschichte der Besternung in der historischen Sprachwissenschaft. Eine historiographische Notiz. *Zeitschrift für vergleichende Sprachforschung* 89/2:185–190.
- Lass, Roger. 1997. *Language change and historical linguistics*. Cambridge: Cambridge University Press.

- Lass, Roger. 2017. Reality in a soft science: the metaphonology of historical reconstruction. *Papers in Historical Phonology* 2:152–163.
- Lichtenberk, Frantisek. 1994. Reconstructing heterogeneity. *Oceanic linguistics* 33:1–36.
- Longobardi, Giuseppe and Cristina Guardiano. 2009. Evidence for syntax as a signal of historical relatedness. *Lingua* 119:1679–1706.
- McMahon, April. 1994. *Understanding language change*. Cambridge: Cambridge University Press.
- Meillet, Antoine. 1903 to 1937 (first to eighth edition). *Introduction à l'étude comparative des langues indo-européennes*. Paris: Hachette.
- Minkova, Donka. 2003. *Alliteration and sound change in early English*. Cambridge: Cambridge University Press.
- Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Bern and Munich: Francke.
- Pulgram, Ernst. 1959. Proto-Indo-European Reality and Reconstruction. *Language* 35/3:421–426.
- Reiss, Charles. 2018. Substance free phonology. In *The Routledge handbook of phonological theory*, ed. S.J. Hannahs and Anna Bosch, 425–452. Abingdon and New York: Routledge.
- Ringe, Don and Joseph Eska. 2013. *Historical linguistics: toward a twenty-first century reintegration*. Cambridge: Cambridge University Press.
- Ringe, Don and Ann Taylor. 2014. *A linguistic history of English, volume II: the development of Old English*. Oxford: Oxford University Press.
- Salmons, Joseph. 1993. *The Glottalic Theory: survey and synthesis*. Journal of Indo-European Studies Monograph 10. Institute for the Study of Man.
- Saussure, Ferdinand de. 1879. *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. Leipzig: Teubner.
- Schleicher, August. 1861. *Compendium der vergleichenden Grammatik der indogermanischen Sprachen*. Weimar: Böhlau.
- Trask, R.L. 1996. *Historical linguistics*. London: Arnold.
- van Oostendorp, Marc. 2013. A consumer guide to phonological evidence. *Nordlyd* 40/1:274–293.
- Walkden, George. 2019. The many faces of uniformitarianism in linguistics. *Glossa* 4:52/1–17.
- Weiss, Michael. 2015. The Comparative Method. In *The Routledge handbook of historical linguistics*, ed. Claire Bowerman and Bethwyn Evans, 127–145. Abingdon and New York: Routledge.

Patrick Honeybone
Linguistics and English Language
University of Edinburgh
patrick.honeybone@ed.ac.uk