Evidence for Basque as an Indo-European Language: A Reply to the Critics

Gianfranco Forni

Why should the reader bother to read this article?
As it is obvious that

A) this article is no light reading: it’s very long, and it’s full of arcane sound laws, detailed etymologies, annexes with pages after pages of further details;
B) many readers of JIES may not be familiar with Basque;
C) several commentators peppered their comments with colorful and overtly demeaning terms;¹
D) all commentators heavily criticized this article – except Koch;

why should readers bother to take this article seriously and invest several days to study it in any detail? I hope the reasons provided below will justify the reader’s attention to the article and not just the comments:

• if you scratch the surface of colorful rhetoric in most comments, you will often find preciously few hard facts beneath this superficial gloss: several sweeping and categorical claims are based on extremely limited evidence; for instance, most commentators do not attack any meaningful amount of my “likely” etymologies (i.e., etymologies #1 to #142 in my article): criticisms to specific likely etymologies range from 0% in Bengtson, to 8% in Koch, 13% in Kassian, 16% in Prósper, and 46% in Gorrochategui

¹For example, “illustrious amateurs”, “visionary”, “Bongo-Bongo method”, “amateur unaware of le métier”, “the current work is the author’s first”, “a beginner”, “amateur theories”, to cite just Gorrochategui, Lakarra and Kassian.
and Lakarra (the only comment with a sizable amount of specific criticisms\(^2\));

- few etymologies attract consistent criticism: just 20% of my “likely” etymologies are criticized by more than one commentator; if most of my etymologies were obviously and hopelessly wrong, you’d expect most commentators to criticize the same, numerous blunders – but this was not the case;
- my final comments below address all key criticisms (including detailed ones concerning individual etymologies) and I hope you may find quite a few of my “counter-comments” reasonably convincing;
- an article should be judged by its own merits – not based on the author’s background and track record.

In a similar vein, why should the author accept to have this article published in JIES in spite of harsh criticism by most commentators? I have two main reasons:

- as you will see below, much of this loud criticism does not stand detailed, fact-based, quantitative scrutiny;
- publishing in JIES ensures a wide circulation of my work among the world’s linguists; this is a unique opportunity I don’t want to miss, as I’m confident that professional readers will be able to form their own, personal opinions about it, and some may end up endorsing it, at least in part.

Structure of these final comments

In the following pages I will systematically reply to key comments I received, first lumping and addressing together topics that several commentators have in common, then addressing specific comments by individual commentators.

1. Responses to common comments

Goals and methodology

The goal of my article: classification, pure and simple

Several commentators (e.g., Bengtson, Gorrochategui

\(^2\)Hence I will devote more space to Gorrochategui and Lakarra’s comments, than to other commentators.
and Lakarra, but not Prósper\(^3\) seem to misconstrue the key goal of my article. My goal is solely to **classify** Basque, i.e., to gather adequate evidence that it belongs to a known language family (IE, as it happens). I only intend to provide evidence for the “**what**” side of this question,\(^4\) i.e., **what** family Basque belongs to; I believe it is absolutely premature to venture into wild conjectures about the “**why**” side, trying to answer such questions as “**what** PIE sub-family does it belong to?”, “when and how did it diverge from PIE?”, “how and when was it in contact with other languages during its prehistory?”. Such questions have proven formidably hard to answer even for well-established language families (say, PIE, Semitic or Uralic); debate over such controversial topics as subgrouping, “glottochronology”, *Urheimat*, language contacts and loans has been hot and is often still unsettled, sometimes decades or even centuries after a language family was recognized.

Every first attempt at language classification, even if eventually successful, inevitably suffers from errors and imprecisions, which may take decades to correct and refine, without making the initial classification invalid. I think I gathered enough evidence to claim that most of Basque native, basic lexicon has an IE origin, and I report this finding as a fact, but I deliberately refuse to venture into the minefield of creating conjectures to try and explain the detailed causes of this fact. After all, the origin, chronology and homeland of Indo-Europeans are still hotly debated after two centuries,\(^5\) but the fact that these questions are still unsettled is not jeopardizing the validity of the IE language family *per se*. I prefer to stick to

---

\(^3\)Prósper writes: “The […] question concerning us here […] is not where speakers of Basque come from or when they first met with speakers of Indo-European languages, but whether the language itself can be classified as Indo-European.”

\(^4\)To quote Kassian, my goal is “just to establish a mere classification”.

\(^5\)In his comments, Koch offers an enlightening summary of the still sorry state of theories about the pre-historical origins of the (linguistically very well-established) IE family: “the answer will depend on which theory of the origin and dispersal of the Indo-European languages is selected” as alternative theories provide very different timings, ranging from 3rd millennium BC to around 6,000 BC to the Ice Age.
experimentally observable facts and not venture into wild
guessses about their possible causes.

In particular, I do not claim that Basque is a member of
the Celtic sub-family; I just say it seems to share several
terms and sound laws with (parts of) it; but this may well be
due to areal contact. I find Kassian’s lexical evidence and
Koch’s phonological evidence against Basque as a Celtic
language rather convincing; I see their comments as a
useful contribution to a nascent debate about where Basque
belongs within IE: based on their preliminary comments, it
seems that a likely answer can be: “not with Celtic”.

How important is grammar for language classification?

First of all, in sub-section “Methodology used in the
present research” of my article, I wrote

[...] to prove that language L belongs to family F,
one must prove that the majority of native [...] basic
lexicon and grammar of L can be derived from the
lexicon and grammar of proto-F through regular
sound changes

but this should be expressed, more precisely, by replacing
“grammar” with “grammar morphemes”, i.e.:

[...] to prove that language L belongs to family F,
one must prove that the majority of native [...] basic
lexicon and grammar morphemes of L can be derived
from the lexicon and grammar morphemes of proto-F
through regular sound changes

Many commentators (e.g., Gorrochategui, Lakarra
and Bengtson) insist that grammar morphemes should play
a key role in language classification – though they also
admit that they often are short, few, and unstable, and
Kassian explicitly states that a match in basic vocabulary is
stronger evidence than a match in grammar morphemes,
and the latter can serve as additional proof. In section 6
(“Discussion, Conclusions and Next Steps” – sub-section

---

6“Proto-Celtic and its Proto-Brittonic descendant are decisively excluded
as ancestors of Basque by at least 10 of the sound changes between Proto-
Indo-European and Pre-Basque proposed by Forni”.

The Journal of Indo-European Studies
“Grammar Etymologies”) of my article I addressed this topic already. I agree, in principle, on the highly diagnostic value of grammar morphemes for classification efforts, but sometimes this “luxury” is not available; if grammar morphemes of language L disagree with family F, but lexicon agrees, can one derive no valid conclusions about its classification? Suppose you find a language L whose basic lexicon is overwhelmingly from family F, but whose grammar morphemes are at odds with family F. What do you conclude? I can only conclude that language L belongs to family F. Alternative explanations seem less likely: I know of no language that belongs to a family but borrowed most of its basic lexicon (across all semantic areas) from one or more languages outside that family – except in the case of creoles, where the concepts of “family” and “loan” becomes very blurred.

**How important is typology for language classification?**

The consensus answer among historical linguists is: typology is practically irrelevant for language classification. Therefore, typological remarks in sections 2 and 3 in Gorrochategui and Lakarra’s comments are irrelevant.

**How many lexical etymologies do you need to classify a language?**

Language classification usually starts from basic vocabulary, which tends to be more stable and less subject to loans. All commentators seem to agree on this. That’s why I’ve been reluctant to expand my etymologies to less-than-very-basic vocabulary, and I’ve added some 40 further etymologies only as an appendix. My goal is not to provide an etymological dictionary of all 500+ native Basque terms identified by Trask: I intentionally limited my analysis to what I consider to be basic lexicon, and discarded the rest from the scope of my research. Interestingly, commentators themselves seem to seriously disagree about the “proper” number of lexical items needed for language classification.

---

7Pace Prósper, I believe that personal and place names are not part of basic vocabulary; that’s why I intentionally did not include them in my lexical evidence.
classification purposes: suggested amounts vary widely across comments, from 50 (Kassian⁸) to 1,000 (Koch), i.e., by an order of magnitude. For additional details on the choice of basic lexicon, see my reply to “Objection: to classify Basque as IE, basic lexicon is insufficient: many more etymologies are needed, in order to provide a statistically robust sample” in section 6 (“Discussion, Conclusions and Next Steps”) of my article.

**How many sources do you need to classify a language?**

Gorrochategui and Lakarra (section 4) and Prósper protest I should have used more sources than I did. I think what matters is not how many sources one uses, but how incrementally useful each additional source is. Suppose that using six sources provides 90% of the information you need, and each additional source provides an extra 1%, or less, of useful information, but requires an extra 10% of effort to be taken into account.⁹ to maximize output-to-effort ratio, for both the researcher and readers, it may be more efficient to stop adding extra sources when the ones you collected are “good enough” already, for most purposes. This was my choice. Anyway, if the sources I overlooked are fundamental, the key information they contain has probably been incorporated into the comments I received, so by responding to these comments, I should be able to take into account such key extra data from important sources I did not use.

By the way, for these final comments I add one more source:

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Source</th>
</tr>
</thead>
</table>

As for citing sources properly, Prósper writes that when I cite Pre-Basque reconstructions, “one is often at a loss as to whether some changes are widely acknowledged or just F.’s hypotheses”; actually, in my article, each Pre-

---

⁸And Kassian even uses as few as 11 lexical items to reject the assignment of Basque to Celtic.

⁹I use these simplified, fictitious percentages just to explain my line of reasoning – they should not be understood as “real”, measurable values.
Basque reconstruction is always followed by its source (T, T-E, or both); as explained in subsection “Introduction to lexical etymologies” of my article, in the rare cases when I came up with a Pre-Basque reconstruction that is different from Trask’s, I explicitly marked it by putting “recte” before my own reconstruction. There are 24 such cases in my 142 “likely” etymologies, i.e., my Pre-Basque reconstructions differ from Trask’s in 24/142 = 17% of the cases (in the “likely” etymology section).

Why did nobody realize that Basque is IE?

Many comments can be summarized as “I can’t believe it” or “it can’t be true”: if so many linguists have researched the origins of Basque and they all failed, how can a “beginner” with “half a dozen handbooks” succeed? On the other hand, my etymologies are often described as “non obvious” or “non transparent”. That’s precisely why previous classification attempts have failed: there are few obvious similarities between Basque and IE in general, vs. so many obvious differences (including many typological ones, which we know are irrelevant, though). I think what happened is the following:

- Basque looks un-IE, so few researchers ever bothered to look at IE as a possible origin;
- direct comparisons between modern Basque and reconstructed Proto-IE can find very few promising similarities and a lot of obvious differences, so even those who tried probably stopped short of finding anything useful;
- based on this research (and others I started) I realized that, if an isolated language is very different from a given proto-language, but ultimately belongs to the family the proto-language is reconstructed for, a different approach can be more productive:
  1. compare the isolated language with surrounding languages of the target family, too – not just with the reconstructed proto-language; areal phenomena may have led to similar historical phonetics, so similarities may be more recognizable;
  2. if “obvious” similarities in both form and meaning are hard to find, you have three
alternatives:
a) accept that the comparison has failed;
b) compare words that have a similar phonetic shape, but fairly different semantics, and accept lax semantics to salvage these cases as valid comparisons; or
c) be strict with semantics and look for diverging phonetics instead.

In broad terms, historical linguists often adopt approach “a)” if they are so-called “splitters”, or approach “b)” if they are “lumpers”; pace Bengtson\(^{10}\), I think approach “c)” is better; after all, if you consider how phonetically divergent many modern IE languages are, and how different reflexes of a same PIE root have become, approach “c)” is de facto the one that has been often productively used for the IE family so far.

**Is a classification invalid if parts of its sound laws are still sketchy?**

I believe it is not. It is typical of the very first proposals of new classifications to only provide a first sketch of sound laws, without fully identifying conditioning factors that might turn apparent irregularities into regularly-conditioned changes.\(^{11}\) It is also typical that such first sound laws focus on consonants much more than vowels: the latter may take decades to settle; think for example of how vowels in Schleichert’s tale changed in one century (M-A 45-47), from its original 1868 version, to Hirt’s 1939 update, to Lehmann and Zgusta’s 1979 revision; or think of PIE’s laryngeal theory (or rather, theories);\(^{12}\) none of these major revisions of reconstructed

---

\(^{10}\) “Sound laws” (phonetic correspondences) are secondary, arising from the obvious etymologies already recognized as plausible evidence for a linguistic relationship. It seems that Forni has reversed the logical order of these steps, comparing Basque words with IE words of a similar meaning, and forcing them into “cognacy” with a series of implausible “sound laws”.

\(^{11}\) For example, Verner’s law was established decades after Grimm’s law for Germanic (adducing IE stress as a conditioning factor for the first *Lautverschiebung*).

\(^{12}\) Also see, e.g., Prósper’s comments on current disagreements about PIE
Evidence for Basque as an Indo-European Language: 
A Reply to the Critics

PIE vowels resulted in a rejection of previous classifications of IE languages as belonging to the IE family.\textsuperscript{13}

Is it really “too easy” to find matches, if some initial consonants were lost?

Prósper alleges that the posited loss of initial consonants “opens the way to more or less any etymology: you only have to posit a PIE word, always beginning with a consonant, and claim it has been lost\textsuperscript{14} (if unaccented, the whole syllable is finally lost), which demonstrates it was PIE after all. This bears fruit because it is designed never to fail to do so”, and so the Basque term “always complies because half the root has been swept away”. This is a misrepresentation of my sound laws: I do not posit the loss of all initial consonants (only of *m- and some plosives: see sound law sets #2,6,7,8,17), nor of all initial unaccented vowels (only of *e-: see sound law set #22). For a detailed, quantitative discussion of lost vs. retained phonemes from PIE to Pre-Basque, see sub-section “Sound laws” in section 6 “Discussion, Conclusions and Next Steps”, where I show that “PIE reconstructions are long (over 4 phonemes on average) and Pre-Basque reflexes preserve almost 3 phonemes on average. Both values are well above the threshold of just 1 or 2 matching phonemes, which might engender the suspicion that matches are due to chance”.

Are semantics in my comparisons too lax?

Commentators disagree here: some say my semantics are too strict, but my sound laws are too odd; others contend that “the great part of matches is semantically either optional or odd” (Kassian) or “As a mere statistic reflecting on the rigour of the method as applied to vowel and laryngeal systems.

\textsuperscript{13}Similarly, in the Semitic field we still don’t have a full account of vowels, but no specialist seems to believe this hinders the recognition of language family members; as a further example, after key aspects of Uralic consonant treatment had been established by Setälä and Budenz, it took about a century to achieve a solid treatment of Uralic vocalism.

\textsuperscript{14}Oddly, Prósper also writes “he departs from the basic assumption that many sounds were forbidden in absolute initial position in (Pre-)Basque”, but cites no specific examples: I think exactly the opposite is true, but if I made mistakes, I’ll be glad if she points them out.
semantics, I note that there are 122 examples where the Basque and Proto-Indo-European do not match exactly” (Koch) – though neither cites specific examples, unfortunately. On this topic, see annex 5 in my article, which shows that only 9% of “likely” etymologies are based on a loose semantic match.

**How should PIE affixes be legitimately used in a classification exercise?**

Kassian15 and Prósper16 notice that some combinations of PIE roots and affixes I used in my etymologies have limited or no attestations in other IE languages; they seem to imply that this undermines the credibility of such etymologies. Any IE language can (and does) use IE affixes, so the frequently reconstructed presence of IE affixes in my etymologies should be a further confirmation that Basque is IE (also see last objection on lexical etymologies in section 6 “Discussion, Conclusions and Next Steps”, as well as annex 3 “Root vs. non-root etymologies”).

However, not all combinations of PIE roots and affixes can legitimately be projected back to the PIE stage: indeed, some may be later innovations in specific languages or sub-families, where such IE affixes were still productive and were used in new combinations that cannot, strictly speaking, be reconstructed for the common PIE stage, based on available evidence.

In my etymologies, I failed to explicitly distinguish cases where a combination of a root and affixes could reasonably projected back to (i.e., reconstructed for) common PIE, from cases were such combination was more likely to be a later innovation – though based exclusively on IE morphemes anyway.

In the following cases, for instance, it is likely that PIE affixes reconstructed for Proto-Basque are local innovations, that cannot probably be projected back to PIE, but consist of PIE morphemes anyway:

15 For example, “Morphology is uncontrolled”; “ad hoc affixation and compounding”.

16 For example, “A suffix or enlargement is then chosen so as to match the result to satisfaction”.

*The Journal of Indo-European Studies*
<table>
<thead>
<tr>
<th>Etymology</th>
<th>Criticized by</th>
<th>Further comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>15. <em>bi, biga</em></td>
<td>Gorrochategui and Lakarra, Prósper</td>
<td>Final –a in Basque <em>biga</em> may be a Basque-internal (possibly Proto-Basque) ending, so we’d have PIE <em>dwén(u)</em>(^ {17}) (M-A 97, 308) &gt; Proto-Basque *beu + *-a &gt; <em>beva</em>; compare, e.g., Old Irish <em>da</em> ‘two (neuter)’ (M 110)</td>
</tr>
<tr>
<td>17. <em>bizi</em></td>
<td>Prósper</td>
<td>V 685 reconstructs PIE <em>gwíh₃</em>-ue/o- ‘to live’ and *gwíh₃-uó- ‘alive’, so PIE *gwíh₃-wo- can be safely reconstructed; *-to- / -tV- extensions are also amply attested, e.g., in Sanskrit *jivá- ‘living’, Lith. <em>gyváta</em> ‘life’, etc.; i.e., the combination of PIE root <em>gwíh₃</em>- with extensions *-wo- + *-to- is commonplace in IE</td>
</tr>
<tr>
<td>74. <em>otz, hotz</em></td>
<td>Kassian</td>
<td>*-en- stem, as Armenian <em>otn</em> &lt; <em>phóten</em> (according, e.g., to Pisani and Belardi, who, <em>contra</em> Meillet <em>et al.</em>, do not think that Armenian –n &lt; PIE accusative <em>-m</em> here)</td>
</tr>
<tr>
<td>93. <em>oin</em></td>
<td>Kassian</td>
<td></td>
</tr>
</tbody>
</table>
| 95. *orri* | Kassian | *
| 107. *(h)il-* | Kassian | *
| 32. *ez* | Kassian | likely a later formation from a |

\(^{17}\) With *h₃* (not *h₁*, as I mistakenly wrote in my etymology).  
\(^{18}\) Kassian writes “Luwian *munula-s* ‘moon’ simply does not exist”; it’s actually hard to find in common sources (such as K); the source is Stuart E[ward] Mann, *An Indo-European Comparative Dictionary*, Hamburg: Helmut Buske Verlag, 1984-1987, col. 759.
Proto-Basque *en ‘not’ (< PIE *ne, *n ‘not’) + Proto-Basque *si ‘is’ (< PIE *h₁es-ti ‘is’), with a perfect parallel in Goidelic and Brittonic ni ‘not’ < *ni ess < *ni esti ‘it is not (the case that)’

108. labur  Kassian, ‘short’  Prósper M-A 319 reconstructs *mrghus ‘short’ for Latin brevis ‘short’ etc.; a (diminutive) *-IV- extension may well be a Proto-Basque innovation

125. hor-, hura, har-, haie- 'that'

112. mihi, mi, min 'tongue'

54. itsaso 'sea'

97. argal 'thin'

*-n- extension may be PIE or later; this does not affect the validity of the etymology (just as it does not affect the validity of parallel etymologies of Tocharian, Germanic and Iranian reflexes)

M-A 299 reconstructs PIE *makróś as “a […] word for ‘thin’”, with reflexes such as Latin macer ‘lean, meager, thin’, Old Norse magr ‘thin’, Hittite maklant- ‘thin’; V 356-357 reconstructs PIE *mh₂k-ro- ‘long’ for Latin macer ‘thin, lean’; so the meaning ‘thin’ is well-attested across IE, and it is immaterial whether we need to project this specific meaning back to PIE; a *-lo- diminutive
Evidence for Basque as an Indo-European Language:
A Reply to the Critics

may be an internal Proto-Basque formation, with an exact morphological and semantic parallel in Latin *macellus*, which remains valid and compelling even without projecting the exact semantics and morphology back to PIE. M-A 128 tentatively reconstructs PIE *[teNH₂G]*- ‘shallow water’ for, e.g., Latvian *tīgas* ‘deep spot in water’, Greek *tēnagos* ‘shoal, shallow water’, and tentatively connects Latin *stāgnun* ‘standing water, pool, swamp’. V 585 tentatively reconstructs PIE *steH₂G-* for Latin *stāgnun*, though, and proposes different cognates. M353-354 reconstructs PIE *st(e)H₂G-* ‘pool’ for Proto-Celtic *stagra* ‘river, stream’, and connects it to Latin *stāgnun*. Semantically, these proto-forms and reflexes are all fairly coherent.

Morphologically, if we combine M-A and M, we get a PIE *(s)teH₂(n)g-* (with “s-mobile” and optional *-n*- infix), which gets a *-rV*- extension in Celtic. Though we probably cannot project the *-rV*- extension back to PIE, we can reconstruct *st(e)H₂ng-* as the shape of the PIE root which underlies Basque *zingira*, with a later *- (V)rV-* extension, with clear parallels in Celtic.

can be more precisely explained as deriving from PIE *m(e)ud-* with a (later) affixation which parallels Proto-Celtic *mussako-* (M 281)
To summarize, I believe that the above etymologies remain valid, with the important proviso that the specific combinations of PIE morphemes in them probably cannot be projected back to common PIE, but are more likely to be Proto-Basque innovations.

**Personal pronouns**

Prósper, Kassian, Gorrochategui and Lakarra (section 5.3) variously criticize my etymologies of personal pronouns. I agree that some etymologies I had previously proposed for 69 ni ‘I’, 117 hi ‘thou’, 45 gu ‘we’ and 84 zu ‘you’ were fairly convoluted (and this was especially true in an earlier draft of my article that Gorrochategui and Lakarra may be referring to in their comments). In particular, deriving a first person singular pronoun from a first person plural one is a quite uncommon process, and all the entanglement and analogical leveling I posited for the personal pronoun system looked fairly contrived, though not totally implausible. I think a much simpler explanation is possible, at least for first and second person singular pronouns. K 111-113 reconstructs accusative and oblique first person singular pronoun PIE *h₁mne, which was extended to nominative usage in Lydian, Lycian and Luwian amu ‘I’; K also cites Tocharian parallels¹⁹: see K 111-113 for all details. To explain Basque ni ‘I’ and hi ‘thou’, all we have to do is start from PIE accusatives *h₁mne,²⁰ *te, which would regularly yield *ne, *he (by sound laws #2 and #6-7 respectively), whence ni, hi (for e/i vocalism, compare widespread Celtic forms mi, my, me and ti, ty, te for these pronouns).

As for 84 zu ‘you’, it is solidly attested as second person singular unmarked pronoun (vs. hi ‘intimate’: T 96), but there is evidence that it was originally used as plural (see T 196 and T-E 379). However, zu would be the regular phonetic outcome of PIE *tuh₃, while for second person plural we have PIE *swe (M-A 417) or *us- (K 111), which may have yielded Proto-Celtic *swis ‘you (pl.)’ (M

¹⁹ AD 265 also reconstructs *mne for Tocharian B ῆnas ‘I, me’.
²⁰ This also reduces by one unit (from 11 to 10) the amount of etymologies with a loose semantic match between Basque and PIE in Appendix 5, as Basque ni ‘I’ is now derived from ‘me’, not ‘we’.

The Journal of Indo-European Studies
Evidence for Basque as an Indo-European Language: A Reply to the Critics

Volume 41, Number 1 & 2, Spring/Summer 2013

365) (> Old Irish si etc.) and Hittite šum- ‘you (pl.)’ (K 115-116), though the details are unclear. A possibility worth exploring is that zu might be a conflation of singular *tu- and plural *su-, which would have become phonetically identical (by sound laws #4 and #18), generating the innovation zu-ek for the plural (T 196). As for the plausibility of a co-existence of hi < *te- and zu < *tu-, compare, e.g., the fact that in several parts of Northern Italy two 2nd person singular subject pronouns synchronically coexist side by side: a colloquial, intimate te (< Latin accusative tē), and an unmarked tu (< Latin nominative tū).

Gorrochategui and Lakarra say I derive 45 gu ‘we’ < PIE *wei “through completely gratuitous phonetic changes”. Actually, *wV- > g- is perfectly regular (sound law set #20 has 10+ examples), while *ei > *o (whence –u) may either be explained as vowel rounding conditioned by *w- (as I posited) or by analogy with other PIE *-o- endings in plural personal pronouns (such as *no- ‘we (obl.)’, *wo- ‘you (obl.)’; a similar *e > *o analogy-driven change can be found, e.g., in Tocharian B we- ‘we’ < *wo- (by analogy) < *wei- (AD 266).

We could thus summarize the possible evolution of Basque 1st and 2nd person pronouns as follows:

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person</td>
<td>ni &lt; *h1mne</td>
<td>gu &lt; *wo- &lt; *we-</td>
</tr>
<tr>
<td>2nd person</td>
<td>hi &lt; *te-</td>
<td>zu &lt; *su &lt; *swe / *us</td>
</tr>
<tr>
<td></td>
<td>zu &lt; *tsu &lt; *tu- (?)</td>
<td></td>
</tr>
</tbody>
</table>

2 Responses to individual comments

In this section, I address specific comments by individual commentators (comments shared by several commentators have been addressed in previous section already).

Gorrochategui and Lakarra

Internal etymologies with dubious or unspecified semantics

Gorrochategui and Lakarra seem to make their best efforts to provide internal etymologies, though sometimes “à tort et à travers” (to quote their own words); in several
instances, this leads to dubious semantics, in particular in the following cases, where I think readers may easily find the semantics in my etymologies more convincing than the Basque-internal alternatives suggested by the commentators:

<table>
<thead>
<tr>
<th>Basque term and meaning</th>
<th>Semantics of my etymology</th>
<th>Semantics of Gorrochategui and Lakarra’s etymology</th>
</tr>
</thead>
<tbody>
<tr>
<td>18. buztan ‘tail’</td>
<td>‘tail’ (in Celtic parallels)</td>
<td>‘above’ + ‘blow, fart’</td>
</tr>
<tr>
<td>39. gibel ‘liver’</td>
<td>‘liver’</td>
<td>‘black (?) matter’, supposedly from gi- ‘prefix of matter’ + -bel, a root with unspecified semantics (possibly the same as beltz ‘black’?), recurring is such terms as ‘purple’, ‘fallen leaf’, ‘slate’, ‘stomach’</td>
</tr>
<tr>
<td>53. ilun ‘dark, darkness, night’</td>
<td>‘unlit’</td>
<td>‘space or time of the dead’</td>
</tr>
<tr>
<td>89. ile, ule ‘hair, fur, wool’</td>
<td>‘body hair, wool’</td>
<td>‘grinder, carver’, supposedly from –le suffix of ‘agent’ + a reconstructed root *non ‘to grind, carve out’</td>
</tr>
<tr>
<td>103. ero ‘kill’</td>
<td>‘die’</td>
<td>‘to cause to grind or beat’, supposedly a causative of *e-non, probably the same root as above, here glossed as ‘to grind, beat’, though ‘to close’, from a reconstructed root *her ‘to close’, supposedly seen also in eri ‘finger’</td>
</tr>
<tr>
<td>118. hiru, hirur ‘three’</td>
<td>‘three’</td>
<td>‘to close’, from a reconstructed root *her ‘to close’, supposedly seen also in hirur ‘three’</td>
</tr>
<tr>
<td>134. eri ‘finger’</td>
<td>‘finger’</td>
<td>‘to close’, from a reconstructed root *her ‘to close’, supposedly seen also in hirur ‘three’</td>
</tr>
<tr>
<td>136. gordin ‘raw’</td>
<td>‘gore, raw meat’</td>
<td>‘to become hard’</td>
</tr>
<tr>
<td>156. her(t)ze, er(t)ze, este ‘intestine'</td>
<td>‘womb, stomach; belly, abomasum; entrails’</td>
<td>‘closed / closing thing?’, supposedly from *her ‘to close’ + -(z)e ‘nominalizer’</td>
</tr>
</tbody>
</table>
I find it particularly puzzling that a (reconstructed) root *her ‘to close’ should be used to try and justify internal etymologies for such diverse terms as ‘intestine’, ‘finger’ and ‘three’.

In other cases, Gorrochategui and Lakarra don’t even provide semantic clues to their reconstructions, so I honestly would not know what to do with cases like the following:

<table>
<thead>
<tr>
<th>Basque term and meaning</th>
<th>Semantics of my etymology</th>
<th>Semantics of Gorrochategui and Lakarra’s etymology</th>
</tr>
</thead>
<tbody>
<tr>
<td>19. ebaki ‘to cut’</td>
<td>‘to divide’</td>
<td>a reconstructed root *ban with unspecified meaning and origin</td>
</tr>
<tr>
<td>23. eman ‘give’</td>
<td>‘in’ + ‘hand’</td>
<td>a reconstructed root *bo(n) with unspecified meaning and origin</td>
</tr>
<tr>
<td>48. ibai ‘river’</td>
<td>‘river’</td>
<td>‘water’ + a reconstructed verb *e-ban-i with unspecified meaning and origin</td>
</tr>
<tr>
<td>65. lur ‘earth, land, soil, dirt’</td>
<td>‘dirt’</td>
<td>a reconstructed Pre-Basque initial *d-, which would yield an original *dur, with unspecified meaning and origin</td>
</tr>
<tr>
<td>104. handi ‘big’</td>
<td>‘measure’ (in PIE) &gt;</td>
<td>unspecified semantics; supposedly from a reconstructed root *han-, with unspecified meaning, and a suffix –di ‘adjectival suffix’ (to be found under comments to 67. mendi)</td>
</tr>
<tr>
<td></td>
<td>‘quantity, size, greatness, magnitude’ in Celtic</td>
<td></td>
</tr>
<tr>
<td>108. labur ‘short’</td>
<td>‘short’</td>
<td>unspecified semantics; supposedly from a “subroot element” la-, with unspecified meaning (occurring, according to the commentators, also in labar ‘edge of precipice’, with no apparent semantic connection); nothing is said about the origin or meaning of –bur</td>
</tr>
</tbody>
</table>
Chronology of sound laws

Gorrochategui and Lakarra's criticisms of my sound laws are particularly weak, as they are often skewed by a major misunderstanding of the chronology of sound laws. Sound laws typically start and cease to operate during a specific time frame, so it is perfectly possible that some sound law which I posit for Proto-Basque may have ceased to operate after Pre-Basque. By definition, Pre-Basque is the earliest possible internal reconstruction of an ancestor of Basque, based on dialect variation, known loans, and synchronically observable morpho-phonological alternations. In spite of this, Gorrochategui and Lakarra often use Pre-Basque-to-Basque sound laws to criticize earlier sound laws, which may well have ceased to operate before Pre-Basque. A clear example of their chronological confusion concerns the loss of PIE *m (my sound law set #2). It is well known that Basque words do have /m/, with various possible origins, listed in section 7.1 of their comments. However, the fact that an /m/ phoneme was re-instated after Pre-Basque, as a result of loans and internal developments, does not contradict the fact that Pre-Basque had no */m/ phoneme, and that a sound law that deletes PIE *m may well have operated before Pre-Basque, but ceased to operate (i.e., ceased to "delete" secondary or borrowed m’s) after Pre-Basque. There are hundreds of well-established sound laws in IE languages alone, that are known to have operated only during a certain time frame; e.g., PIE *s > h in ancient Greek (under some constraints), while secondary /s/ did not undergo the same process; however, no one (to my knowledge) criticized, e.g., the sound law *sV- > hV- for Greek on the ground that there are words in Greek that begin with sV- (as we know that such instances of sV- have a secondary origin, and arose after *sV- > hV- had ceased to operate).

Similarly, in section 7.2 they criticize my sound law set #19 (Environment-conditioned palatalization of /s/) on the ground that there exist, synchronically, examples where <z> is in contact with a front vowel, and where <s> is in contact with a non-front vowel. This fact might not contradict my sound law, if such examples turned out to
derive from different vocalic environments before Pre-Basque. As a parallel, it is known that Latin /ka-/ yielded /tʃa/ in Old French, but no one believes that this obvious sound law is falsified by the existence, in modern French, of such words as car (with a secondary c-, < Latin quà ré). Additionally, Gorrochategui and Lakarra partially misunderstand or misrepresent my sound law: I posit that *s was palatalized by a following *e or by a preceding *e or *i, so adducing several examples of <z> followed by <i> does not contradict my sound law anyway.

In section 7.3 (and in the appendix, under 33 ezur ‘bone’ < *enazur) they criticize my sound law set #18 (which occurred before Pre-Basque) because there isn’t “any known change” of this kind in the Pre-Basque to Basque period that their internal reconstructions can address.

Again, in the appendix, under etymology 75 (sehi, sein ‘boy, child’ < *seuh-nV-), they contend that “there is no proof of such early *-eu- > -e-; in fact, I posit this sound law (in set #22) for the transition from Proto-Basque to Pre-Basque, so the fact that no trace of such change can be found later on (from Pre-Basque to Basque) is not in contradiction.

In the second part of section 7.5, they say I do not adequately explain the Anlaut alternation d- > l-, which happens after Pre-Basque (e.g., in Latin loans); as I explained in “Introduction to sound laws” in my article, I only reconstruct up to Pre-Basque – I’m not interested in reviewing, discussing, challenging or amending any post-Pre-Basque sound laws, which have been sufficiently well established by Michelena, Trask, etc.

In short: it is obviously unwarranted to say that a sound law I posited for the PIE-to-Pre-Basque period is “impossible” or “groundless” just because it is not observable in the PreBasque-to-Basque period.

Additional cases of this critical misunderstanding may also be found with regard to my comments on specific etymologies below (e.g., 26. erre ‘burn, bake, grill, roast’, 74. hotz ‘cold’, 90. larru ‘skin’, 95. orri ‘leaf’), as well as next sub-section.
Alleged “implausibility” of some sound laws

Some of the sound laws I posit, such as #20 *w > g and #4 *tu > *tsu are dismissed with very weak, or downright wrong, criticisms. The change *w > g is simply dismissed because it “lacks any historical parallels”. This may mean one of two things: either they mean it should be dismissed because there are no historical parallels within Basque (i.e., after Pre-Basque) – and then we’d have, again a case of misconstrued chronology (see previous sub-section), i.e., it is perfectly possible that *w > g operated only before Pre-Basque, and so there is no trace of this after Pre-Basque; or else, the “lack of historical parallels” may point to a more general lack of *w > g in known IE languages, but this, too, would be obviously wrong, as, e.g., this shift is well known for Armenian.

*tu > tsu (#4) is dismissed as a “groundless change” or a “very unnatural phonetic change”. If by “groundless” they mean there is no trace of it after Pre-Basque, they may well be right, but then we’d have yet another case of misunderstood chronology, as discussed in the previous sub-section (i.e., again, such change may well have taken place before Pre-Basque but not after). If, on the other hand, their criticism is based on the lack of similar phenomena in other languages, I beg to differ; general fricativization of /t/ before vowels is obviously well-attested (e.g., zweite Lautverschiebung in German), but a selective fricativazion or asibilization of /t/ before /u/ is also visible in, e.g., PIE *tu, *twV > Greek /su, sV/ and in Japanese (where [ʦu] is the phonetic realization of /t/ + /u/).

Pre-Basque */h/ phoneme

Gorrochategui and Lakarra insist that */h/ was phonemic in Pre-Basque, contra Trask (T, section 3.11, p. 157-163); see my sound law set #17. However, the only specific etymologies they refute on this ground are four (see sections 7.4 and Appendix of their comments). Three of them, namely 48. ibai ‘river’, 89. ile, ule ‘hair, fur, wool’ and 134. eri ‘finger’, have already been dealt with in the previous sub-section. The fourth one, 121. eri ‘sick, ill’ is traced back to Gascon heri; however, as far as I could find,
Gascon herì is an infinitive (‘to hit, wound’ < Latin ferìre), whereas a participle such as herìt ‘hit, wounded’ would seem more appropriate as an origin for Basque erì ‘sick, ill’ (the semantics would be fairly plausible, though not straightforward); in any case, T 189 reconstructs a Pre-Basque form *eli (with *-l-, not -r-), which further weakens the Gascon etymology.

Palatalization of *s (sound law set #19)

Gorrochategui and Lakarra criticize this sound law for two reasons: first, because of alleged internal contradictions in my article; second, because of alleged counter-examples. The latter are unwarranted, as I already explained in sub-section “Chronology of sound laws” above, so let’s take a closer look at my alleged internal contradictions.

As for 138 izaba ‘aunt’ and 140 osaba ‘uncle’, I accept their criticisms: see sub-section “Kinship terms” below.

33 ezur is presented as a contradiction, but it’s not: it derives from Pre-Basque *enazur (T 178, 283, 394), and if *enetsur had already evolved into *enatsur before sound law set #19 (as I posit), unpalatalized <z> is perfectly regular.

A similar phenomenon may explain 87 hatz < *peds, i.e., again, *e > a before sound law #19.

As for 81 zazpi ‘seven’, in my article I simply assume that it was *septsu when sound law set #19 was operating (so *e was not in contact with *ts and could not palatalize it), and underwent metathesis only later.

For 47 hortz ‘tooth’ I reconstruct *odints, where *n prevents *i from palatalizing -*ts; in other words, I assume *n was lost after sound laws #19 had ceased to operate.

On the other hand, Gorrochategui and Lakarra are right in pointing out that 40 gizon ‘man’, 114 ezne ‘milk’ and 55 itz ‘word’ seem to contradict sound law #19. In fact, these instances require a chronological re-arrangement of two of my sound laws:

- in sound law set #17 I posit *∅ > z; this is incorrect, for two reasons: first, no phoneme *z should exist yet (it should emerge only after sound law set #19), so I
should have written \( *\theta > *s \); second, if \( *\theta > *s \) had taken place before sound law set #19, we should have expected Pre-Basque **gison and **esene instead of *gizon (T-E 387) and *ezene (T 316); I therefore should amend sound law set #17, in that \( *\theta \) must have evolved into /z/ after sound law set #19, not before it.

- similarly, I included *-Vkt- > *-Vits- in sound law set #15; however, if this change had taken place before sound law set #19, we should have expected **its instead of itz, as Gorrochategui and Lakarra correctly point out; by the same token, we should also expect **bederatsu and **gaits; so this sound law should be amended to *-Vkt- > *-Vitz- and should be placed after sound law set #19.

Gorrochategui and Lakarra are also right that my etymology for 123 haserre 'angry; anger' does not account for apical /s/, since my sound laws do not include a palatalization of *s by a preceding *a; so etymology #123 should, as a minimum, be demoted from “likely” to “tentative”; the segmentation has-erre (with erre ‘burn’) can be found in T-E 113 as well, but there has- is derived from hats ‘breath’: if my etymology is now weakened on phonological grounds, Trask’s internal etymology may be a better alternative, ultimately leading to an outright rejection (instead of simple demotion to “tentative” status) of my own etymology.

As for 78 sudur ‘nose’, initial *e- in my reconstruction is not “gratuitous” as they’d have it: it is motivated as a prefix (*h1en-: see etymology #78 for details).

On the other hand, Gorrochategui and Lakarra are right that 193 suge < *es-uge is just a guess, based on the fact that s-, in my system of sound laws, may only derive from an original *es-, so I would tentatively reconstruct a Proto-Basque *esuge on purely phonological grounds; I then observed that, if a segmentation *es-uge was attempted, then *-uge might be a regular reflex of PIE *h2ogwh-, but since I’m anyway unable to propose an etymology for the hypothetical initial segment *es-, this is just a speculation, so I left 193 suge in the section for basic lexical items for which I could find no IE etymology.
Other sound laws

I agree that my treatment of PIE syllabic sonorants *m, *n, *l, *r (section 7.5; sound law sets #2, 12, 13) is very sketchy and provisional, since in many cases it is supported by just one or two etymologies, so this area will definitely need refinement when (or if) new etymologies and conditioning factors are discovered. Since very few etymologies (about a dozen) involve these sounds, a revision of these sound laws might entail the rejection of very few etymologies, if any, so the impact of my still imprecise treatment of syllabic sonorants (due to currently insufficient evidence) should not make a major dent in my etymologies anyway.

As for the non-coherent treatment of PIE initial non-aspirated voiced plosives (sound law set #8) and non-initial plosives (sound law set #10), which they criticize in section 7.5 of their comments, I agree that some specific sound laws in these two sets (in particular, set #8) are based on very limited evidence (one or two etymologies) and might thus need to be changed when additional evidence becomes available; again, since these sound laws affect very few etymologies, any future changes to them cannot invalidate a significant amount of my etymologies anyway. Also see my response to Prósper’s criticisms of sound law set #7, below, where I put forth a more coherent hypothesis for the evolution of these sounds.

As for criticism to sound laws affecting vowels (#2, 16, 22), I addressed vowel reconstruction issues in sub-section “Is a classification invalid if parts of its sound laws are still sketchy?” above (which see), and I address specific etymologies criticized here in other parts of my comments.

Numerals

We have a complex situation here (see section 5.1 of Gorrochategui and Lakarra’s comments). I think an IE origin for 15 bi ‘two’ < *dwih₁, 118 hiru, hirur ‘three’ < *treies, *tisores, 76 sei ‘six’ < *seks and 81 zazpi ‘seven’ < *septm is extremely likely (for criticisms to sound laws underlying these etymologies, see above). In particular:

• the alternation hiru / hirur and lau / laur may be explained etymologically for ‘three’ (masc./fem.),
and as an analogical remodeling on hirur for laur

- I do not assume sei to be “an ancient loanword”, but rather a direct reflex of PIE *seks, via regular sound laws in my system
- zazpi with metathesis < *septsu < *septm is probably the most obvious etymology among numerals

As for 91 lau ‘four’, the parallel with Old Welsh lau ‘palm, hand’ < PIE *ph₂meh₂ is surely striking, and there are other instances of *m > u (see sound law set #2). However, if we apply my sound laws in the order I posited, we should get something like *elaua > *elaga, as Gorrochategui and Lakarra correctly point out; therefore, loss of –a must pre-date *VuV > *Vwv > VgV (sound law set #20), i.e., this part of sound law set #22 (loss of word-final vowels in words with more than two syllables) must actually be earlier than sound law set #20, and therefore also earlier than loss of *e- in pre-tonic position; this way we would have *elaua > *elau > lau. Incidentally, Basque would not be the only IE language that apparently lost (or never had) a reflex of PIE *kʷetwor-: Anatolian languages did not have it either (M-A 311).

I agree that 149 bortz, bost < *penkʷ-ro-st- looks contrived (that’s why it’s a “tentative” etymology), so a loan from some reflex of Proto-Celtic *bostā ‘palm, fist’ is a more likely explanation, though not all details are clear to me (see etymology #149 for more information).

As for 10 bederatzi ‘nine’, it is reconstructed as Pre-Basque *bederatzu by T; residual octal systems often have recourse to compounds to generate a word for ‘nine’ on their transition to decimal systems; it is true, as Gorrochategui and Lakarra point out, that this usually results in a subtraction from 10 (as, e.g., in Finnish), but usage of ‘eight’ as a basis for ‘nine’ is also attested (e.g., Ossetic farast ‘nine’ < 'beyond eight’). Interestingly, T-E 128-129 lists bederatzi under bat ‘one’, so a segmentation bed-eratzi, with bed- ‘one’ seems quite defensible; In T 273, Trask wrote “Both zortzi and bederatzi seem to share a final element –tzi, of unknown function” and “the word for ‘nine’ was originally *bederatzu”. PIE *h₃ektoh₁ would regularly yield Proto-Basque *oitsu, so my etymology for bederatzi does not look so far-fetched. For intervocalic *d
> r, see T 136 and P17 in T-E 28. So, based on this information from Trask, we can safely go back to a segmented proto-form *bed-era-tzu, and possibly to *bed-eda-tzu, with *bed- safely < bat / bede- ‘one’: we are just a few simple steps away from *bede eta oitsu ‘one and eight’. However, there seems to be no other trace of *oitsu in Basque, unless someone comes up with a convincing etymology for *zor-21 in 200 zortzi ‘eight’, which would support a segmentation zor-tzi and a possible origin of –tzi < *oitsu. Therefore, it seems more prudent to move bederatzi to the “tentative etymology” section.

145 bat, bede ‘one’ already belongs to the “tentative etymology” section, and 111 hamar ‘ten’ should probably be moved there, too.

To summarize, I would say Basque numerals 2, 3, 6, 7 are safely IE; numerals 1, 4, 5, 9, 10 might ultimately be, in various ways, including loans and internal innovations; 8 still does not fit in, unless it can be successfully explained as zor+tzi (which I’m unable to do).

Kinship terms

In section 5.2 and in the appendix, Gorrochategui and Lakarra criticize some kinship terms for which I provided an etymology.

Kinship suffix 8. -ba is widely attested in Basque, with a possible variant -be (see T 260-270, T-E 123 for details). The fact that the distribution of related PIE *-kwa (M 260) ’kinship suffix’ is sparse22 does not make the comparison invalid – especially if we consider that one IE sub-family where it happens to be attested is neighboring Celtic.

127 alaba ‘daughter’: though the counter-examples offered by Gorrochategui and Lakarra against a derivation of alaba < alu offer no specific example of –lu > -l- in compounds, I accept that this etymology may be

---

21 Assimilation of initials in contiguous numerals is a fairly common phenomenon; cf. e.g., Slavic words for ‘nine’, which begin with d- (instead of **n-), by analogy with words for ‘ten’; e.g., OCS desëti ‘nine’, desëti ‘ten’; we might thus conjecture that a proto-sequence *sei, *septsu, *oitsu became *sei, *septsu, *soitsu; however, this would only explain the initial z-, but not the -r- in zortzi ‘eight’.

22 This criticism is also found in Koch’s comments.
implausible based on general Basque morpho-phonological rules.

131 neba ‘brother of a woman’ seems to be contradicted by eme ‘female’ (i.e., if they both derive from Occ. hemne, as I surmised, why should the outcome of – mn- be different in the two words?), so my etymology should be rejected – unless someone comes up with a convincing conditioning factor that can explain the different outcomes.

I also accept that the etymologies for 138 izaba ‘aunt’ and 140 osaba ‘uncle’ are weak and should rather belong under the “tentative etymologies” section, at least until sound laws are refined enough to either promote them to a higher likelihood, or definitely exclude them.

For 75 sehi, sein ‘boy, child’ see my comments above (in sub-section “Chronology of sound laws”).

For 165 ume ‘child’, -kume ‘offspring’ (in compounds), Gorrochategui and Lakarra actually provide additional evidence for a derivation from umme < *unbe, for a segmentation *un-be and even for k-/0- alternation (also see T 136 and P14 in T-E 27), so I’d say the evidence they adduce supports my etymology, rather than invalidates it.

They criticize 4 anaia ‘brother’ because it allegedly does not drop the final vowel (contra sound law set #22); however, I derive this term from *h2ed-gnh1-yV-; the vowel that should drop is this unspecified *-V; given that a form anai is amply attested (in B, G, L, LN, Z) and that final –a may well be a petrified article, I’d say the final vowel may well have regularly dropped. Anyway, since there are half a dozen other proto-phonemes in this reconstruction that are regularly accounted for, I would not say that a marginal doubt on the reflex of a single, final vowel can invalidate, per se, this whole etymology.

To summarize, I accept that etymologies 127, 131, 138 and 140 are weak and might ultimately have to be rejected. However, in my article, 127 and 131 are internal etymologies, while 138 and 140 are supposed to be loans, so even if these 4 etymologies were to be discarded, this would not decrease the number of basic terms derived directly from PIE (i.e., those in the “Likely etymologies”
section).

Various words meaning ‘up, top, high’

In the appendix, Gorrochategui and Lakarra criticize my etymologies for 35 gain ‘(on) top’, 37 garai ‘high place’, 41 goi ‘high (place)’, 42 goiz ‘morning’, 43 gora ‘up’, 49 igan, igon ‘ascend’, which I derive from PIE *uper- (M 398) – or *h₄uper- (M-A 292). They object that -r- could not disappear and that g- could not be “added”. The latter objection is yet another case of their misunderstanding of the chronology of my sound laws: see my comments to their criticism of sound law #20 *w- > g-above. As for the loss of -r-, this is a sporadic phenomenon, according to T-E 29 (P25 “Sporadic /r/ loss”). Anyway, it is probably better to review this set of etymologies and re-group them differently, without positing /r/ loss:

- 37 garai and possibly 43 gora\(^{23}\) < PIE *(h₄)uper-*, with Brittonic comparanda such as MW gor, gwar (M 398)
- “r-less” 35 gain, 41 goi, 42 goiz and 49 igan, igon < PIE *h₄upó ‘up’ (M-A 289, 292) > *uo- > *wo- > go-, or from the underlying verbal root *h₄up- (ibidem) also found, e.g., in Hittite úpzi ‘the sun rises’ (a semantically very interesting parallel for 42 goiz ‘morning’, by the way).

Other etymologies

12. ber- ‘self’: Gorrochategui and Lakarra correctly point out that we should reconstruct *ber- for Pre-Basque, instead of my *bede; in fact, this reconstruction can also be found in T-E 388 and 392, so it is not likely that -r- < *-de. A *-IV- extension (with subsequent\(^{24}\) regular rhotacism) is more likely, as in Latin sólus (IEW 882-884).

26. erre ‘burn, bake, grill, roast’: in my reconstruction, e- is part of the root, and adding the verbal prefix *e- to it would regularly result in *e-er- > *err-; as for the alleged “impossibility” of *-tr- > -rr- before Pre-Basque, we have yet another example of Gorrochategui and Lakarra’s recurrent

---

\(^{23}\)If –ra in gora is an allative suffix (as per T-E 207), then gora should belong with the other group.

\(^{24}\)We should actually reconstruct *belV for Pre-Basque.
confusion about the chronology of sound laws, as discussed above.

40. **gizon** ‘man’: the fact that an internal etymology (with similar semantics) is available does not automatically invalidate my own one. By the way, *gi- ‘meat’ is labeled as “hypothetical” in T-E 205.

44. **gorotz** ‘dung’: I’m well aware of an alternative etymology < Latin *crocea ‘saffron-coloured’ (see T 261, T 284 – with “probably” – and T-E 209), but its semantics seems convoluted and over-stretched: ‘saffron-coloured’ > ‘yellowish garments and coverings’ > ‘covering of dung and straw placed over a haystack’ > ‘dung’; I think my semantics ‘dirt’ > ‘dung’ is more straightforward.

46. **hartu** ‘to take’: it is true that I do not account for *-rr- > -r- from Proto-Basque to Basque, so this etymology should be demoted to “tentative”, because of incomplete sound correspondences.

63. **joan** ‘to go’: the commentators confirm that there must have been a *-n- in the root, as I, too, wrote, following T-E 387; they propose, of course, an alternative internal etymology (from a reconstructed root *da-, with unspecified meaning and origin, as is too often the case in their comments). However, a derivation from PIE root *menH-/monH- is absolutely straightforward, being based on few and regular sound laws (incl. #2 *m- > 0- with 18 examples) and having a perfect parallel in Celtic (Proto-Celtic: *mon-i- ‘to go’), so I really see no good reason to change or reject my etymology.

67. **mendi** ‘mountain’ is obviously a noun, and mountains are, for sure, inanimate, so I do not understand why the commentators take –di as an adjectival ending (also found in “names of domesticated animals”). Since **mendi** is clearly reconstructed as Pre-Basque **bendi** in T-E 287, 388 and 392 (with no segmentation whatsoever), a comparison with Proto-Celtic **bando- ‘peak, top’ (M 54) > Old Irish **benn ‘peak, mountain’, MBret. **ban ‘peak’, etc., is the simplest and most obvious explanation for this term.

74. **hotz** ‘cold’: the fact that final sibilants may become strong in Basque (i.e., may be a secondary phenomenon) is no proof against this specific instance of –tz being primary; in fact, *-tz is reconstructed for Pre-
Basque in T-E 386 and 394.

90. larru ‘skin’: yet another case of the commentators’ confusion on the chronology of sound laws: the fact that e- was not lost after Pre-Basque does not invalidate the theory that such loss may have operated earlier: see sub-section “Chronology of sound laws” above.

94. alu ‘vulva, vagina’ < Latin alvus ‘stomach, uterus’ (also ‘belly’) is a much more compelling and straightforward etymology than mine (in fact, in annex 5 in my article I classified this among “Etymologies with a loose semantic match between Basque and PIE”). I believe my own etymology should be rejected, and replaced with this loan.

95. orri ‘leaf’: I place sound law *-ly- > -rr- before Pre-Basque (in set #23), so again the treatment that *-ly- might have undergone after Pre-Basque is irrelevant: see sub-section “Chronology of sound laws” above.

100. eke, ke, kei ‘smoke’ is reconstructed as Pre-Basque *eke(e) in T-E 248-249, 389 and 392.

109. uzki ‘anus’: a connection with 18. buztan ‘tail’ is not out of question, though I do not understand what the precise phonetic evolution should be (in particular, why was b- lost in one case but not the other?). And the segmentation involved here (uzki < *buz-gi and buztan < *buz- + *gan-) has been criticized on semantic grounds in the sub-section ‘Internal etymologies with dubious or unspecified semantics’ above. All in all, this seems a fairly weak internal etymology.

116 ugatz ‘udder’ regularly loses its final consonant because of sound law set #22, while in 52 ikuzi ‘to wash’, –i is a very common verbal morpheme attached to verb root -kuz-.

119. adar ‘horn’: A link with Old Irish adarc is discussed in T-E 76 and ultimately rejected. Adar can be derived for PIE with fully regular sound changes; as for rhotic dissimilation, see T 145 and, more specifically, P24 in T-E 29: “In the sequence /rVrr/, the first (tapped) rhotic is unstable, and commonly dissimilates to /l/, /d/ or zero”.

126. odol ‘blood’: the “Basque etymology” cited by the commentators is not an etymology; it is simply an arbitrary segmentation o-dol, with *dol ‘blood’, that gives us no
useful information and begs the question (where is *o- from? where is *dol from? why should odol be a reduplicated word?). I reconstruct *orol for Pre-Basque, with *-r- > -d- as per P17 (/d/ ~ /r/ fluctuation) (T-E 28).

137. inurri, zinaurri 'ant': the fact that z-/0- alternation has Basque-internal parallels strengthens my etymology, as it makes it unnecessary to posit a double loan from two different Celtic sources.

141. sartu 'to go in, enter, insert' and 142. sortu 'to be born' are rejected on phonetic grounds, though the semantic alternatives offered by the commentators are not particularly appealing (‘thicket’ and ‘body’, respectively). I acknowledge that the commentators surely have a much better command of post-Pre-Basque phonology than I have. However, this might result in the rejection of two loans, which would not weaken the set of core etymologies that derive Basque words directly from PIE.

152. euskara 'Basque language': I acknowledge that my etymologies are conjectural; that’s why I put them in the “Tentative” section. I’m fully aware of Irogoyen’s etymology (extensively discussed in T 320-322 and briefly in T-E 186 and 226). It is very compelling on semantic grounds, but partly doubtful because it relies on a poorly attested, defective verb -io- / -iño- (T-E 226), which, with some special pleading, can be derived from a reconstructed root *enau- or *-inau-, to which a participial extension -(t)si should be added to obtain *enau(t)si25. However, given the additional phonetic problems that the commentators find in my etymologies, I accept that mine should most probably be discarded in favor of Irogoyen’s.

154. haritz ‘oak’: my etymology had already been classified as “tentative” on phonetic grounds; the variant hareitz, cited by the commentators, makes is weaker still; so do the parallels that seem to point to a *-leitz- segment, which would imply a different segmentation (ha-reitz). I think this is enough to discard my tentative etymology.

155. hega(l), ega(l) 'wing': I posit an optional *-l- extension for Proto-Basque, not Basque: compare similar extensions, e.g., in 97. argal, 102. erle, 107. (h)il-. This

25 As far as I understand, *enausi is a reconstructed form, not an attested one (as the commentators seem to imply by citing it without asterisk).
Evidence for Basque as an Indo-European Language:

A Reply to the Critics

extension may have become unproductive later.

162. ortzi, ortze, ost- ‘sky, storm, thunder’ is reconstructed as Pre-Basque *ortzi (and *ortze is explicitly discarded) in T-E 316, where Trask declares it “OUO” (of unknown origin). The most frequent meanings of this word (see T-E 316-319) seem to be centered around ‘thunder, storm, lightning’, paralleling several IE reflexes < PIE *por- (see examples in my article); a dental extension is also found in Armenian orot ‘thunder’ < *porado-: I think, on balance, my etymology is actually compelling enough, on semantic, morphological and phonetic grounds, to be promoted to “likely” status.

Prósper

Unclear / underspecified comments

24 entzun ‘to hear’ is criticized because, according to Prósper, *h₂eus- ‘ear’ should “yield *aus- like everywhere else” (see sound law set #1). 230 behi, bei ‘cow’ and 231 edo, ero ‘or’ are criticized because *h₂eu > *eu- and *gʷeh₃u- > *beu- are “unjustified evolution[s]”. I honestly do not know what to do with these generic comments, especially as Prósper does not support them with adequate evidence. Anyway, they are obviously too weak to invalidate these etymologies, so I guess I can be excused if I do not worry too much about them. Anyway, see my general comments on vowel reconstruction, above.

Comments on specific sound laws

On sound law sets #7 and #17, Prósper writes “The very idea that word initial voiced aspirates collapse into a single h and then disappear is questionable and has no parallels”. The exact trajectory through which word-initial plosives disappeared is hard to reconstruct at this stage; an alternative trajectory (with the same ultimate outcome) might be a fricativization in step #7, followed by loss of fricatives in step #17, e.g., something like *p- > *φ- > 0-; *t- > *θ - > 0-; *k- > *x- > 0-; *bh- > *β- > 0-; *dh- > *δ- > 0-; *gh- > *y- > 0-. This alternative trajectory is probably more plausible on typological and areal grounds.
As for sound law set #13, see my comments on PIE syllabic sonorants above (in sub-section dedicated to Gorrochategui and Lakarra).

Prósper’s comments on other sound laws (sets #9, 10, 19, 22) are made within the context of specific lexical etymologies and are therefore addressed in the next sub-section.

Comments on specific etymologies

2 *ain(n)tzin, antzin* ‘before, in front of’ and 48 *ibai* ‘river’ are in mutual contradiction, according to Prósper, because they both begin with PIE *h₂e*- and “contain the same suffix”, so it’s not clear to her why we should get different initial vowels in Basque (i- vs. a-), as if the two Proto-Basque words were stressed differently, i.e *ántsun > antzin* vs. *abáni > ibai* by sound law set #22. I can reply as follows: firstly, the PIE suffix is actually reconstructed differently for PIE in M 23 *h₂ep-h₃on-* ‘river’ (with *o) vs. M 39 (tentative) *h₂ent-h₃en- (with *e). Secondly, it is my understanding that, though Proto-Celtic is a long-established sub-family of PIE, there is still no consensus on precisely how stress evolved within this family, so I might be forgiven for not having been able to develop a theory of stress in Proto-Basque. For similar comments on the evolution of vowels, see above. Having said that, the different outcome of *h₂e-* in *ibai* and *antzin* may be explained in several ways, including:

1. it is a contradiction, so one of the two etymologies is wrong (this is probably what Prósper implies);
2. stress was actually different in the two words in Proto-Basque (though we still have no model to explain why), so the evolution of *ántsun > antzin* vs. *abáni > ibai* is regular, by sound law set #22;
3. stress was identical in the two proto-words (*antsún- and *abáni*), but some other conditioning factor (e.g., the presence of *n in *antsún-?) was at work and generated a different outcome;
4. one of the two words is native (possibly *ibai?) while the other is a borrowing from Celtic (possibly *antzin*, whose variant *aintzin* reminds of the *sandu / sanđdu* alternation < *sanctu-, a loan of ultimate Latin origin).
I presently don’t have enough information to choose among these alternatives. Whichever of them is correct, however, at least one of the two etymologies can be confirmed, while the other is either discarded, or confirmed, or turned into a loan, so the worst possible effect of this comment by Prósper is to invalidate one of my “likely” etymologies (if at all).

11 beltz ‘black’ is traced back to Pre-Basque *bel-, which is perfectly compatible with the Aquitanian parallel cited by Prósper.

33 azur, ezur, hezur ‘bone’, Prósper criticizes my reconstruction *h₁en-h₁eh₁-tor, but does not explicitly explain why; however, she accepts M 115 reconstruction *h₁en-h₁oh₁-tro- for Middle Irish inathar, so she probably has an issue with the root grade (*h₁oh₁- instead of *h₁eh₁-?) and/or with the suffix (*-tro- instead of *-tor?). A suffix *-tor seems to be secured by, e.g., Greek étor; as for the grade, a strong grade would yield an intermediate proto-form *enotsur instead of *enetsur, which anyway has 5 out of 6 phonemes regularly corresponding to Pre-Basque *enazur, so I believe Prósper’s comment is not enough to invalidate my etymology (also see my comments of vowel reconstruction, above).

34 gaitz ‘bad’: Prósper writes that PIE *weh₂k-to- “probably does not exist as such”, presumably because she believes a different grade *wh₂k-to- should be posited (as in M 405). I have no problem with that: by sound law set #1, both *eh₂ and *h₂ > a, so the validity of my etymology is unaffected by this alternative reconstruction.

67 mendi ‘mountain’: according to Prósper “[t]hat a PIE root *bend- exists cannot be taken for granted, since it violates the combinatory rule forbidding two voiced stops”. Actually, this PIE root can be found, e.g., in M 54 and IEW 96f. I’m also glad to learn that there are several place names in –bendā that might support my etymology (I could form a better understanding if only Prósper had shared with us some more details about the possible meaning and origin of this element).

For 69 neska ‘girl’ I reconstructed PIE *gnh₁-eto-, which Prósper finds “unlikely”, presumably because she believes a different grade *gənh₁-eto- should be posited;
actually, zero-grade forms of this root are also attested before suffixes beginning with a vowel; anyway, also PIE *\text{\textgreek{gh}en}\text{-}\text{et}- + *\text{\textgreek{s}k}- can yield neska, too, by loss of pre-tonic *\text{\textgreek{e}}- (sound law set #22), if we reconstruct Proto-Basque *\text{\textgreek{y}en\text{\textgreek{e}}}tska.

71 On ‘good’: the conditional loss of Pre-Basque *b- before /o/ is a regular sound law governing the evolution of Pre-Basque into Basque, as explained, e.g., in T 133 and in T-E 27 (P9), i.e., Basque on derives regularly from Pre-Basque *\text{\textgreek{b}on}, so this etymology confirms the validity of PIE *\text{\textgreek{d}w}- > Pre-Basque *b- (sound law set #9).

75 sehi, sein ‘boy, child’: I’m glad to learn that this etymology might have further parallels in Celtic appellatives (such as SVNVA), too.

77 su ‘fire’ and 235 itsu ‘blind’: I’m not sure I understand the criticism here: both terms have regularly undergone palatalization of *s (as per sound law set #19) – not just su, as Prósper seems to imply	extsuperscript{26}: the first because of a preceding *\text{\textgreek{e}}- (which regularly disappeared by sound law set #22), the second because of a preceding i- (again, by sound law set #19). Actually, based on evidence from itsu, we should probably add “i- + -ts > -ts” to sound law set #19, and reconstruct an intermediate form *\text{\textgreek{i}nts}\text{su} which is not explicitly included in etymology #235: *\text{\textgreek{h}}\text{\textgreek{a}endh}\text{o}- > *\text{\textgreek{endh}}\text{o}- > *\text{\textgreek{endh}}\text{\textgreek{u}}- > *\text{\textgreek{ents}}\text{\textgreek{u}}- > *\text{\textgreek{ints}}\text{\textgreek{u}}- > *\text{\textgreek{its}}\text{\textgreek{u}}-, possibly with *\text{\textgreek{en}} > *\text{\textgreek{in}} in pretonic, closed syllables (vs. *\text{\textgreek{en}} > an in all other environments), thus enriching sound law set #16 as well.

78 sudur ‘nose’ < *\text{\textgreek{h}}\text{\textgreek{1}ens}h\text{\textgreek{3}od-o-ro-s} ‘having smell inside’: according to Prósper, my PIE reconstruction “is not only irregular, but requires a prefix that matches Greek directional *\text{\textgreek{e}ns}”. The alleged reason why this reconstruction should be “irregular” is not spelled out, so I cannot comment specifically, but one can note that:

- if Prósper has a problem with root grade, both root grades *\text{\textgreek{h}}\text{\textgreek{3}ed}- and *\text{\textgreek{h}}\text{\textgreek{3}od}- would initially > *\text{\textgreek{o}d}- by sound laws set #1, so a change in posited root grade makes no difference to this etymology;
- an *\text{\textgreek{o}}- extension of this root can be found, e.g., in Armenian hot (an o-stem, in fact);

	extsuperscript{26} “Palatalization […] should have equally happened in the second.
Evidence for Basque as an Indo-European Language: A Reply to the Critics

- extension *-ro- (for other roots) is common, and even comes as *-o-ro- (again, for other roots), e.g., in Germanic.

So it seems safe to reconstruct *-h₁e/od-o-ro-. As for directional *h₁en-s, it not exclusively Greek; it is also found, e.g., in Venetic es ‘unto’ (V 300). For these reasons, I find Prósper’s comment too generic and weak to warrant a rejection of this etymology.

81 zazpi ‘seven’: Prósper correctly points out that I posited an unconditional loss of PIE *-p- (in sound law set #10), which is obviously contradicted by zazpi < *septm; as all other instances of loss of PIE *-p- are between vowels, sound law set #10 should be amended as follows: PIE *V₁pV₂ > V₁V₂, i.e., *-p- was lost between vowels.

83 zorri ‘louse’ is criticized because it is presented as an ad hoc exception to the general law *sw- > b- (sound law set #9). However, there are at least six solid cases for *sw- > b-, so if you come across the comparison of Basque zorri ‘louse’ vs. Proto-Celtic *sworo- ‘louse’ (M 365) > Middle Irish sor, from PIE *sworó-s, what should you do? You could:

1. ignore it as a mere coincidence
2. accept that it cannot be a coincidence, and posit an exception to the general sound law (as I did)
3. accept that it cannot be a coincidence, and interpret the Basque word as a Celtic loan

In my article, I opted for option #2. However, option #3 might be preferable if examined in isolation (i.e., the reasoning would be: since zorri violates sound laws, it cannot be inherited, and must be a loan instead). Yet, there seems to be at least another instance of *swVr- > *sVrw- > *sVrr-, namely 28 esan ‘to say’ < Pre-Basque *e-serr-an < PIE *swer- ‘to say’, so it seems safer to posit a conditioned exception to sound law *sw- > b-.

110 egun ‘day, today’: Prósper writes that my segmentation of PIE *dei-wo- is wrong, and that *deywos does not mean ‘day’; I used the following sources: V 167 reconstructs PIE *dei-u-o- ‘god, divinity (god of the clear sky)’ > Latin deus ‘god, deity’; V 170 has PIE *di-eu- ‘God of the clear sky, Clear Sky’ > Latin dies ‘day, daytime’; M
101 has PIE *diy-ew- ‘day’ (IEW 184f.) > Proto-Celtic
*diyw(o)- ‘day’. M-A 301 also has PIE *dye(u)- ‘day’ < *dei-
‘shine’. Martirosyan reconstructs PIE *dyeus ‘heaven, day, Sky-God’ > Armenian tiw ‘day’. V 315 reconstructs PIE
*dieu- / *diu- as the proto-form underlying several IE
reflexes meaning ‘god’ or ‘day’. The list could go on, but I
think it is obvious that we are dealing with a PIE root
*d(e)i-, frequently found with *-u-(o)- extension, whose
reflexes mean ‘day’ in Celtic, Italic, Greek, Sanskrit, Armenian, etc.

133 belaun ‘knee’: the close resemblance with Middle
Welsh pen(g)lin, Middle Breton pennglin, Old Cornish
penclin ‘knee’ and Old Irish glúin (M 162) is undeniable
and can hardly be dismissed as coincidental, given the
semantic identity and the long sequence of phonemes
involved. Brittonic forms are traced back to Proto-Celtic
*kwenno-glúnos ‘knee-cap’ (M 162), whose first element
*kwenno- ‘head’ has no PIE etymology (M 177). The
proto-form *glúnos ‘knee’ < PIE *gonu is exclusively Celtic,
too. Therefore, *kwenno-glúnos is an exclusively Celtic
compound, with no PIE parallels. We have at least four
theoretical possibilities to explain Basque belaun in this
context:

1. it’s a mere coincidence, and it is not related to Celtic
words;
2. the whole Basque language derives from Proto-Celtic,
and belaun is inherited from Proto-Celtic *kwenno-
glúnos;
3. Basque cannot be derived from Proto-Celtic, but
there was some PIE compound (**kwendo-gnú-no-??)
that was lost in all other IE languages, except Celtic
and Basque;
4. the Basque word is a Celtic loan.

I would discard explanation 1 as all too unlikely, 2 as
unproven (see discussion on the relationship between
Basque and Celtic, above), and 3 as quite unlikely, so we’re
left with #4 as the most likely explanation.

148 bihotz ‘heart’ is a tentative etymology: Prósper is
right that one cannot reconstruct PIE *swi-; if my
etymology is correct, the PIE antecedent of Basque bi-
Evidence for Basque as an Indo-European Language: A Reply to the Critics

should be reconstructed as *s(e)we (V 549), so the etymology should be amended to *swe + * kords > Proto-Basque *be-hord-s (instead of *bi-hord-s), which might yield Basque bihotz anyway (see my introductory comments on vowel reconstruction, above).

167 ze-, zer- ‘interrogative stem’ (a tentative etymology) is criticized because “[t]here is no demonstrative PIE pronoun *se-“. Of course, I meant to refer to the PIE demonstrative pronoun commonly reconstructed as *so-, which has several shapes, incl. fem. *sih2- (M 337) or *seh2 (V 568) or *seh (M-A 418), and generates reflexes with different vocalism in various IE families: compare, e.g., Proto-Celtic *so- ‘this’ (M 350), *si- ‘she’ (M 335), *sindo- ‘this’ (M 336), *sondo- ‘that’ (M 351), Latin si ‘if’ < Proto-Italic *se-i (locative, V 561), etc. Since the following vowel may vary, this PIE pronominal stem is, more agnostically, reconstructed as *s-, e.g., in V 568 and M 335, while K752 explicitly reconstructs three different shapes for PIE, namely *si-, *so, *se-, based on Hittite vocalization. We thus may want to change the PIE reconstruction to *so/e/i- in my etymology #167, without affecting its (tentative) validity, however.

209 egotzi ‘throw’: Prósper writes that a PIE participle *yoh1-to- ‘thrown’ “is an invention”, because I allegedly “freely posit[] the grade [I] need[]”; actually, *-to- forms from strong-grade roots are permissible in IE (e.g., Greek mortós, etc.).

Kassian

All key comments by Kassian have already been addressed in previous sections.

Bengtson

Bengtson makes no specific comments about any of my etymologies (except that some are “non obvious”), and only criticizes 3 of my 23 sets of sound laws (#2, 7, 17) which have already been dealt with above, since other commentators have criticized them as well. So, basically, all Bengtson really tells us is, he has a better theory than mine, to which I reply as follows:
readers are invited to analyze my article and Bengtson’s works and decide which they find more convincing; I keep subscribing to Trask’s criticism of attempts to link Basque to North-Caucasian and the highly conjectural “Sino-Dené-Caucasian” phylum;

• interestingly, Bengtson’s theory is not cited as a valid classification by any other commentator, 27 except, possibly, Kassian: I think this is an indication that his theory does not have much currency among mainstream specialists.

Koch

Koch did not analyze my article in detail28, but believes a sizable amount of my etymologies might be acceptable. 29,30 However, Koch seems to share the “can’t be true” attitude of other commentators: rather than accepting the most logical conclusion that Basque is IE because most of its basic lexicon is IE, Koch prefers an alternative explanation: Basque was a non-IE language which “absorb[ed] layer upon layer of Indo-European loanwords”; therefore my valid IE etymologies “represent successive strata of pre-Roman loanwords”, i.e., “possible pre-Roman layers of Indo-European influence on Basque”, made possible by “a minimum of 3000 years of contact” with “waves or stages of Indo-European influence […], prior to the Roman Conquest”. As an alternative, but less likely, explanation, Koch suggests that “Forni could be onto something in his suggestion of Basque as a creole”.

In the context of Koch’s comments, we may be facing four major alternatives here:

27 (For example, Prósper writes “it is widely held that Basque is a linguistic isolate” and “other attempts to connect Basque with other language groups have foundered on the rocks”; Gorrochategui and Lakarra write “The Basque language is […] [a] language isolate[…]”; etc.).

28 “I have not challenged sound laws, etymologies, and percentages”, “it is inherently unlikely that all of the 167 etymologies and 121 sound laws are wholly without basis”.

29 “we say Forni has a case”.

30 I’m not analyzing Koch’s few comments on specific etymologies, as they are few (8% of my likely etymologies), and most of them have already been addressed in previous sections anyway.
Evidence for Basque as an Indo-European Language: A Reply to the Critics

1. my article is indefensible, because of various reasons put forth by other commentators or else, most etymologies in my article are defensible, and therefore
2. Basque is IE, or
3. Basque is a creole, mostly based on an IE lexicon, or
4. Basque is not IE, but I found additional layers of pre-Roman loans

I hope that my “counter-comments” are sufficient to dismiss option 1. Options 2 or, possibly, 3 are my interpretations of my findings. Option 4 is favored by Koch (though Koch thinks option 3 is less probable, but still possible).

Why my valid IE etymologies cannot all be pre-Roman borrowings

The main reason why I don’t believe option 4 is a likely explanation is that, if most of my etymologies are pre-Roman loans, then Basque would be a very unusual language: one with very, very little native (supposedly non-IE) inherited lexicon. This is unlikely for a theoretical reason and an empirical one.

The theoretical reason is the following: basic lexicon is resistant to borrowing, but loans may creep in it as well, so the probability that a given item in the basic lexicon is borrowed is low, but not null. Each single item may have different probabilities of being borrowed (e.g., verbs and pronouns are less likely to be borrowed than nouns), but, for the sake of simplicity, let’s assume that each item in the basic lexicon has the same probability of being borrowed: call such probability $p$, where $p$ is small but not null. If $p$ is the probability that an individual item is borrowed, the combined probability that $n$ items are ALL borrowed is $p^n$. Koch is telling us that perhaps more than half of my likely etymologies are valid, but represent loans, not inherited lexicon. How probable is this? If we consider only the “likely” etymologies that I regard as inherited

---

31 Probabilities are usually represented as numbers between 0 and 1, rather than as percentages: e.g., a 5% probability is expressed as 0.05, 20% as 0.2, 80% as 0.8, etc.
from PIE, i.e., \#1 to \#126\textsuperscript{32} and assume only half of them are valid, i.e., 63 of them, we have \(n=63\). Even if \(p=0.7\) (a very generous assumption indeed: it would mean that each basic term has a 70% probability of being borrowed “over the course of at least three millennia”), \(p^n\) would be

\[(0.7)^{63} = 0.00000000174251\]

i.e., the combined probability that ALL 63 basic terms are loans is less than one in five billion, or

\[1 / 0.00000000174251 = 5,738,831,575\]

You may think I deftly chose numbers that favor my thesis, so let’s see what happens if we are even more generous: suppose that the probability that a given basic term is borrowed during millennia of contact is higher still: 80%, or 0.8. Suppose that just 50 of my 126 etymologies are valid; the combined probability that they are all loans would anyway be one in 70 thousand, or 0.0014%:

\[(0.8)^{50} = 0.000014272476927 \rightarrow 1 / 0.000014272476927 = 70,065\]

The empirical reason why I dismiss option 4 as highly unlikely is much simpler to understand, but is a consequence of the theoretical discussion above: precisely because borrowing of a predominant portion of basic lexicon is so unlikely, there is no historically known language that underwent such a major overhaul of its basic lexicon. Even English (possibly with 50%+ terms of ultimately Latin origin), Japanese and Korean (with 60%+ terms borrowed from Chinese) preserve 80%+ (possibly even 90%+) of native, basic lexicon (i.e., borrowing is down to 10% for their basic lexicon).

As far as I know, the only historically attested cases of wholesale replacement of basic lexicon are creoles and instances of language substitution (i.e., cases when the original language spoken by a population is replaced altogether by the language of a dominant culture).

\textsuperscript{32}Thus excluding my ”internal etymologies” 127-131 and “likely loans” 132-142.
4. Summary of changes

In this section I summarize all changes to my original article introduced in these final comments as a consequence of criticism by commentators.

Changes to sound laws

#7: initial stops may have been lost via fricativization, rather than via *h
#10: non-initial *-p- was lost between vowels, instead of unconditionally
#15: *-Vkt-, *-Vkt- > -Vitz took place after #19
#17: *y > z also took place after #19
#19: *ts was palatalized not only by a preceding *e-, but also by a preceding *i-
#22: *-V > -0 (loss of final vowels) in words with more than two syllables actually took place before #21

Changes to etymologies

2 ai(n)tzin, antzin ‘before, in front of’ might be borrowed from Celtic, rather than inherited.
10 bederatzi ‘nine’ should be demoted to “tentative” if no other reflex of Proto-Basque *oitsu ‘eight’ is found.
12 ber- ‘self’ < *swe-IV- instead of *swe + *de.
33 azur, ezur, hezur ‘bone’ might be < *h₁en-h₁oh₁-tor instead of *h₁en-h₁eh₁-tor.
34 gaitz ‘bad’ probably < *wh₂k-to- instead of *weh₂k-to-.
35 gain ‘top; on top’ < *h₄up- instead of *h₄uper-.
41 goi ‘high place, height; high’ < *h₄up- instead of *h₄uper-.
42 goiz ‘morning’ < *h₄up- instead of *h₄uper-.
43 gora ‘up’ < *h₄up- instead of *h₄uper- if it must be segmented as go- + -ra ‘allative suffix’.
46 hartu ‘to take’ should be demoted to “tentative” because *rrr- > -r- is left unexplained.
49 igan, igo(n) ‘ascend’ < *h₄up- instead of *h₄uper-.
68 neska ‘girl’ possibly < *g₀nh₁-eto- instead of *g₀nh₁-eto-.
69 ni ‘I’ < *h₁mne ‘me’ instead of *no- ‘we’.
84 zu ‘you’ possibly a conflation of *tuh₄ and *swe.
94 alu ‘vulva, vagina’: a loan from Latin alvus, instead of a reflex of *bh₁el-n-.

₄₃ Please refer to previous sections for details.
111 **hamar** ‘ten’ should be demoted to “tentative”.
123 **hasserre** ‘angry, anger’ should be either demoted to “tentative” or rejected, because -s- is not accounted for, and an alternative internal etymology seems viable (**hats** ‘breath’ + **erre** ‘burn’).
127 **alaba** ‘daughter’ < **alu** + -**ba** should be either demoted to “tentative” or rejected, because it seems to be a morphophonologically irregular compound.
131 **neba** ‘brother of a woman’: my etymology should be rejected, as it is incompatible with **eme** ‘female’.
138 **izaba, ize(b)a, izo** ‘aunt’ should be either demoted to “tentative” or rejected.
140 **osaba, ose(b)a** ‘uncle’ should be either demoted to “tentative” or rejected.
141 **sartu** ‘to go in, enter, insert’ should be either demoted to “tentative” or rejected.
142 **sortu** ‘to be born’ should be either demoted to “tentative” or rejected.
148 **bihotz** ‘heart’: bi- < *swe* not *swi*.
149 **bortz, bost** ‘five’: a loan from Celtic (Proto-Celtic *bostā* ‘palm, fist’) is more likely that < PIE *penkw-ro-st-.*
152 **euskara** ‘Basque language’: Irigoyen’s internal etymology is stronger than mine.
154 **haritz** ‘oak’ probably < *ha-leitz* (based on internal evidence); this segmentation invalidates my tentative etymology.
162 **ortzi, ortze, ost**- ‘sky, storm, thunder; sky god’ should be promoted to “likely” status.
235 **itsu** ‘blind’: probably via *intsú-*.  
The above changes have the following impact on my 126 “likely” etymologies < PIE:

- two are reclassified as loans (#2, 94)
- four are demoted to “tentative” (#10, 46, 111, 123)
- the others change slightly, but remain likely
- one “tentative” etymology is promoted to likely (#162)

so the net effect is that my 126 likely etymologies < PIE decrease by 5, or 4%.
5. Conclusions

The above analysis of comments results in minor changes to sound laws and etymologies contained in my article. I therefore believe its conclusions remain valid: most of the Basque native basic lexicon derives from PIE via regular sound laws; the most likely explanation is that Basque is IE; a possible alternative explanation is that Basque originated from a creole with mainly IE lexicon.

34 But is probably not a Celtic language, though its phonology and lexicon were influenced by Celtic languages through areal contact.