

## PREFACE TO PROTO-INDO-EUROPEAN LINGUISTICS

### URINDOGERMANISCHE SPRACHFORSCHUNGEN

#### 1. On Proto-Indo-European Linguistics – Urindogermanische Sprachforschungen

Welcome to *Proto-Indo-European Linguistics – Urindogermanische Sprachforschungen*, a new digital journal in Indo-European linguistics.

The timing of the founding of the journal in 2016 is no coincidence, as Indo-European linguistics celebrates the 230th anniversary of Sir WILLIAM JONES's announcement of the descent of the Indo-European languages from a common source, now known as Proto-Indo-European. This year also marks the 200th anniversary of a key work of FRANZ BOPP (1816) and coincides with the 100th anniversary of BEDŘICH HROZNÝ's announcement and proof of the Indo-European (IE) character of Hittite in 1915 and 1917 respectively.

Beyond these celebrations there is an even more compelling reason to launch a new journal: the recent publication of the *glottal fricative theory* (GFT), containing the recently revised Proto-Indo-European (PIE) reconstruction theory, and its ground-breaking ramifications. The GFT originally appeared in JOUNA PYYSALO's dissertation *System PIE: The Primary Phoneme Inventory and Sound Law system for Proto-Indo-European* in 2013 and has now been digitally tested and proven to be correct in *Proto-Indo-European Lexicon* (PIE Lexicon) at <http://pielexicon.hum.helsinki.fi>.

PIE Lexicon, the computerised version of the GFT, presents the first purely comparative reconstruction theory implemented as an operating system ever in the history of Indo-European linguistics. In practice, the entire PIE reconstruction has been revised to only contain comparative components exclusively inferred by means of the principle of postulation, i.e. FICK's rule of two witnesses. The reconstruction of the GFT is therefore the inductive equivalent of the IE data, and hence PIE Lexicon is capable of automatically generating the data from PIE on the basis of digitised sound laws. This logical equivalence of Indo-European and Proto-Indo-European means that at this very moment the field is in a transition phase from Indo-European linguistics into Proto-Indo-European linguistics, a development reflected in the title of the new journal, *Proto-Indo-European Linguistics*.

**1.1** Although *PIE Linguistics* (PIEL) functions as the publishing arm of PIE Lexicon, the journal employs a peer-review system and publishes articles by any author as long as our standards are met. The initially adopted anonymous peer review system will be turned into open peer review as soon as possible due to the superior qualities of the latter.

Although *PIE Linguistics* focuses on articles containing new proposals, we are confident that the correctness of a proposal is more important than its novelty. In addition, specifying a real reconstruction problem – whether a solution is offered or not – is more valuable than a new proposal that is false in content. These preferences reflect the fact that the mission of PIEL is to identify, describe and solve each and every research problem of Indo-European linguistics. Consequently, our editorial policy will accept articles that describe real problems even if no solutions are presented and simultaneously reject novel proposals if they are found to be incorrect.

As such, *PIE Linguistics* offers a consistent and coordinated editorial policy, which is often absent in scientific journals adjusted to serve the benefits of the publishing companies rather than science itself. Although plurality has its undeniable functions, we feel that the researchers should own, publish, and distribute the most important scientific publications in order to secure the progress of their own field, and have the right to define what science and its progress are in the field in question. The privilege of creating and publishing science is not only the supreme prerogative of science, it is also its greatest opportunity second only to that of the discoveries themselves.

By adopting these principles, the scholars of the field will save precious time and the publishers and funding will cease to be the masters and become servants of science. Instead of a biblical flood of publications that no one is able to control, we need less, but better targeted articles solving the remaining fundamental problems of Indo-European linguistics and defining the new, emerging problems brought forth by the new horizon of the digitised Proto-Indo-European parent language, the next great object of research once the problems of the Indo-European level have been solved.

**1.2** In order to solve the key problem of the modern publishing system, the absence of a mechanism for replication of the results, *PIE Linguistics* has already deployed digital tools for the verification and/or falsification of the suggestions in the submitted articles. These tools, essentially the GFT and the comparative method of reconstruction in digitised form, supported with language technology tools, will allow our editors to immediately test the content of submitted articles.

The GFT in itself is a synthesis of the correct research results presented during the two centuries of research in Indo-European linguistics, chosen,

perfected and combined into a single unified theory in PYYSALO 2013. As such the GFT is built upon the successful proposals of all great schools of reconstruction, the Paleogrammarians, the Neogrammarians, the laryngeal theory, and monolaryngealism, to which correct proposals of numerous independent scholars have been added. The GFT therefore stands on the shoulders of giants. However, a next-generation theory is initially not easy to understand and perhaps to accept for the mainstream, whence we have adopted the strategy that once led the leading Neogrammarians, KARL BRUGMANN and HERMANN OSTHOFF, to success. Just as they established *Morphologische Untersuchungen* to introduce and develop their new theory, a new journal, *PIE Linguistics*, is launched today for the GFT.

**1.3** By 2017 all key conjunctures of the GFT have been digitised, tested and verified by means of the *foma* programming language by MANS HULDEN (2009), a digitised predicate calculus, a standard tool of expression of natural science. A provably consistent set of digitised sound laws (a.k.a. *foma* rules) has been coded and is now being published online in the Proto-Indo-European Lexicon Project, the generative etymological dictionary of Indo-European languages, at the address <http://pielexicon.hum.helsinki.fi>.

Currently PIE Lexicon automatically generates stem forms of some 120 most ancient Indo-European language by means of *foma* scripts, chronologically arranged sound laws in code format. As the main bulk of the earliest sound laws has already been coded, only some later, mostly well known sound laws remain to be added. As a result Indo-European linguistics has become a branch of natural science by the end of 2016 in terms of the key feature, predictability: Exactly as ISAAC NEWTON was able to predict the movements and positions of physical objects after presenting the laws of gravity and calculus required to accomplish the computations, Indo-European linguistics is now capable of generating (i.e. predicting) the Indo-European words by means of digitised sound laws.

## 2 An outline of the research history of Indo-European linguistics up to this day

**2.1** In the 1850s AUGUST SCHLEICHER was the first to recognise comparative Indo-European linguistics as a branch of natural science.<sup>1</sup> This conclusion, forgotten but not lost, is now proven by PIE Lexicon's capability of automatically generating (i.e. predicting) the data, the trademark of natural science. Accordingly, Indo-European linguistics is an empirical science with a single foundation, the Indo-European data. All conjectures must be derived exclusively from this basis by means of direct observation of measurable features of the data. The comparative method of reconstruction is the gold standard, i.e. the sole justification for a reconstruction is the principle of postulation, according to which

“durch zweier Zeugen Mund wird alle Wahrheit kund” (AUGUST FICK)<sup>2</sup>

This principle – as explained by HOLGER PEDERSEN (1962:274) – means that

“If a word [or an object of any level] is found in the two branches, then it was also to be found in the original language which divided into these branches.”

This principle, when applied recursively to the full data, legitimises the reconstruction and leads to the inductive restoration of Proto-Indo-European as the logical equivalent of the data as they have been preserved.

**2.2** In 1786 Sir WILLIAM JONES sensationally announced the genetic relationship between several European languages such as Greek, Latin, and Sanskrit, all of which derived from his conception of the proto-language, a ‘common source, which, perhaps, no longer exists’. Although Sir WILLIAM's idea of an independent proto-language has been endorsed ever since, in practice the researchers in Indo-European linguistics have repeatedly opted for the exact opposite view of the proto-language with far-reaching effects.

**2.3** Not long after the first generation of pioneers, FRANZ BOPP, RASMUS RASK, and JACOB GRIMM, had laid down the foundations of the field, the Paleogrammarian school exchanged Sir WILLIAM's paradigm of an independent Proto-Indo-European language for postulation of Sanskrit as the proto-language *an sich*. Although the Sanskrit-centrism of the Paleogrammarians scored

---

<sup>1</sup> For two modern interpretations on SCHLEICHER's views on Indo-European linguistics as natural science, see KOERNER (1982:2) and FOX (1995:24).

<sup>2</sup> The ‘principle of postulation’ first appeared in the cover page of FICK's etymological dictionary (1870-1871).

successes in the reconstruction of certain segments of the proto-phoneme inventory in which Sanskrit was the most conservative language indeed, the all-out generalisation of the hypothesis led to an unnecessary rift in the field. In elevating Sanskrit as the proto-language the mainstream had chosen the easy (non-comparative) way, hence appealing to the great majority of the establishment, soon to be shocked when a generation of young scholars, to be known as the Neogrammarians, realised on external grounds that the European languages were more conservative than Sanskrit with regard to the vowel system. Therefore, as the Neogrammarians correctly pointed out, Sanskrit was not the proto-language, as the paradigm of the now doomed Paleogrammarian school had erroneously believed.

**2.4 KARL BRUGMANN**, the leading Neogrammarian, caused a sensation in 1876 by inserting two articles (1876a, 1876b) in volume 9 of GEORG CURTIUS' *Studien zur griechischen und lateinischen Grammatik* without authorisation of CURTIUS, leading to the publication of some of BRUGMANN's new ideas. In a note added to the same volume CURTIUS disavowed BRUGMANN, who in turn launched a new journal, *Morphologische Untersuchungen auf dem Gebiete der Indogermanischen Sprachen* (MU), with HERMANN OSTHOFF in 1878. Through MU and other journals the Neogrammarians ultimately scored a success when their advanced system of eight proto-vowels Neogr. \*a e o ā ē ō ə ǎ replaced the two cover symbols (Paleogr. \*a ā) of the former, Sanskrito-centric vowel system.

Despite this correction, the Neogrammarian theory was never completely free from Sanskrito-centrism. On the contrary, OSTHOFF and BRUGMANN introduced a new Sanskrito-centric component that was absent in the Paleogrammarian theory. This segment of reconstruction was nothing else than the most famous Neogrammarian theory of all, that of *syllabic sonants*, which assumes *a priori* that Sanskrit, where 'svarabhakti vowels' are usually not found, reflects the original state of affairs, whereas the rest of the group has innovated the svarabhakti vowels, allegedly emerging from syllabic sonants losing their syllabicity.

This internal, Sanskrito-centric approach is not (nor was it ever) the necessary one, because the vastly extended data now at our disposal securely imply that the alleged 'svarabhakti vowels' of the Indo-European languages have external parallels, which implies their PIE origin instead of emergence from the syllabic sonants. In comparisons such as RV. *dr̥ś-* : OIr. *drach-*, Gr. *δρακ-* Sanskrit can be understood as representing a zero grade form PIE \**drahki-* (≈ Neogr. \**dr̥k-*), whereas the Celtic and Greek forms represent a full grade PIE \**dreaḥki-* (≈ Neogr. \**drak-*), therefore preserving the vowel instead of innovating it (see §3.2.a).

**2.5** After BEDŘICH HROZNÝ had interpreted (1915) and proven (1917) Hittite to be the oldest known Indo-European language, HERMANN MØLLER's idea, presented in the days of the Neogrammarians, of the existence of segmental laryngeals in PIE was sensationally confirmed – to a degree: A single laryngeal was directly preserved as Hitt. *ḫ*, as was shown independently by three scholars, JERZY KURYŁOWICZ (1927), ALBERT CUNY (1927) and EDGAR STURTEVANT (1928). As the attention of the field turned to the laryngeal and its relation with the PIE vowel system, the Neogrammarian theory of syllabic sonants was accepted without scrutiny, and their advanced reconstruction of the vowels was radically remodeled in the new laryngeal theory (LT).

MØLLER's intention when proposing his theory was by no means restricted to stating that PIE might have had a laryngeal. On the contrary, he believed (or wanted to believe) in a genetic relationship between the Indo-European and Semitic languages. In order to justify his views, it was MØLLER's turn to replace Sir WILLIAM's paradigm of an independent, comparatively postulated Proto-Indo-European parent language with a Proto-Semitic one, implemented as follows:

- (a) The Indo-European languages and PIE, being genetically related, share the Proto-Semitic root structure  $C_1C_2\cdot C_3$  (appearing in two vocalizations  $C_1eC_2\cdot C_3$ , :  $C_1C_2\cdot eC_3$ ).
- (b) The three Indo-European vowel qualities /e a o/ reflect the 'laryngeals' \*E A O (= \* $h_1 h_2 h_3$ ), actually 'gutturals' coinciding with their counterparts in MØLLER's Proto-Semitic.

Together these two features, when analogy is admitted to dismiss the exceptions, constitute a methodology to substitute (P)IE vowel qualities and quantities with alleged Proto-Indo-Semitic 'laryngeals' in a now well-known manner.

In addition to its non-genetic (Semitic) typology, MØLLER's suggestion would also have meant a step backwards in the reconstruction of the proto-language, since it reduced the more advanced, comparatively inferred Neogrammarian system of eight vowels into a one-vowel system with \* $h_xe$  \* $eh_x$ , the near-equivalent of the outdated Sanskrito-centric vowel typology Paleogr. \*a \*ā.

This mixture of unverifiable, false and correct proposals and a non-genetic typology made the laryngeal theory a true φάρμακον for Indo-European linguistics. The Neogrammarians correctly judged MØLLER's scholarship as unsound, and no takers were found for the toxic mixture among their successors either. On the other hand, the Semitic core of Møller's ideas made it a remedy for the main developers and proponents of the laryngeal theory such as ALBERT CUNY, JERZY KURYŁOWICZ, ÉMILE BENVENISTE, HEINER EICHNER, and MANFRED MAYRHOFER, all also Semitic linguists or native speakers of Semitic languages.

They swallowed the theory whole, since it was a far easier option than managing the Indo-European data by means of the comparative method of reconstruction.

Independently of the two opposite receptions, the non-genetic origin of the laryngeal theory made it incompatible with the Indo-European data from the very beginning. The subsequent research history of the theory is a succession of redefinitions of the properties of the laryngeals  $*h_1$   $h_2$   $h_3$  in order to make these and the Semitic root axiom compatible with the Indo-European data, especially with that of the Old Anatolian group of languages, Hittite, Palaic, Cuneiform Luwian and Hieroglyphic Luwian.

The high water mark of the laryngeal theory was already reached more than a generation ago, when HEINER EICHNER argued in a series of articles beginning in the 1970s that  $*h_2$  was the only laryngeal preserved in Hittite.<sup>3</sup> In their responses CRAIG H. MELCHERT (1987) and FREDERIK KORTLANDT (1984, 2004) presented the remaining alternatives for a partial loss of  $*h_3$ , thus exhausting the feasible theoretical distributions of lost/preserved laryngeals in Old Anatolian.

The situation, described in detail with research history and its analysis, is presented in PYYSALO 2016 and can be summarised in the following table, in which “h” stands for preserved, “H” for lost and “h/H” for a laryngeal preserved/lost in specified environments in Old Anatolian:

MØLLER [PUHVEL]	$*h_1$	$h_2$	$h_3$	[H <sub>4</sub> ]	[H <sub>5</sub> ]	[H <sub>6</sub> ]
BENVENISTE	$*H_1$	$h_2$	$h_3$			
EICHNER [KURYŁOWICZ]	$*H_1$	$h_2$	H <sub>3</sub>	[H <sub>4</sub> ]		
MELCHERT	$*H_1$	$h_2$	h/H <sub>3</sub>			
KORTLANDT	$*H_1$	h/H <sub>2</sub>	h/H <sub>3</sub>			

By the beginning of the new millennium all theoretically meaningful permutations of lost/preserved h/H<sub>x</sub> had thus already been presented and the theory emptied. Despite this, all models of the laryngeal theory remain incompatible with Hittite, Palaic, Cuneiform Luwian, Hieroglyphic Luwian, and the Indo-European data in general. In such circumstances trilaryngealism and the laryngeal theory no longer have a theory. This state of affairs is permanent, because no additional realistic distributions exist.

---

<sup>3</sup> EICHNER (1978:162, fn77): “Von den in der Literatur für anatol.  $h$ - ←  $*h_3$ - genannten Beispielen ist keines sicher, alle können auch mit  $*h_2$ - angesetzt werden (Material bei F. O. Lindeman, Einführung in die Laryngaltheorie, Berlin 1970, § 27).”

For Indo-European linguistics this means that the tables have been permanently turned. In 1988 HEINER EICHNER could still maintain that there were only two scientifically possible candidates for the solution of the Proto-Indo-European laryngeal problem, monolaryngealism (*Monolaryngalismus*) and trilaryngealism (*Trilaryngalismus*).<sup>4</sup> By 2016, however, only one option, monolaryngealism and its revised continuation, the glottal fricative theory, remains.

**2.6** The fourth and final major historical school of reconstruction, monolaryngealism, was proposed in 1951 by LADISLAV ZGUSTA, who provided the first preliminary proof that only one laryngeal, PIE \*H, can be postulated on the basis of Hitt. ḫ. In so doing ZGUSTA implicitly discarded MØLLER's Semitic root hypothesis and in addition presented the phonetically (and phonologically) well-motivated idea that this 'laryngeal' had no colouring effect.<sup>5</sup> As ZGUSTA, however, accepted the rule of compensatory lengthening, his phoneme inventory for the vowels and laryngeals consisted of system of six items, Zg. \*e a o H eH aH oH.

Further improvements in the theory were introduced from the 1960s onwards, in particular by OSWALD SZEMERÉNYI (1967, 1970, 1996), who in addition to defining the laryngeal as a glottal fricative (IPA h) accepted the existence of original long PIE vowels, thus introducing a system Sz. \*a e o ā ē ō ə h (see SZEMERÉNYI 1996:40) without the compensatory lengthening rule.

This kind of system is also the basis of JOHANN TISCHLER's *Hethitisches etymologisches Glossar* (1977-2016), the monolaryngealist Anatolian etymological dictionary, which is still the most reliable source of its category, despite being partly outdated especially for the first volume.

Despite the undeniable success in the simplification of the PIE fricative system (from the obsolete LT \*h<sub>1</sub> h<sub>2</sub> h<sub>3</sub> to a single PIE \*h) and the respective expansion of the vowel system into PIE \*a e o ā ē ō ə, almost identical to the Neogrammarian eight cover symbols, also early monolaryngealism suffered from transition problems especially in two directions:

- (a) As pointed out by EICHNER (1988), monolaryngealism never presented a theory fully independently of the Neogrammarians – as especially the critical relationship between 'schwa' and PIE \*h = Hitt. ḫ remained

---

<sup>4</sup> EICHNER (1988:124): "Hinsichtlich der **Anzahl** der postulierten urindogermanisch-grundsprachlichen Laryngale läßt sich beobachten, daß gegenwärtig nur noch zwei Konzeptionen von einem jeweils größeren Zahl gut unterrichteter Forscher vertreten werden, nämlich der Monolaryngalismus (ein Laryngal) und der Trilaryngalismus (drei Laryngale)."

<sup>5</sup> The reason for this is that the place of articulation of vowels (i.e the place where qualities of vowels are actually produced) is in the mouth cavity and not in the larynx.

unexplicated.<sup>6</sup> The criticism is well-aimed, because in the laryngeal theory these issues have been explained ever since MØLLER's interpretation of DE SAUSSURE's \*A (= Neogr. \*ə) as a 'laryngeal' consonant and KURYŁOWICZ's identification of MØLLER's \*A with Hitt. ḫ.<sup>7</sup>

- (b) Secondly, early monolaryngealism never clarified its relationship with the laryngeal theory, especially with regard to the phenomenon explained by the 'colouring effect' of LT \*h<sub>2</sub>. ZGUSTA posits \*H as being explicitly non-colouring, which SZEMERÉNYI's PIE \*h obviously is simply on the basis of its phonetic interpretation as a glottal fricative. Denying the colouring effect, as phonetically realistic it is, leaves the connection between PIE \*h and the 'a-vocalism' of the cognates without explanation. However tentative the laryngealist explanation might be, it strikes a chord, because in a vast majority of the correspondences PIE \*h (= Hitt. ḫ) indeed correlates with 'a-vocalism'.<sup>8</sup> In other words, the comparative method itself supports the laryngeal theory in this observation, which is also embedded in the traditional ablaut scheme Neogr. \* ā: ə (DE SAUSSURE \*eA : A), deeply rooted in the research history.

Leaving these issues unexplored certainly delayed the acceptance of monolaryngealism, as the theory created more problems with regard to them than it could offer solutions for.

### 3 The digitised glottal fricative theory, comparative method, and Indo-European linguistics

**3.1** The glottal fricative theory has been designed for two main purposes: to draw the line between the correct and incorrect proposals presented during two centuries of history of the field and to identify and solve the remaining open research problems of Indo-European linguistics. The two aims supplement each other and actually form a single agenda in the greater task of creating a common

---

<sup>6</sup> EICHNER (1988:128): "Er [= der Monolaryngalismus] bildet im Grunde die Fortsetzung der Brugmannischen Auffassungen vermerht um die Ansicht, daß man nach der Entdeckung der anatolischen Evidenz nicht mehr ganz ohne Laryngal auskommt."

<sup>7</sup> Finally, it can be mentioned that EICHNER, not unlike CUNY (1912), assumed that the laryngeals \*h<sub>x</sub> had a vocalic variant/allophone \*ə<sub>x</sub>.

<sup>8</sup> In this sense it is the absence of the 'a-colouring' in forms implying PIE \*e/ē in environment PIE \*H and the true cause of the colouring effect (as the colouring of \*H itself is not feasible) that require an explanation.

understanding of the field for the professionals. The bird's-eye view on the research history of Indo-European linguistics indicates that in addition to the real data problems there are man-made ones unwittingly caused by the researchers themselves.

The four most important problems, which have often been inherited in transitions between theories, are separately discussed below in §§3.2-3.5. As will be shown, each historical school contains a cumulative mixture of true and false propositions: Rather than steadily improving the general theory of Indo-European linguistics by adopting the true propositions of their predecessors and adding their own, subsequently proven inventions, they have been unable to recognize all correct features of the older theories and have added false proposals, therefore augmenting problems. As an outcome even the modern state-of-the-art theories such as trilaryngealism and monolaryngealism carry an extra weight, making them unsustainable in the long run.

**3.2** The most frequent reason for the emergence of major errors in the reconstruction of Proto-Indo-European is the replacement of the comparative method, the gold standard, with inferior methodologies elevated as alternative paradigms. In these cases the quasi-paradigm yields and accumulates false results that increasingly plague the reconstruction until replaced with the comparative solution. The most significant problems caused by this, now corrected in the glottal fricative theory, are:

(a) *The PIE vowel system.* The Paleogrammarians abandoned Sir WILLIAM's concept of an independent proto-language and identified Sanskrit with Proto-Indo-European. This led to a dead end especially in the reconstruction of the PIE vowel system and the PIE velar system. Despite the Neogrammarian success in replacing the Sanskrit-centric vowel system (Paleogr. \*a ā) with a comparative solution of eight cover symbols for vowels (Neogr. \*a e o ā ē ō ə â), the tables were turned again after MØLLER had replaced this comparative Neogrammarian paradigm with his Proto-Indo-Semitic one. After KURYŁOWICZ's (1935) and BENVENISTE's (1935) interpretations this has become the mainstream of today, leading back to square one.

The early problems of the vowel system have now been corrected in the GFT through the postulation of the comparative counterparts of the eight Neogrammarian cover symbols for vowels, upgraded with the rules for \*A and PIE \*H. The core of the solution, sufficient to reconstruct of any attested vowel patterns of the cognates, stands as follows:

1) The eight Neogrammarian correspondence sets for vowels represent the following respective PIE starting points:

Neogr. \*ə = PIE \*a    Neogr. \*a = PIE \*ae/\*ea    Neogr. \*ā = PIE \*aē/\*ēa  
 Neogr. \*ā̃ = PIE \*o    Neogr. \*o = PIE \*oH    Neogr. \*ō = PIE \*ō  
 Neogr. \*e = PIE \*e    Neogr. \*ē = PIE \*ē

2) From the laryngeal theory the colouring (or properly assimilation) rules for \*eA (DE SAUSSURE) and \*Ae (MØLLER) are accepted, rewritten as PIE \*ea → \*aa and PIE \*ae → aa. Identical assimilation rules are added for PIE \*ē: PIE \*ēa → \*āa and PIE \*aē → aā.<sup>9</sup>

3) The \*A-loss rules of DE SAUSSURE (but without his compensatory lengthening) and MØLLER for the environments \*eA Ae are accepted with the addition that the ‘A-loss’ occurs in all environments for unaccented PIE \*a.<sup>10</sup>

The ‘laryngeal problem’ turned out to be the great watershed of Indo-European linguistics of 20th century: The anti-laryngealists perished refusing to take the φάρμακον, the three-laryngealists perished overdosing, and the progress of monolaryngealism – though it survived because it avoided these mistakes – was delayed for decades, because the correct ingredients and doses were not actively sought from the beginning. Had the Neogrammarians been more objective in their evaluation of the ideas of DE SAUSSURE and MØLLER, not all of which were wrong – some of them were even indispensable – Indo-European linguistics might have been able to avoid this disaster and progress much faster than it actually did.

(b) *The PIE sonorant system.* Immediately after their comparative reconstruction of the PIE vowel system, the Neogrammarians replaced the comparative method with a Sanskrito-centric paradigm in their reconstruction of the syllabic sonorants. Although it remains true that Indo-Iranian is the only group preserving the syllabic liquids PIE \*r̥ \*l̥, the sonorants PIE \*l̥ \*r̥ \*m̥ \*n̥ never yielded ‘svarabhakti vowels’ in any ancient IE languages. This is now confirmed by the latest and far more complete data, in which the vowels in question are paralleled by at least two IE subgroups (FICK’S RULE). Therefore the principle of postulation implies their PIE origin (see PYYSALO 2013: 181-344 for key examples). Consequently the comparative solution already used by certain Paleogrammarians is restored in the GFT and used to explain these vowels as follows instead of deriving them from syllabic sonorants:

<sup>9</sup> Contrary to the consonantal interpretations of DE SAUSSURE, MØLLER and KURYŁOWICZ, \*A is considered a vowel corresponding to Neogr. \*ə, reinterpreted as PIE \*a. In addition there is no compensatory lengthening, i.e. PIE \*ea → aa → Lat. a, OInd. a, etc.

<sup>10</sup> The unaccented vowel PIE \*a → Ø everywhere. Simultaneously the Neogrammarian ‘schwa’ rule applies only to the respective accented vowel, i.e. PIE \*á → Lat. a : OInd. i.

1) PIE \* $\text{ṛ}$  \* $\text{ṛ}$  \* $\text{ṃ}$  \* $\text{ṇ}$  are postulated for the proto-language, exactly as was done by the Neogrammarians, but they resulted in respective consonants (IE l r m n) except for \* $\text{ṛ}$  continued in Indo-Iranian when in a consonantal environment.<sup>11</sup>

2) The ‘svarabhakti vowels’ of OSTHOFF and BRUGMANN are explained as original on the basis of comparison with IE parallels preserving identical vowels. In practice this means applying reconstruction à la KARL VERNER (1877:125) who compared, for instance, the items RV. pūrṇá- ‘voll, gefüllt’ : Goth. full- ‘full’ : ORus. pūlnŭ- ‘voll’: Several branches confirm an identical vowel in the first syllable, i.e. the principle of postulation implies PIE \*pulno- for the correspondence.

(c) *The PIE fricative system.* Although overstated and postulated for false reasons, HERMANN MØLLER’s general conjecture of the existence of ‘laryngeals’ contains a seed of truth in the form of a single laryngeal \*H, i.e. the glottal fricative PIE \*h/fi = Hitt. ḫ. In general, however, MØLLER’s laryngeal theory was never postulated according to the comparative method, but rather replaced it with a Semitic typology.

The number of MØLLER’s laryngeals was correctly restricted to a single PIE \*H already by ZGUSTA (1951). SZEMERÉNYI (1967, 1970) further defined the phoneme as a glottal fricative and explicitly removed the Proto-Semitic root constraint from the reconstruction of PIE as ‘not binding’ (1967:92-93).<sup>12</sup> A voiced alternative PIE \*fi was added to PIE \*h by PYYSALO (2013), who in the same study also specified the answer to the most difficult problem of all, that of the properties of the ‘laryngeal’ and its relation to the Indo-European vowel system, especially with regard to the ‘a-vocalism’ and syllabicity. The comparative solution of the GFT combines the correct proposals of the Neogrammarian vowel system, the laryngeal theory, and monolaryngealism into a single solution as follows:

<sup>11</sup> Except for the continuation of PIE \* $\text{ṛ}$  \* $\text{ṛ}$  in Indo-Iranian, the consonantal outcomes of PIE \* $\text{ṛ}$  \* $\text{ṛ}$  \* $\text{ṃ}$  \* $\text{ṇ}$  are now widely documented in the original environment with PIE \*h/fi (≡ Hitt. ḫ). After the loss of PIE \*h/fi the respective outcomes in all IE languages are l r m n without svarabhakti vowels or vocalisation of nasals (see PYYSALO 2013: 332-334 for a summary).

<sup>12</sup> In 2015 VOYLES and BARRACK (<http://linguistlist.org/issues/26/26-4746.html>) condemn the laryngealist methodologies in Linguist List (26.4746) in the introduction of their book ‘On Laryngealism’ writing: “Chapter 6 ‘Logic and laryngealism.’ places the laryngeal theory within the spectrum of other false theories which have been proposed from time to time in various sciences - such as the phlogiston theory in chemistry. Hence laryngealism is shown to have been one of the growing pains in the history of the science of linguistics.”

1) The cover symbol Neogr. \*ə (= DE SAUSSURE \*A) is interpreted as the vowel PIE \*a (with colouring and loss rules already specified above), i.e. MØLLER's interpretation of \*A as a laryngeal and KURYŁOWICZ's identification of \*A with Hitt. ḫ are rejected.

2) A single 'laryngeal', phonetically a glottal voiceless/voiced fricative PIE \*h/h̥ ≡ Hitt. ḫ, is postulated for Proto-Indo-European.

3) PIE \*h/h̥ (≡ Hitt. ḫ) always appears in connection with PIE \*a, i.e. the phonemes stand in strict phonotactic selection in the diphenemic pairs PIE \*ha \*ah \*h̥a \*aḥ.<sup>13</sup>

In practice the comparative solution means that neither the traditional (Neogr. \*pāter-) or the modern (LT \*ph<sub>2</sub>ter-) reconstruction is correct as such, but only the two together result in the correct reconstruction: Accordingly, both Neogr. \*ə (= PIE \*a) and LT \*h<sub>2</sub> (= PIE \*h) are reconstructed simultaneously in the GFT in PIE \*pahter- 'father' and so forth.

**3.3** The second most frequent reason for major errors in the reconstruction can be traced back to the Neogrammarians, more specifically to KARL BRUGMANN himself. In the *Neogrammarian manifesto*, the preface to *Morphologische Untersuchungen* co-signed by HERMANN OSTHOFF, BRUGMANN proposed a novel twofold interpretation of sound changes (and sound laws).

(a) On one hand, according to BRUGMANN:

"First, every sound change, inasmuch as it occurs mechanically, takes place according to laws that admit no exception. That is, the direction of the sound shift is always the same for all the members of a linguistic community except where a split into dialects occurs; and all words in which the sound subjected to the change appears in the same relationship are affected by the change without exception."<sup>14</sup>

The emphasis on the regularity of sound change traces back to AUGUST SCHLEICHER, the leading theoretician of the Paleogrammarian movement, whose idea of *Ausnahmslosigkeit der Lautgesetze* his pupil AUGUST LESKIEN made the famous slogan of the Neogrammarian movement.

<sup>13</sup> The syllabic and colouring features are thus not properties of the 'laryngeal \*h<sub>2</sub>', but of 'schwa', i.e. PIE \*a, exactly as stated in the colouring rules of DE SAUSSURE and MØLLER.

<sup>14</sup> BRUGMANN and OSTHOFF (1878: xiii-xiv): "*Erstens*. Aller lautwandel, so weit er mechanisch vor sich geht, vollzieht sich nach *ausnahmslosen gesetzen*, d.h. die richtung der lautbewegung ist bei allen angehörigen einer sprachgenossenschaft, ausser dem fall, dass dialektspaltung eintritt, stets dieselbe, und alle wörter, in denen der der lautbewegung unterworfenen laut unter gleichen verhältnissen erscheint, werden ohne ausnahme von den änderung ergriffen."

What strikes one as odd in young BRUGMANN's description is his way of describing the relation between sound changes and sound laws. According to BRUGMANN 'sound change (...) takes place according to laws that admit no exception'. This is in fact true if and only if the sound laws, which rather describe than govern the sound changes, are correctly formulated. This, in turn, was and still is far from always the case in all four main historical reconstruction theories.

(b) On the other hand BRUGMANN admits the existence of irregular changes, and insists that if a form does not behave according to the sound laws, it is analogy.<sup>15</sup> Due to the later prestige of the Neogrammarians, lifting up this doctrine as the standard of linguistic research, this marks a decisive turning point in Indo-European linguistics: In reality, irregularity has an alternative explanation to analogy, namely an erroneously formulated sound law that fails to describe the sound change perfectly.

In other words, BRUGMANN does not distinguish between sound laws, potentially fallible, and the respective sound changes, that indeed admit no exceptions. Consequently, if a form appears to be irregular, it is not necessarily an analogy, but a failure in the formulation of the sound laws. The replacement of this explanation has deprived Indo-European linguistics of one of the most powerful tools of science, the self-correcting process, ever since it was introduced. After irregularities became explained as analogy, few real corrections to sound law problems have been introduced. Consequently, a considerable portion of suggested sound laws have remained without the final touch, as the explanation for the exceptions has been sought in an error in language instead of a human error in the formulation of the sound laws.

This error has been removed from the GFT by restoring the self-correcting procedure in Indo-European linguistics: Instead of analogy, the preferred explanation for irregularity is incompleteness or a human error in the related set of sound laws. Accordingly, the sound laws themselves are scrutinised and corrected in order to eliminate the apparent irregularities and allow for regular generation of the data in Indo-European linguistics in the future.

---

<sup>15</sup> BRUGMANN and OSTHOFF (1878: xiii-xiv): "*Zweitens*. Da sich klar herausstellt, dass die form-association, d. h. die neubildung von sprachformen auf dem wege der analogie, im leben der *neueren* sprachen eine sehr bedeutende rolle spielt, so ist diese art von spracherneuerung unbedenklich auch für die *älteren* und *ältesten* perioden anzuerkennen, und nicht nur überhaupt hier anzuerkennen, sondern es ist dieses erklärungsprincip auch in derselben weise zu verwerten, wie zur erklärungs von spracherscheinungen späterer perioden [...]". For BRUGMANN insisting that analogy should be used automatically if the sound laws failed, see (1879a:6): "In allen anderen fallen, in denen wir abweichung vom allgemeingiltigen gesetz finden, haben wir eine association (analogie) zu statuieren."

**3.4** The third cause for major errors in Indo-European linguistics, characteristic for early phases of scientific enterprises and relatively young branches of science, is the absence of a complete set of unified symbols for the real objects of the field, metaphorically the counterpart of the periodic table in chemistry. The early periods, often also characterised by insufficient understanding of the development of the field and its research history as a whole, tend to contain a blend of proper, combined and non-existing terms for basic elements, such as oxygen, water and aether in chemistry. In the absence of a common, shared basis of the fundamentals of the field confusion and chaos often reign, causing further misunderstandings until the fundamentals have been established for good.

By 2016 also these problems are being resolved, as the GFT also contains the primary phoneme inventory of Proto-Indo-European. This matrix, consisting only of indivisible PIE phonemes, has been obtained through segmental analysis in which all historical reconstruction phonemes of combinatory origin have been defined in terms of primary ones (e.g. Neogr. \*k<sup>u</sup> = PIE \*k+u) and the non-existing ones (e.g. LT \*h<sub>1</sub> h<sub>3</sub>) have been eliminated. As an outcome of the segmental analysis presented in PYYSALO 2013, the GFT postulates only the following primary phonemes for Proto-Indo-European:

PIE *o	*e	*a	*i	*u	*l	*r	*m	*n	*h	*s	*k	*p	*t
PIE *ō	*ē	*ā <sup>?</sup>	*ī	*ū	*ļ	*ŗ	*m̥	*n̥	*h̥	*z	*g	*b	*d

This system consists of fourteen phonemes with two variants for each item: vowels alternate on the axis of short/long, sonorants on the axis vowel/consonant and obstruents (fricatives/stops) on the axis voiceless/voiced.

The segmental analysis mostly deals with the classical stop system of the proto-language, dating back to the Neogrammarians and by and large accepted in both the laryngeal theory and monolaryngealism. The classical stop system is naturally divided into two subsets: the *Decem-Taihun isogloss* reflects the problem of the four manners of articulation of the stops in the series T : Th : D : Dh represented in rows I–IV, and the *Centum-Satem isogloss* reflects the problem of the three places of articulation of velars represented in columns 3–5:

	1.	2.	3.	4.	5.
I	*p	*t	*k	*k̥	*k <sup>u</sup>
II	*ph	*th	*kh	*k̥h	*k <sup>u</sup> h
III	*b	*d	*g	*g̥	*g <sup>u</sup>
IV	*bh	*dh	*gh	*g̥h	*g <sup>u</sup> h

Although this classical stop system is correctly inferred in terms of the comparative method, the very real possibility that some of its phonemes might be secondary, i.e. consist of two or more primary PIE phonemes, has not been systematically explored, although several correct suggestions for individual classes of phonemes have been presented by scholars. The full segmental analysis implemented in both main parts of the classical stop system leads to the following simplifications in the GFT:

(a) The Decem-Taihun isogloss, the four manners of articulation Neogr. T Th D Dh, is derived from the series PIE \*p t k in environments  $\pm$  PIE \*h/ĥ as indicated below:

1) The series *tenues* (Neogr. \*t p k):

Hitt. t, Gr. τ, RV. t	PIE *t	(*t)
Hitt. p, Gr. π, RV. p	PIE *p	(*p)
Hitt. k, Gr. κ, RV. k/c	PIE *k	(*k)

2) The series *tenues aspiratae* (Neogr. \*th ph kh):

Hitt. t, Gr. θ, RV. th	PIE *tah/*tha	(*th)
Hitt. p, Gr. φ, RV. ph	PIE *pah/*pha	(*ph)
Hitt. k, Gr. χ, RV. kh/c	PIE *kah/*kha	(*kh)

3) The series *mediae* (Neogr. \*d b g):

Hitt. d, Gr. δ, RV. d	PIE *ĥ—d/*d—ĥ	(*d)
Hitt. b, Gr. β, RV. b	PIE *ĥ—b/*b—ĥ	(*b)
Hitt. g, Gr. γ, RV. g/j	PIE *ĥ—g/*g—ĥ	(*g)

4) The series *mediae aspiratae* (Neogr. \*dh bh gh):

Hitt. d, Gr. θ, RV. dh	PIE *dĥa/*dāĥ	(*dh)
Hitt. b, Gr. φ, RV. bh	PIE *bĥa/*bāĥ	(*bh)
Hitt. g, Gr. χ, RV. gh/h	PIE *gāĥ/*gĥa	(*gh)

(b) *The Neogrammarian velar system* with three series is simplified to a single velar PIE \*k appearing with PIE \*i/ĭ (= I), PIE \*u/ŭ (= U) and diphonemic PIE \*ha ah ĥa aĥ in various combinatory clusters:

## 1) The plain velar series (Neogr. \*k kh g gh):

Hitt. k, Gr. κ, RV. k/c	PIE *k	(*k)
Hitt. k, Gr. χ, RV. kh/c	PIE *kah/*kha	(*kh)
Hitt. g, Gr. γ, RV. g/j	PIE *ǵ—g/*g—ǵ	(*g)
Hitt. g, Gr. χ, RV. gh/h	PIE *gah/*gha	(*gh)

## 2) The labiovelar series (Neogr. \*kʷ kʷh gʷ gʷh):

Hitt. ku, Gr. π/τ, RV. k/c	PIE *kʷ	(*kʷ)
Hitt. ku, Gr. φ/θ, RV. kh/c	PIE *kahʷ/*khaʷ	(*kʷh)
Hitt. gu, Gr. β/δ, RV. g/j	PIE *ǵʷ—gʷ/*gʷ—ǵʷ	(*gʷ)
Hitt. gu, Gr. φ/θ, RV. gh/h	PIE *gahʷ/*ghaʷ	(*gʷh)

## 3) The palatovelar series (Neogr. \*k̑ kh̑ g̑ gh̑):

Hitt. k, Gr. κ, RV. ś	PIE *k̑	(*k̑)
Hitt. k, Gr. χ, RV. ś	PIE *k̑ah/*k̑iha/*k̑ahI/*k̑ahI	(*k̑h)
Hitt. g, Gr. γ, RV. j (Av. z)	PIE *ǵ̑—g̑I/*g̑I—ǵ̑	(*g̑)
Hitt. g, Gr. χ, RV. h (Av. z)	PIE *ǵ̑ah/*ǵ̑iha/*ǵ̑ahI/*ǵ̑ahI	(*g̑h)

**3.5** The fourth cause for major problems in Indo-European linguistics is the absence of a scientifically neutral and commonly accepted expression of inference, basically a version of predicate calculus with symbols and definitions for the commonplace expressions of the field, allowing us to produce argumentation in mechanical and testable form.

In order to eliminate this defect the GFT endorses language technology and, more broadly, digital humanities (DH), in practice by implementing *foma*, a digitised predicate calculus, as the standard tool of expression of IE linguistics in PIE Lexicon, to be strengthened with other digital auxiliaries in the future. This contribution results in:

(a) A completely transparent and unambiguous scientific means of expression in the generation of Indo-European languages from Proto-Indo-European – and vice versa.

(b) A fully automatic proof procedure generated by the *foma* code reader, allowing testing and verification/falsification of propositions in a mathematically precise manner. As such this digital extension provides Indo-European linguistics with a scientifically solid procedure of verification/falsification, simultaneously defining truth (*Wahrheit*) as provability (*Beweisbarkeit*).

(c) Coding of the comparative method of reconstruction and the standard methodologies of natural science into digital tools attached to the PIE Lexicon operating system in the decades to come.

As a result of these improvements, the early period in Indo-European linguistics, characterised by misunderstandings and a lack of satisfactory scientific discussion between the proponents of the mainstream and the challenging schools, will end during the 21st century.

Since the language technology implemented in PIE Lexicon so far represents only a slight segment of wider digital humanities (DH), it can be readily understood that these improvements increase the external momentum of the field significantly also by making it more understandable and easier to pass forth: The transparency already achieved makes it possible to explain the core functionality of PIE Lexicon, Indo-European linguistics in a digitised form, for a layman in fifteen minutes, and no doubt significant improvements in the teaching and learning of the Indo-European languages can be achieved in the future.

**3.6** The improvements outlined above mean that Indo-European linguistics is reaching maturity as a branch of natural science. The main bulk of direct data problems have already been solved or are approaching their solution, and the non-data-related problems have been eliminated from PIE Lexicon, the digitised version of GFT, in the following manner:

(a) The naïve (non-comparative) methodologies of the historical reconstruction theories have been consistently replaced with the solutions implied by the comparative method of reconstruction, the rules of natural science and methodologies strictly based on these.

(b) The self-correcting process of science has been restored by replacing the hitherto preferred explanation for irregularity, analogy, with improvement and correction of the sound laws. In addition, the task of removing as much irregularity as possible has been made an independent goal in PIE Lexicon, aimed at solving all remaining open or disputed problems of Indo-European linguistics.<sup>16</sup>

---

<sup>16</sup> As an example of solving a neglected problem and taking the result as a part of the solution, we refer to FORTUNATOV's law, for the occurrence of which the required environments (VHLT, VLHT) have now been specified (see PYYSALO 2013:224-243). Consequently, the irregular

(c) The early, problematic phoneme paradigms of the proto-language have now been replaced with a single primary phoneme inventory of Proto-Indo-European (see above). This inventory is capable of generating all phonemes of all Indo-European languages, except for the few remaining open research problems discussed below.

(d) Digital technology has already facilitated the translation of the traditionally formulated version of the GFT (PYYSALO 2013) into a next-generation reconstruction theory, the digitised PIE Lexicon. As the theory is a true synthesis of and compromise between all main schools of reconstruction of the past two centuries, correcting and completing their successful proposals, and its correctness has been digitally tested and proven, the digitised GFT should be relatively easy to accept for the scholars in the field. As such the GFT remains on the table and is constantly improved by means of new digital tools and the publication of additional data.

**3.7** The future of Indo-European linguistics, like that of other humanities, is digital. Accordingly, the study of language change in the 21st century will resemble the discovery of laws of nature in other disciplines seeking the origins of the physical world in its various forms. The sooner we reach this future and digitally manage our method, methodologies and data, the better: Once the momentum has been reached, Indo-European linguistics will transform into a branch of natural science with its own empirical method and methodologies.

Indo-European linguistics of the 21st century will be open access and open source, providing us with further advantages in solving our problems: As the data are freely available, their arrangement in etymologies has been verified, and the corresponding proto-forms are equivalent to the data, they can be used by scholars for further development and definition of higher-level problems such as the structure and formation of the proto-language itself, which will definitely be one of the central fields of research in the 21st century.

In this context the GFT and its digital version, Proto-Indo-European Lexicon, contribute to Indo-European linguistics by offering a synthesis of the research in the form of a fully computerised next-generation discussion matrix.

From a developer's point of view PIE Lexicon initially focuses on coding the entire comparative method of reconstruction with digital technology and the digital presentation of the most ancient Indo-European material at <http://pielexicon.hum.helsinki.fi>. In terms of broad strategy, PIE Lexicon will digitise especially the following key targets:

---

explanation is no longer needed. On the contrary, FORTUNATOV's revised law provides another criterion for identifying the presence of an original PIE \*h/fi.

(a) The Indo-European sound law system, i.e. the sound laws of the most ancient Indo-European languages and dialects. In practice, the sound laws chosen and completed in the GFT are written in the digital source code, the *foma* programming language, a computerised predicate calculus. The digitised sound laws of each language are then chronologically arranged in a *foma* script capable of automatically generating the data of that language from the corresponding reconstructions.

The key content of the GFT, the revised sound law system describing the oldest sound changes, has already been translated into *foma* scripts for over 120 languages or language forms and digitally tested and verified by means of language technology. According to the tests, the predictions of PIE Lexicon already have an accuracy rate of about 99,5 %. This capability in the generation of the data, available for the first time in the history of the field, places Indo-European linguistics between physics and biology in accuracy. The result is encouraging, because it indicates that the reconstruction, i.e. the output of the comparative method, is nearly equivalent to the data already in the very beginning of the digitalisation. This capability is comparable to that of astrophysics to calculate and describe the original state of the universe, but at a more advanced stage, because PIE Lexicon already predicts (generates) the existing reality of the Indo-European languages almost flawlessly.

(b) A fully digitised Indo-European language family tree, the first of its kind, will be generated on the basis of the sound laws common to the individual sub-groups, languages and dialects. As the main bulk of the Indo-European sound laws have been coded, also this task is already being managed and will be completed during 2017 except for some late sound laws for the data not yet uploaded. As a result, the common sound laws of individual languages determine the exact relationships of the Indo-European languages in a language family tree format soon publicly available for scrutiny.

(c) The decision method of Indo-European etymology, originally outlined by AUGUST SCHLEICHER in 1852 (see PYYSALO 2013: §5.1.4 for an explicit version), is an algorithm allowing scholars to discover the Indo-European etymologies. The digital technology for the coding of the decision method already exists and requires only calibration according to the special features of the IE data.

The coding of the decision method, allowing for testing, duplicating and finding etymologies, will begin in PIE Lexicon immediately when the priorities allow and make etymology a completely mechanical procedure no longer depending only on human capabilities. Ultimately, as the digitalisation of the IE data has reached a sufficient level, this feature will enable the field to evaluate all etymologies ever proposed, choose the correct ones and compile research histories of all Indo-European morphemes.

(d) The rules of the mainstream laryngeal theory (LT), trilaryngealism, will be digitised. Once the *foma* scripts of the LT are available, the sound law systems of the LT and the GFT will be compared, tested and evaluated in identical data. The success rates of the theories in the generation of the data will be measured and tools to assess the theories will be made available. This will enable objective public criticism, discussion and improvement of the two theories, as well as motivate the presentation of a challenging theory: In order to be meaningful in the first place the glottal fricative theory needs to at least equal, but in reality perform considerably better than the laryngeal theory, the most commonly accepted state-of-the-art.

(e) The Indo-European data themselves, beginning from the most archaic IE language forms, but extending to the modern ones in a later phase of the project, will be digitised as a whole. Initially the data focus of PIE Lexicon has been on solving the most challenging Indo-European etymologies of the Old Anatolian languages (Hittite, Palaic, Cuneiform Luwian and Hieroglyphic Luwian).

In 2017 a new phase, in which the most ancient data arranged under the PIE initials, starting with PIE \*t-/d-, will be published, will begin. Every etymology and, consequently, root will appear in a generative, digitally verified form, which allows the addition of related data if identified in any IE language later on. In this manner PIE Lexicon will become the most complete etymological dictionary of Indo-European languages in existence in the years to come. The PIE Lexicon entries will be linked to digital Indo-European dictionaries already active and appearing in the internet in the future.

**3.8** The editorial policy of *PIE Linguistics* is committed to the comparative method of reconstruction and the general principles of natural science (see PYYSALO 2013: §1 for an extensive description). We consider Indo-European linguistics a branch of natural science within humanities, as originally envisioned by SCHLEICHER, and we require the same from our partners and contributors for readily understandable reasons: Our mission is to solve the open research problems and to identify the unsolved ones in order to accomplish a qualitative and quantitative leap from Indo-European to Proto-Indo-European linguistics, a process which by definition means a transition to natural science.

As the open research problems of the field are the focus of PIEL, familiarity with the glottal fricative theory (PYYSALO 2013, 2017) and its digital version, PIE Lexicon, will also be required from our contributors. At this point the solved and unsolved problems are being defined in the GFT, the most advanced theory in Indo-European linguistic to date. As is commonplace in natural science, the GFT defines truth as provability, i.e. a proposition or a system of propositions is valid if the related forms have been correctly generated from complete data without

non-existing forms being generated. Reversely, the remaining open research problems of Indo-European linguistics are revealed by failure in their generation in PIE Lexicon as the erroneously generated phonemes are shown in red. The remaining margin of error of is 0,5%, which leaves the following problems for PIEL to focus on:

(a) The PIE accent/tone problem and its comparative solution is the fundamental problem of Indo-European linguistics of today. Only a few observations concerning the PIE accent/tone problem were treated in PYYSALO 2013 (and in the GFT in general), and the problem remains open or at least contested to a large extent: Although a relatively large number of rules have been presented, the distinction between correct and incorrect proposals (in the sense of the GFT) has not been made, and it remains unclear whether all the key problems have even been identified or not, not to mention their solution or the verification of the correct proposals. Accordingly, a single, universal solution that explains all features of the Indo-European languages reflecting the PIE accent/tone is urgently required, and the person or party accomplishing this task will be no doubt be credited for the next, perhaps most vital ‘Einsteinian’ revolution of the field.

(b) On the segmental level there are currently about a dozen minor sound law problems, and yet another dozen (or so) will emerge in the data not yet published in PIE Lexicon. These problems are related to individual dialects, languages or subgroups at most and are hence of later origin and therefore tagged ‘minor’ ones (in contrast to, for instance, the PIE accent/tone problem). A list of errors in the generation of the data in PIE Lexicon is already available on the ‘mismatches’ page at

<http://pielexicon.hum.helsinki.fi/?alpha=ALL&view=mismatch>.

On this page the errors, including those related to the PIE accent/tone problem, are revealed (the page is constantly updated when new data are posted, thus also including possible new errors). As the reader can observe, the same problems, belonging to a dozen subclasses, are repeated within this data, including problematic phonemes such as TochAB. ä, all of which will be described *in extenso* in respective articles to appear in *PIE Linguistics*.

(c) In terms of quantity, the most formidable problem of Indo-European linguistics is the etymology of the entire Indo-European corpus. At this point, there are several generations of etymological dictionaries and, in addition, numerous etymologies presented in articles and books to which one can add the possibility of finding new etymologies by direct application of SCHLEICHER’S decision method. Characteristically the sources quote multiple alternative suggestions, but as a rule no consensus or uniform method of choosing the correct etymologies is included. Consequently, no cohesive attempt of a thorough comparison of the entire Indo-European corpus in order to identify all etymologies,

morphologically arranging these under the PIE roots and eliminating the failing suggestions has been made. In order to improve the situation, *PIE Linguistics* counts this task among its goals and will publish articles containing comparatively provable etymologies.

*Jouna Pyysalo & Mans Hulden*

## References

- BENVENISTE, Émile. 1935. *Origines de la formation des noms en indo-européen I*. Paris: Adrien-Maisonneuve.
- BOPP, Franz. 1816. *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechischen, lateinischen, persischen und germanischen Sprache*. Frankfurt am Main: Andreätschen.
- BRUGMANN, Karl. 1876a. Nasalis sonans in der indogermanische Grundsprache. *Curtius' Studien zur griechischen und lateinischen Grammatik* 9: 285-328.
- 1876b. Zur Geschichte der stammabstufenden Declination. *Curtius' Studien zur griechischen und lateinischen Grammatik* 9: 363-406.
- 1879a. Zur geschichte der nominalsuffixe -as-, -jas- und -vas-. *KZ* 24: 1-99.
- BRUGMANN, Karl & Osthoff, Hermann (1878). Vorwort. *Morphologische Untersuchungen auf dem Gebiete der indogermanische Sprachen*, Vol 1. Leipzig: Hirte, iii-xx.
- CUNY, Albert. 1912. Notes de phonétique historique. Indo-européen et sémitique. *Revue de Phonétique* 2:101-132.
- 1927. Reflexions sur le type χρῆ (ζῆν, etc. . . .) et le type ἦχῶ. (*Symbolae Grammaticae in honorem Ioannis Rozwadowski*, 1:85-94.) Kraków: Drukarnia Uniwersytetu Jagiellońskiego.
- EICHNER, Heiner 1978. Die urindogermanische Wurzel \*H2reu- 'hell machen'. *Die Sprache* 24: 144-162.
- 1988. Anatolisch und Trilaryngalismus. In: Bammesberger 1988ed, *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*. Heidelberg: Winter, 123-151.
- FICK, August. 1870-1871. *Vergleichendes Wörterbuch der indogermanischen Sprachen*. Göttingen: Vandenhoeck & Ruprecht.
- FOX, Anthony. 1995. *Linguistic Reconstruction. An introduction to theory and method*. Oxford: Oxford University Press.
- HROZNÝ, Bedřich. 1915. Die Lösung des hethitischen Problems. *Mitteilungen der Deutschen Orientgesellschaft* 56: 17-50.
- 1917. Die Sprache der Hethiter, ihr Bau und ihre Zugehörigkeit zum indogermanischen Sprachstamm. (*Boghazköi-Studien*, 1-2). Leipzig: Hinrichs.
- HULDEN, Mans (2009). *Finite-State Machine Construction Methods and Algorithms for Phonology and Morphology*. PhD Dissertation, University of Arizona.
- JONES, Sir William. 1788. Anniversary Discourse (February 2nd 1786). *Asiatick Researches* 1: 415-431.
- KOERNER, Konrad. 1982. The Schleicherian Paradigm in Linguistics. *General Linguistics* 22: 1-39.
- KORTLANDT, Frederik H H. 1984, "PIE \*H- in Armenian". *Annual of Armenian Linguistics* 5, 41-43

- 2004, "Initial laryngeals in Anatolian", *Orpheus* 13-14, Memorial Volume to Georgi Rikov, 9-12.
- KURYŁOWICZ, Jerzy. 1927a.  $\text{ə}$  indo-européen et  $\text{h}$  hittite. (*Symbolae Grammaticae in Honorem Ioannis Rozwadowski*, 1:95-104.) Kraków: Drukarnia Uniwersytetu Jagiellońskiego.
- MELCHERT, H. Craig. 1987. Reflexes of  $*\text{h}_3$  in Anatolian. *Die Sprache* 33: 19-28.
- MØLLER, Hermann. 1879. Review of F. Kluge, *Beiträge zur Geschichte der germanischen Conjugation*. *EngStud.* 3: 148-164.
- 1880. Germanisch  $\bar{a}$   $\bar{e}$   $\bar{o}$  in den endungen des nomens und die entstehung des  $\text{o}$  (a2). *PBB* 7: 482-547.
- 1893. Review of Bechtel 1892 (Hauptprobleme der idg. Lautlehre seit Schleicher). *Zeitschrift für deutsche Philologie* 25: 382-384.
- 1906. *Semitisch und Indogermanisch I (Konsonanten)*. Kopenhagen: Hagerup.
- 1911. *Vergleichendes indogermanisch-semitisches Wörterbuch*. Göttingen: Vandenhoeck & Ruprecht.
- PEDERSEN, Holger 1962. The discovery of language. Linguistic science in the nineteenth century. Trans. John W. Spargo. Bloomington (IN): Indiana University Press.
- PYYSALO, Jouna 2013. System PIE: The Primary Phoneme Inventory and Sound Law System for Proto-Indo-European. PhD Dissertation, University of Helsinki. Publications of the Institute for Asian and African Studies 15. Helsinki: Unigrafia Oy. <https://helda.helsinki.fi/handle/10138/41760>
- 2016. Does the Laryngeal Theory have a Theory? On the Incompatibility of a Viable Three-Laryngeal Model with the Anatolian Data. (*Wék<sup>w</sup>os* 2:195-215)
- 2017. The Solution to the Proto-Indo-European Laryngeal Problem: The Glottal Fricative Theory of Hittite  $\text{h}$   $\equiv$  PIE  $*\text{h}/\text{h}$  (To appear in *Wék<sup>w</sup>os* 3)
- RASK, Rasmus Kristian. 1818. *Undersøgelse om det gamle Nordiske eller Islandiske Sprogs Oprindelse. Et af der kongelige Danske Videnskabers-Selskap kronet Prisskrift*. Copenhagen: Gyldendal.
- SAUSSURE, Ferdinand de. 1878 *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*. [= Mém.] ( $\rightarrow$  Rec. 1-268) 1922. *Recueil des publications scientifiques de Ferdinand de Saussure*. [= Rec.] Heidelberg: Winter.
- SCHLEICHER, August. 1852. *Die Formenlehre der kirchenslawischen Sprache, erklärend und vergleichend dargestellt*. Bonn: H.B. König.
- STURTEVANT, Edgar. 1928. Original h in Hittite and the Medio-Passive in r. *Language* 4: 159-170.
- SZEMERÉNYI, Oswald. 1967. The new look of Indo-European. Reconstruction and typology. *Phonetica* 17: 67-99.
- 1970. *Einführung in die vergleichende Sprachwissenschaft*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- 1996. Introduction to Indo-European Linguistics. Translation from *Einführung in die vergleichende Sprachwissenschaft*; 4th edition, 1990 with additional notes and references. Oxford: Clarendon Press.
- TISCHLER, Johann. 1977-2016. *Hethitisches etymologisches Glossar. Mit Beiträgen von Günter Neumann und Erich Neu*. [= HEG] (IBS 20). Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.
- VERNER, Karl. 1877. Eine Ausnahme der ersten Lautverschiebung. *Zeitschrift für vergleichende Sprachforschung* 23: 97-130.
- ZGUSTA, Ladislav. 1951. La théorie laryngale. *Archiv Oriental* 19: 428-472.