

Ars Glossopoetica

Essays presented to Raymond Armar Brown on occasion of his 80th birthday

January 15, 2019

Edited by Jörg Rhiemeier

joerg_rhiemeier@web.de

Table of contents

Preface *Jörg Rhiemeier*3

Nouns and verbs in a logical language with minimalist phonology *Leonardo Luiz e Castro*5

Comments on Peter Schrijver’s Language Contact and the Origins of the Germanic Languages
Tristan McLeay11

Confessions of a Self-Segregationist *John Quijada*15

Re-creating Old Europe: Conlanging and European Paleolinguistics *Jörg Rhiemeier*21

Preface

Jörg Rhiemeier

Dear Ray (and everybody else who reads this),

this collection of papers was edited to celebrate and commemorate your 80th birthday on January 15, 2019. I have been knowing you since I joined the CONLANG mailing list and embarked on the project which eventually led me to the creation of the Old Albic language in the spring of the year 2000. I have always known you as a very knowledgeable, creative and witty contributor to the CONLANG list, and I am certainly not the only person who knows you this way.

In fact, you always reminded me of the founding father of the diachronic-naturalistic school I confess to be a part of, the great, unsurpassed John Ronald Reuel Tolkien. Like Tolkien, you are an academically trained scholar of philology who makes use of his knowledge and his linguistic creativity in his conlanging projects, and, in your case, in his witty comments on various matters concerning conlangs and the art of conlanging on the CONLANG list, as Tolkien would without doubt have done if he had had such a forum at his disposal, and probably did in the meetings of the Oxford Inklings.

Your own conlanging projects are very diverse, and each is interesting in its own way. It is regrettable that the *briefscript* project never reached the degree of completeness many of us would have liked to see. It is remarkable because you, Ray, did not imitate the many designers of such maximally concise languages who use a huge phoneme inventory in order to be able to form shorter morphemes. Instead, you set off in the opposite direction to reach the same goal: you start with such a *small* phoneme inventory and simple syllable structure that you can use the 26 letters of the Latin alphabet as syllabic characters. This has the advantage that the language is rather easy to pronounce (since only very basic phonemes shared by most languages of our planet are used, and difficult combinations of these basic phonemes are avoided), yet, a high degree of efficiency is reached at least in written form, as any possible sequence of Latin letters yields a pronounceable string of syllables. Even I, who tends to be sceptical of engineered languages in general and prefers diachronic naturalistic fictional human languages, was following this project with great interest and appreciate the beauty of this construction.

Likewise, *TAKE*, your take (no pun intended) on Greek stripped of its inflections analogous to Giuseppe Peano's *Latin sine flexione*, is an ingenious and beautiful language. Everything in this language is very well considered, as can be seen on your web pages which, as with all of your conlangs, beautifully and interestingly document your train of thought behind the decisions you made. It was a bit of a disappointment to me (and others who followed this project) that you dropped the notion of a "Western Hellenistic Alternative Timeline" ("WHAT") in which Alexander the Great lived longer, conquered Italy and added western Europe to the Hellenistic world, resulting in a Europe with Hellenic languages replacing the Romance ones, and Greek as its sole classical language. This world would have provided a great playground to diachronic conlangers where they could build many Hellenic conlangs. However, your decision to abandon the "WHAT" is very understandable: you wished to limit the scope of your project, and feared that the result

would be cheap Hellenic fake-Romance languages made by applying the sound changes of Italian, French, Spanish etc. to ancient Greek. And finally, your decision to drop the “WHAT” does not mean that any of us who is interested in such a project cannot pursue it at their own peril.

Your third project, *Outidic*, is an example of two rather uncommon and interesting types of conlangs. First, it is a *fictional auxiliary language*, i.e. a language that represents the creation of a fictional language inventor who would build an international auxiliary language. This is rarely done: most fictional languages represent a *natural* language of an imaginary world. Second, it is a fictional language set not in some sort of fictional realm but *in the real world*. This is something which *Outidic* has in common with, among others, my own conlang creations such as Old Albic or Roman Germanech, and I decided to admit it to the *League of Lost Languages*, the collaborative framework I set up for fictional languages set in the real world.

And currently, you are working on *Britainese*, a Romance language of Britain. Of course, a Romance language of Britain is something that has been done before, the most prominent example being Andrew Smith’s *Brithenig*. But in *Britainese*, you follow an approach very different from Andrew’s. Instead of applying slightly modified sound changes from the history of Welsh to Vulgar Latin, which is quite easily done and has found many imitators, such as Geoff Eddy’s “Goidelic-Romance” language *Breathanach* and my own “Germano-Romance” *Roman Germanech*, you treat it as what it probably would have been, namely a northern extension of the Romance dialect continuum, and examine and extrapolate trends in the northernmost dialects of Gallo-Romance. This is a rather complex and difficult work, and again, your web pages document your well-pondered decisions. Surely, this language will look rather unspectacular, without such “interesting” bits as the initial mutations found in *Brithenig*, but in my opinion, this lack of spectacle and the much greater plausibility makes *Britainese* the more interesting of the two.

After thus briefly reviewing your contributions to the art of conlanging, I shall address here another field of interest that you and I share. This is the linguistic prehistory of Europe. I have not yet been able to get my hands on your dissertation about Pre-Greek languages, but I know your excellent web pages on the enigmatic *Eteocretan* language of ancient Crete, and we have discussed various questions of European paleolinguistics in private e-mail in a constructive and inspiring way. As you know, my own current conlang projects are based on paleolinguistic research, as I lay out in my contribution to the present *Festschrift*. If one, like me, engages in both linguistic research and conlanging based on that research, one must always be aware which hat one is wearing at the moment, and be very careful to restrict the information flow to *one* direction – from the research on real languages to the work on conlangs, in order to avoid ending up being a crackpot.

I, and all the contributors to this *Festschrift*, whole-heartedly wish you a happy birthday and that you stay healthy, especially mentally healthy, for many years to come!

Braunschweig, Germany, January 15, 2019

Jörg Rhiemeier.

Nouns and verbs in a logical language with minimalist phonology

Leonardo Luiz e Castro
leolucas1980@gmail.com

Abstract

One of the goals of language creators over centuries and places is the “the search for a perfect language”. Creating the perfect language is currently discredited as realist goal, but creating logical languages to investigate hypothesis or serve specific tasks still have its place in the communities of conlangers and artificial intelligence researchers. Here, I take part in the birthday celebration of fellow conlanger Raymond Brown by showing my work on creating a logical language with the simplest syllable structure I can design without losing some other features, such as word boundary recognizability and possibility of rhyming words from different classes.

Introduction

Logical languages are a class of constructed languages (conlangs) designed to convey precise meaning by means of unambiguous grammar and word formation rules. Reducing ambiguity requires giving more information per sentence than it is usually given in natural languages. More information per sentence can be achieved by means of longer sentences or larger phoneme repertoire. Using very long sentences to reduce ambiguity introduces no innovation to the world of natural language. Thus, the chosen solution is usually having more phonemes or avoiding phonotactic constraints. For instance, Ithkuil presently has 45 consonants and 13 vowels (multiplied by 7 possible tones). Lojban has 19 consonants, 6 vowels and 14 possible diphthongs. Lojban and Ithkuil are two of the four conlangs that serve as my main inspirations to create Muai.

The other two inspirations are Esperanto and Toki Pona. Esperanto was the first conlang I ever studied and heard about. What I liked the most in Esperanto was how one can precisely infer the class of any word by its form. Singular nouns end in “o”, adjectives in “a” and so on. However, my habit of seeking criticism of everything I like (or dislike) took me to Justin B. Rye’s criticism of Esperanto. One of his points was against the idea that Esperanto is a simple language, because it have a very permissive phonotactics. The, I discovered Toki Pona, a minimalist language with only nine consonants and five vowels. Toki Pona has no diphthongs and no consonant clusters. More impressively, it does not have voiced-unvoiced contrast. The letter P could equally well be pronounced as an B. Naturally, phonological simplicity hinders having some features of logical languages. So, I challenged myself to create a logical language with simple syllable structure, but without very long words either. That’s how I came to invent Muai.

Muai design goals

Muai may be classified as an engineered language (engelang) with some characteristics of loglangs, auxlangs, and artlangs. It is designed to

- be relatively easy to pronounce (similarly to Toki Pona),
- have morphological and syntactic regularity (similarly to Esperanto), and
- offer the possibility of forming unambiguously parseable sentences (similarly to Lojban).

More specifically, Muai has the following features:

- no stress, length, aspiration, voicing contrasts;
- word-breaking detectability (self-segregating morphology);
- possibility of expressing many level of ambiguity or unambiguity;
- word class inferrable by morphology;
- possibility of rhyming words from different classes;
- simple phonotactics.

Muai nominal-verbal roots

Concerning Muai grammar, this essay focus on the formation of nouns and verb. In Muai, nouns and verbs are derived from the same roots. Let us find what nouns and verbs are derived from the root “tenk-”:

tenk- : root related to the verb *to teach*

Nouns:

tenki : teacher

tenku : pupil

Verbs:

tenko : to teach [a pupil]

tenke : to be taught by [a teacher]

It could be said that -i marks the subject, -u the object, -o the active voice and -u the passive voice. However, “tenki” means “teacher” morphologically, not syntactically. That is, “tenki” refers to a teacher even when that teacher is not teaching. In this sense, “tenki” refers to a *potential* subject of the verb “tenko”. Note the following sentence:

.tenki kuomo numansi.

It means “[the] teacher eats apples”. In that occasion, the potential teacher is not the subject of the verb “to teach”. Obvious as the example might be, it illustrate how morphology and syntax are connected in Muai. The examples below show how it works:

fank- : {constructor} to make, to construct, to mount, to create {construction}

fans- : {agent} to act {action}

fent- : {worker} to work with {job, subject, object}

fink- : {searcher, seeker, wanter} to search, to seek, to want {searched, sought, wanted thing}

hont- : {word, symbol, sign} to represent, to signify {meaning, signifiant, concept}

hunt- : {object} to serve for a subject and an action (or verb) {subject and action}
kanf- : {counter, quantifier} to count, to quantify {counted or quantified thing}
kanp- : {area, volume, space} to be delimited by {boundary, frontier}
konp- : {occupier} to occupy {volume, area, field, extension}
kiaf- : {aggressor, attacker} to attack {victim}
kuom- : {eater, drinker, being capable of intake} to eat or drink {food or beverage}
kunf- : {chooser, specifier} to choose, to specify {chosen or specified thing}
link- : {link} to link {linked things}
lunl- : {thing} to be linked, related to {another thing}
mant- : {experiencer} to experience, to feel {experience, feeling}
ment- : {thinker, rational being} to guess, to think {thought}
miak- : to be the mother of
miek- : {lover, liker, enjoyer} to love, to like, to enjoy {loved, liked or enjoyed thing}
mint- : {student} to study {subject, science branch}
munt- : {knower, wise} to know {knowledge, known fact}
piak- : to be the father of
puom- : {fruit} to be a fruit from {plant}
punk- : {point} to be a point located at {position}
sens- : {event} to happen simultaneously to or at the same place as {event}
sioK- : {player} to play {game}
sonk- : {organizer} to organize {thing to be organized}
tenk- : {teacher} to teach {pupil}
tiak- : to be the generator of
tinf- : {giver, donor} to give something to {receiver, taker}
tinH- : {thing, entity} to have as a predicate {predicate, quality, feature, characteristic}
tint- : {thing, predicate} to be {thing, predicate}
tunf- : {agent of movement} to move {moved thing}

All the roots are structured as Muai nominal-verbal roots of first type (the second type will be presented soon). The possible structural pattern of such roots are the following:

CVNC,

CUAC,

CUVNC,

where

C represents any Muai consonant {k, t, p, h, s, f, l, m, n};

N is a nasal vowel {m, n, N} (homorganic to the subsequent consonant);

V is any vowel {a, e, i, o, u};

A is an open vowel {a, e, o};

U is a close vowel {u, i}.

Muai nominal-verbal group roots

Now, let's go back to our previous example sentence:

.tenki kuomo numansi.

The word “numansi” (apple) is derived from the root “numans-” (to be an apple in some aspect). This root differs from the roots showed before by an additional initial CV syllable, forming:

CV{CVNC},

CV{CUAC},

CV{CUVNC}.

The parts between curly brackets might repeat indefinitely if the final consonant of a part is identified with the initial consonant of the subsequent one. The three possible structure are interchangeable in this repetition process. Thus, the structure of the second type of nominal-verbal root is better described as

CV({CVN, CUA, CUVN}C),

where round brackets identify repeatable structures and the commas divide the options inside the curly brackets. For instance, valid roots are

lelionant- (le + lio + nan + t-),

sukonlank- (su + kon + lan + k-),

hosant- (ho + san + t-).

Having three subsequent CUA parts must be avoided because it might generate a homophone with a compound noun (to be explained later).

The initial CV syllables in the roots above mark a “noun group”. Noun groups are similar to noun classes of Bantu languages and were conceived to mark objects of same nature with the same prefix. For instance, all plant species are marked with “nu-”:

numansi : apple

nuhuosi : rose

nupuani : banana

nusiefi : Schefflera sp.

Naturally, “numansi” is a fruit in the example “tenki kuomo numansi”, not a tree or a plant species. Here we invoke the Muai stated feature of “possibility of expressing many level of ambiguity or unambiguity”. If it can be clearly inferred from context that the teacher eats an apple fruit and not the tree, we can write only “apple species”, which is the simplest form of the word in Muai.

It must be said that these initial CV syllables are not proper prefixes. The word “salinki” is not directed related to “linki”. The former means “Linguistics” and latter means “link” (something or someone that connect things). This feature is useful to make more words available to describe new things. Once we define that initial so- marks names of arts, we can choose “sopinti” as “painting” without checking whether the word “pinti” already exists. This makes viable to name the countless things that can be grouped: plants, animals, sports, nations, etc.

Suffixes and compounds

We have learned that the root “tenk” generates “tenki” (teacher), “tenku” (pupil), “tenko” (to teach) and “tenke” (to be taught). But what about “tenka”?

The vowel “a” was chosen to be the “glue vowel” of Muai syntax. So, “tenka” would form a compound with whatever word follows. For instance, “tenka tinti” is equivalent to “tenkatinti” and means “thing(s) of teacher”, maybe referring to books, chalks and whiteboard markers. Another example below illustrate another use of the root of “nafiemanimanti” (woman):

nafiem- : being female,
nimant- : being human,
nafiem-a-nimant- = **nafiemanimant-** : being woman (human female).

Naturally, one could use only “nafiemi” when being human is made clear by context. Or only “nimanti” when gender doesn’t matter.

In the example above, -a- was placed after a root. If, instead of the root, the complete **word** “nafiemi” is followed by -a-, we still have a compound noun, but now the first word becomes the nucleus:

nafiemi-a-sioku means "female related to games",
nafiemi-a-sioki means "female related to players",
sioku-a-nafiemi means "game related to women",
sioki-a-nafiemi means "player related to women" and
sioku-a-nafiemu means "game related to femininity".

These compounds are pretty vague and should be used when some vagueness is desirable, but Muai offers the resources to make more precise compounds. For instance, if **sioki** means "player", we must have a way of saying "football player". In Muai, we do that by means of -o- :

sioki : player
sefuoti : football
sioki-o-sefuoti : player of football, football player

On the other hand, **sioku** means "game" and we can specify "women's game" by means of -e- :

sioku : game
nafiemi : female, woman
sioku-e-nafie : women's game, lit. "game-played-by-woman"

Briefly, -o- specifies the **sioku** of the **sioki** and -e- specifies the **sioki** of the **sioku**. Similarly, all -i words can be followed by -o- and all -u words can be followed by -e-.

We also have the suffixes “-ai” and “-au”. They are similar to English “-ing” or “-ation” when used to describe actions or processes. The word “tenkai” means “teaching”, while “tenkau” means “learning”. Any of them could be translated as “education” as well, but there is yet another suffix, “-iaiu”, which is more adequate to refer to abstract processes, without focusing on the agent or on the patient. Therefore, “education” is best translated to Muai as “tenkiaiu”.

The compounding rules for -ai words are the same as for -i words. The same is true among words ending in -au and -u. For example,

tenkai-o-nakinti: education of children, child teaching;

tenkau-e-sioku: learning from games, game learning.

Nouns ending in -ai and -au are also useful to make interesting rhymes. Muai has pronouns, specifiers and quantifiers ending in the same way. Muai pronouns end in three vowels (kioi, tiei, lai, siai, etc.) that can be pronounced as triphthongs. Specifiers have structure {CAU}, such as “tei” (“and”, “too”), “lau” (“or”, “alternatively”) and “mai” (interrogation particle) and quantifiers follow the pattern {CVCAU}: “munai” (one), “ninai” (two), etc. All of them might appear at the end of a sentence, where rhyme usually takes place.

Conclusion

The rules described here allows us to form innumerable nouns, verbs and their compounds from nominal-verbal roots. The words formed so are easy to pronounce and morphosyntactically regular. Some suffixes may be pronounced as diphthongs or triphthongs, which allows us to rhyme nouns with words of other classes.

A minimalist phonology makes it harder to increase the vocabulary of a conlang without homophones (normally avoided in loglangs). This problem is attacked by deriving nouns and verbs from the same roots. The rules of compounding are designed to make words more precise when needed and more vague when context is enough to avoid confusion.

Muai has a minimalist phonology in the sense of having the simplest syllable structure I could arrange without losing some goals that are more akin to artlangs. Having simple phonology is a typical objective of auxlangs. Hence, Muai is an engelang at boundary with the auxlang and artlang worlds.

References:

[Ithkuil] http://www.ithkuil.net/01_phonology.html

[Lojban] <https://lojban.github.io/cll/2/2/>

[Lojban] <https://events.ccc.de/congress/2005/fahrplan/attachments/555-sm2005lojban.pdf>

[Toki Pona] <https://tokipona.org/>

[Justin B. Rye] <http://jbr.me.uk/ranto/index.html>

[Muai] <http://leolucas.conlang.org/muai/muai.html>

Comments on Peter Schrijver's *Language Contact and the Origins of the Germanic Languages*

Tristan McLeay

So I've read the first substantive section of a book I mentioned in a thread about Dutch Umlaut *Language Contact and the Origins of the Germanic Languages* by Peter Schrijver. A book about language contact and how it influences languages should be great for conlangers. If that's your main interest, do, consider it. If you're interested in the origin of features of English analysed, your time is probably more honestly served elsewhere. I get to his conworlding later on in my text, but it's spread throughout his work wherever necessary to make a point. (I guess it's not really conworlding, since he's trying to explain the real world, but I'm writing for a specific audience.)

The first section is about the development of English. There are certain features of English, known to most who are interested in English as a Germanic language, which distinguish it from its continental siblings, but which relate to Celtic languages, for instance the present progressive and do-reinforcement. From time to time, a linguist wants to identify these as being a Celtic substrate feature; and another linguist claims they cannot plausibly be so described since (a) they are seen on the continent in Dutch and German dialects and (b) they are late developments not seen in OE. Schrijver provides a reason to believe they could have been present in OE days, and adds to the list of borrowed features some which are certainly known from OE. The more interesting part of his text, and the bit which however speculatively undergirds his entire argument, is his painting of the languages used in Britain and Ireland from the beginning of the so-called first century until his interest wanes sometime during the middle ages.

I don't find the counterarguing linguists' arguments particularly convincing, since I've only seen summaries of them that make them sound like they're arguing simultaneously that the feature is old since it's shared and that it's new, since it's not visible in the earliest strata. To this point I've formed no particular opinion. Schrijver argues that they were transmitted from Celtic to English in the periphery in communities that remained Celtic speaking for several generations, and that the speakers who initially used these features were speaking what was regarded as bad English, and it required the collapse of the literary standard for them to enter into use. It could be.

He shockingly attributes to Celtic substrate effects English i and a/u umlaut. This is literally incredible. I don't think you can believe it; I certainly can't. His argument goes something like this:

1. The sound system of North Sea Germanic c. 400 has the vowels /i: e: (æ:) u: o: ɔ: i e a u [o] ai au iu eu/ whereas 300 years later, Old English dialects have /i: e: æ: y: ɔ: a: u: o: i e æ y ø a u o i: o e: o e: a i o e o ea/ - these are orthographical forms with added colons (in his text, macros) and the letter ø thrown in for good measure.

1a. This is a big change and it cannot be justified. The Germanic vowel system was perfectly unremarkable and should not have shifted that way. The most that can be said for it is that it was

common in North West Europe for similar shifts to occur, but that merely provides evidence of a substrate.

1b. However, dialects vary considerably in whether a word has a diphthong or a corresponding monophthong. For instance, Anglian has *sce:p* and *edor*, whereas West Saxon has *sce:ap* and *edor*.

=> Therefore, there was pressure to adopt the vowel system, but exactly how it was adopted is not determined by that pressure (i.e. it was external).

2. Germanic did not have Umlaut, since Gothic doesn't have it and West Dutch only weakly has it (I hear that as saying West Dutch has Umlaut, but he's trying to make you read it as West Dutch doesn't have Umlaut); moreover, each language which does have Umlaut varies in time frame and details.

=> Therefore, the process is not inherited; if there was so much as a pre-existing subphonemic process, we still need to explain why it became phonemic.

2b. Moreover, there was no unity of the Germanic speakers that could have allowed the process to begin in, say, Denmark and then spread throughout the whole territory after the language had divided.

=> Therefore, the process needs independent explanations.

2c. The details are unique in English compared to other Germanic language. This requires explanation.

3. On the other hand, the differences between English and Irish processes are more apparent than real. In English, vowels anticipate following vowels; in Irish, consonants anticipate following vowels - but the anticipation in both cases affects the other segments too; and it is merely a subsequent development that results in the distinction. In details too, there are similarities. For instance, in both languages *ACuCi* -> *ECiCi* where *A* and *E* are a back/front vowel pairs, *C* is some valid consonant or cluster and *u* and *i* are taken literally.

I've probably done a terrible job of explaining his argument; but I really think the argument is terrible. The whole thing requires us to accept assertion after assertion. He creates four criteria for testing if a change is due to a substratal influence, but the only time he tests them they fail to detect a known substrate. There's simply no evidence they work. And if they did work, he applies them so poorly. For instance, rule 2 is that the "initial state of the sound system does not pre-program the language to undergo those changes". He denies that a system in which */bo:k+iz/* is pronounced [bø:kiz] preprograms the output */bøk/* since, effectively, there's no guarantee that it will produce that output on a certain timeframe in a certain system. You can hardly believe a person accepting that rule 2 is prima facie plausible would think that's a fair application.

He also has a dreadfully dull view of language change, as though a change can only spread throughout a language community as far as a previous change, and that change necessarily fully affects every candidate in a language as though there is some iron-clad law. Moreover, only normal changes can happen. Anything too adventurous requires the help of an external language. He really does argue this: "It seems safe to say that the linguist's toolbox of general structural principles, however sophisticated, does not contain the gear necessary to turn the vowel system

of pre-settlement North Sea Germanic into that of Old English." And he follows that up with an argument that I keep reading as preventing umlaut from developing anywhere at any time without the help of another language that had it first - I guess we know how Adam and Eve spoke. I don't believe he truly thinks that of course; but I can't see what it is in his text that could save me from having to make that conclusion.

Now, to describe his linguistic landscape. I actually like this bit. Since I don't believe Umlaut needs the help of a substrate, I don't think it does any work; but it's still quite plausible, or at least fun, even in the absence of need. I have converted it, as best as I can, into chronological order. The order he presents it in honestly doesn't appeal to me; it was hard to follow his logic at times because it seemed like he was contradicting himself. This ordering makes his argument more coherent in my view.

When the Romans began to invade Britain, Celtic was spoken throughout the relevant parts of Britain and an unknown language was used in Ireland. The Celtic that was used was the ancestor of Welsh, Cornish and Breton as well as Irish and Scots Gaelic and related languages. This language already had processes by which later vowels are anticipated in earlier consonants and vowels. If it was uniformly P- or Q-Celtic, then it was uniformly P-Celtic (that is, PIE $k^w \rightarrow p$), but it is not necessary that it was uniform in this respect.

Some Celtic refugees fled to Ireland, and Celtic began to become the linguistic majority over the next few centuries. It underwent many changes, but they didn't substantially affect the vowel anticipation. If British Celtic had been uniformly P Celtic, then the pre-Irish language of Ireland had $/k^w/$ but not $[p]$, and speakers adopting Celtic used $/k^w/$ to approximate the missing stop. This certainly happened on the Ireland, since Latin *pa:scha* \rightarrow OIrish *ca:sca*.

By the time the Romans left Britain, Britain had multiple languages present. In the highlands, we had a fairly traditional form of Celtic. In the lowlands, we had Latin, Latin influenced by Celtic and Celtic influenced by Latin. The Celtic influenced by Latin shows more affinity to Gaulish - for instance, in sharing the Latin $/i: e: u: o: \rightarrow ei ie ou ua/$ shift. Those who spoke Latin and Latin-influenced Celtic were relatively wealthy. Those who spoke Celtic and Celtic-influenced Latin were not.

The Germanic invasions occurred. The wealthy were forced to seek refuge in the highlands, and did so in sufficient numbers to cause Highland British Celtic to become Latin influenced. It is from this form of language that Welsh, Cornish and Breton are derived.

Meanwhile, the less well-connected remained in the lowlands and were largely permitted to integrate, as long as they were prepared to. They spoke the more pure, Irish-like form of Celtic. And as they shifted to Germanic, they did so with a strong Celtic accent, and brought across their tendency to influence earlier consonants and vowels. This Germanic-with-a-Celtic-accent is therefore the basis of Old English. Anyone who was looking for Welsh to find Celtic in Old English was looking in the wrong place, since Welsh is Celtic with a Latin accent, and English is Germanic with a Celtic accent.

Nothing about this strikes me as being ridiculously implausible. But it doesn't seem to do the work it was meant to once you call Umlaut internally driven. And I find some bits a bit difficult. Irish in particular seems to both change a lot and not change at all in his telling. It is a good guide

to British Celtic and it is a substrate influenced mixture. Perhaps I just need to read that section again to get a better understanding of what he's arguing.

The book is intended for beginners to read, so it's not substantially inaccessible. No need to fear it. It's the content, not the presentation, that is bad.

In summary, I don't like this work so far. It feels more like he's come up with a great explanation, and then sought to squeeze something - anything - into it, however poorly it might fit. Conlangers might enjoy his conworlding which really inspires me.

CONFESSIONS OF A SELF-SEGREGATIONIST

By John Quijada

I am guilty of self-segregation. That is, I've created a self-segregating morphology for my conlang. Here follows my sordid tale...

So I was perusing the website of scholar and fellow conlanger Ray Brown, and was intrigued to discover that at one point in his eclectically varied career both within and outside academia, Ray went through a phase during which he pursued the idea of a self-segregating morphology for a logical language. For readers unaware of the term, a self-segregating morphology refers to a language's linear sequence of phonemes being automatically parsable into discrete morphemes by any listener (or computer) familiar with the parsing rules, without having to be necessarily familiar with the grammar of the language itself.

Based on Ray's description of his efforts, he was inspired by various postings on the Conlang Mailing List from the late 1990s and early aughts, focused solely on possible phonological structures for sequences of -CV- syllables by Jeff Prothero, referred to as Plan B or "Bee", and subsequently, Plan C ("Cee") by Jacques Guy. Ray worked on improving and expanding these schemes and came up with a system he called Plan D ("Dee"). Other conlangers who have worked on the same problem (Gary Shannon, John Cowan, and Jörg Rhiemeier) contributed ideas to Ray's work. Common to all of these systems is a fundamental binary structure, easily convertible to binary notation for use by computers.

Ray ultimately abandoned his work, believing that attempts to improve upon Plan B ultimately must collapse in "kludgeyness" and, if I understand correctly, that the constraints of the system would demand it be paired with an oligosynthetic syntax which would likely be insufficiently logical to qualify as a logical language à la Loglan. Additionally, he considered such binary-based systems to be too computer-centric and insufficiently anthro-centric, which I interpret as meaning they were not easily learnable/usable by would-be language speakers.

Nevertheless, the pursuit of a self-segregating morphology remains worthwhile for those of us who profess to be engelang. And so it is that I find myself currently involved in a engelang project requiring a self-segregating morphology. In my case, the motive is not simply theoretical, nor focused on parsability by computers, but rather on actual, practical parsability by human beings wanting to learn to speak the language.

I've Been Forced Into It...

By way of background, since presenting my constructed language Ithkuil in late 2004, I have watched with surprise, humility, and a sense of surreality the notoriety that has come to surround my work well beyond the niche audience of fellow conlangers. Yet, at no time during Ithkuil's creation did I ever expect anyone to seriously attempt learning to speak or write in the language. I've always considered it merely a philosophical construct for contemplation; demonstrating how

a human language might function given no limits on human working memory capacity (to handle the enormous amount of overt morphological information the grammar requires), or cognitive flexibility (to manipulate such a complex morphology in real-time). That is why, when designing it, I deliberately ignored issues of “learnability”, adherence to human linguistic universals, and principles of “natural” language design, not to mention worrying about a phonology and morpho-phonology allowing ease-of-pronunciation or systematic memorization. All I was really interested in were the philosophical design goals of greater overt expressiveness and efficiency.

Nevertheless, the online community of Ithkuil fans have finally convinced me there exists a number of persons genuinely devoted to learning and using a language with Ithkuil’s design goals and capacity for expression, but one more easily learnable. So, as of Autumn 2018, I have succumbed to the pressure and begun work on a new, as-yet-unnamed successor language to Ithkuil that preserves its general morphology and lexico-semantic principles, while drastically altering the morpho-phonology to be more agglutinative, systematic, and modular.

As for establishing a self-segregating morphology for this new language, the problem is three-fold:

- (1) Inter-lexical parsability: the ability for a listener to unambiguously distinguish word boundaries within a sentence and between sentences.
- (2) Intra-lexical parsability: the ability for a listener to unambiguously distinguish the individual morphemes within a word.
- (3) Accomplishing the above two goals while dealing with a complex phonotactic pattern: word-initial C(C)(C)(C)V- onsets, intervocalic -C(C)(C)(C)(C)- conjuncts, and word-final -V(C)(C)(C)(C) offsets.

Inter-lexical parsability

Focusing solely on problem No. 1 above, inter-lexical parsability was fairly straightforward in Ithkuil utilizing tone, where each word of the language carries a tone contour beginning with neutral mid-level tone then switching to a different, non-mid tone beginning with the stressed syllable of the word (usually the penultimate or ultimate syllable). Thus, the return to a syllable with mid-tone signaled the beginning of a new word.

However, the fans have demanded the new forthcoming language be a non-tone language, for easier learnability. As a result, I’ve come up with an entirely new approach utilizing two specialized consonants in conjunction with syllabic stress. The two consonants in question are the glottalic consonants /ʔ/ (written as ’) and /h/. By “specialized” I mean highly constrained in their distribution. The rules are as follows:

- The consonant /h/ occurs only in word-initial position. It may be followed by a vowel or diphthong, or can be the first member of the conjuncts /hl/, /hr/, /hm/, /hn/, or /hw/, these being followed by a vocalic-form.

- The glottal stop occurs only in word-initial position followed by a vowel ('V-), or in syllable-final position preceded by a vowel and followed by a consonant (-V'C-). It cannot occur inter-vocally nor in word-final position.
- No word begins with a vowel. All words seeming to begin with a vowel are preceded by a glottal stop. Learners must be consciously aware to pronounce this glottal stop, as it is crucial for parsing purposes. At the same time, speakers must avoid inserting a glottal stop into disyllabic vocalic conjuncts such as /ea/, /oa/, /ae/, etc. (i.e., they must not be pronounced as /eʔa/, /oʔa/, /aʔe/).
- All words contain one stressed syllable; any other syllables of the same word are unstressed.
- If a word does not have word-initial stress, it must begin with a glottal-stop /ʔ/ or /h/.

Based on the above , the following two rules allow a listener to parse word boundaries in an utterance:

- Any syllable beginning with a glottal-stop or /h/ signals the beginning of a new word.
- The first stressed syllable following a word-initial 'V- or h- (which may be the very same syllable containing the 'V- or h-) constitutes the stressed syllable for that word. Any subsequent stressed syllable before another 'V- or h- signals the beginning of the next word.

As for parsability between sentences, I have decided to “cop out” and rely solely on a speaker’s (i.e., a learner’s) knowledge of the new language’s morpho-syntax to tell him/her that a new sentence has begun. This is rather straightforward for four reasons: (1) The language is verb-initial, (2) the sentence-initial verb will always be the verb associated with the main clause of the sentence, (3) the morpho-phonological structure of verbs is distinct from other word classes, and (4) verbs within any subordinate clause of the sentence are overtly marked as such. Consequently, the learner/speaker can always recognize a main-clause verb and know a new sentence has begun. In the infrequent case where a sentence does not begin with a main-clause verb (certain specialized sentences involving atypical topic-focus constructions, as well as single-word sentences consisting of interjection-like words), the phrase will be prefaced by a specialized monosyllabic particle beginning with /h/ signaling the start of a new sentence.

Intra-lexical parsability and phonotactic complexity

Problem No. 2 above, the parsability of morphemes within words, becomes tricky in this new language, as it was for Ithkuil, given the extremely large number of morphemes shown within a word (verbs in the language inflect for two dozen different morphological categories). The problem is exacerbated by the extreme phonotactic possibilities permissible, as detailed in problem No. 3 above.

The solution I devised for the 2011 version of Ithkuil will also serve for this new language as well. It consists of imposing a “slot”-based morpho-phonology for all words in the language, an idea I borrowed from Northwest Caucasian languages, specifically Abkhaz. By utilizing a slot-based word-structure, the sequential ordering of morphemes in a word becomes very strict, thus

allowing each slot to have its own internal morpho-phonological constraints as to what phonemic combinations it can contain. By ensuring these slot-specific constraints are distinct from those of neighboring slots, the learner/speaker/computer can distinguish which slots of the word are filled or not, as well as unambiguously determine the start and end of each slot-bound morpheme.

The slot structure illustrated below is currently tentative and may still change, however it will suffice to illustrate the principles involved. (The arcane notational scheme is explained on the next page.)

I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV
C _V	V _K	'C _E	wVC _S	(w)V _A	C _R	V _R	C _M	V _K	C _D	C _A	VC _S	V _C	[stress]

Example of a word with all slots filled:

h	ae	'n	woʃ	oa	kr	ou	ln	au	'	l	ömky	uô	[ultimate]
---	----	----	-----	----	----	----	----	----	---	---	------	----	------------

hae'nwoʃoakroulnau'lömkyuô [ha.ɛʔ.nwo.θo.a.kroụl.nauʔ.lœm.kju.ɫ̥]

TRANSLATION: '...instead of hoping that one day I might benefit by officially taking part in entering within and ascending in a corkscrew-like trajectory to reach the very center'

At first blush, parsing the above fully-inflected verb looks like it would be tantamount to untying the Gordian knot. But by understanding the morpho-phonological constraints for each slot, it turns out the word's morphological structure is completely transparent. This is accomplished in three ways:

- (1) by using syllable-initial semi-vowels /w/ and /j/ in certain slots as morpheme boundary indicators
- (2) by distinguishing vocalic morphemes by three types: single vowel (e.g., /a/, /e/, /o/) vs. falling diphthongs (e.g., /aj/, /eu/, /oj/) vs. disyllabic conjuncts (e.g., /i.a/, /u.e/, /o.a/)
- (3) by various constraints on which consonants can or cannot appear in certain positions within certain slots

Here follows an analysis of each individual slot as to its function and its distinct phonotactic signature:

SLOT	LABEL	FUNCTION (i.e., category shown)	PHONOLOGICAL STRUCTURE
I	C _V	Shows a category called Version, having 2 possible values	Two values possible: either /ʔ/ or /h/

II	V_K	A slot showing any one of 3 different categories (Illocution, Valence, or Effect) via a vocalic infix.	To distinguish which of the 3 categories is being shown, each has a different phonotactic structure: (1) Illocution = single vowel, (2) Valence = a falling diphthong, (3) Effect = a disyllabic vocalic conjunct
III	C_E	An optional slot holding a single consonant showing the category of Effect	Single consonant preceded by a glottal stop (which keeps it from being misinterpreted as the C _R root consonant in Slot V.)
IV	VC_S	An optional slot used to shown any derivational affix of the form -VC(C)(C). In this slot, the affix must always be preceded by /w/ (or /j/)	All derivational affixes are of the form -VC(C)(C). The initial /w/ or /j/ preceding the affix ensures it is not misinterpreted as the V _A plus C _R forms of following slots V and VI .The consonantal portion of the affix cannot consist of or begin with /w/, /j/, /h/ or /ʔ/.
V	V_A	A vocalic infix indicating Aspect	A single vocalic form (single vowel, diphthong, or disyllabic vocalic conjunct). If Slots III and/or IV are empty, this vocalic form must be preceded by /ʔw/ so that it does not juxtapose directly with Slot II.
VI	C_R	The lexico-semantic root: C(C)(C)(C)(C)	Consists of one to five juxtaposed consonants (the phonotactic constraints on what consonants can be juxtaposed are irrelevant to this presentation). Cannot consist of or begin with /w/, /j/, /h/ or /ʔ/.
VII	V_R	A vocalic infix identifying which stem of the root it is	A single vowel, diphthong, or disyllabic vocalic conjunct.
VIII	C_M	A bi-consonantal infix indicating Mood; the default Mood value is zero-marked if the following Slot IX is zero-marked as well.	This infix must end in a nasal consonant, to distinguish it from the CA morpheme in Slot XI (which can never end in a nasal consonant).
IX	V_K	Same infix as in Slot II above; this is its alternate or additional position if not placed in Slot II or if Slot II is already filled by a different V _K value	(See info for Slot II above.) May be zero-marked if the preceding Slot VIII is also zero-marked.
X	C_D	Shows a category called Designation, a 2-valued category	There are 2 possible Designations; the first is zero-marked, the other is marked by placing a glottal stop /ʔ/ in this Slot.
XI	C_A	An agglutinative consonantal complex of the form C(C)(C)(C) showing five different morphological categories (Configuration, Affiliation, Extension, Perspective, and Essence)	Consists of one to four juxtaposed consonants (the phonotactic constraints on what consonants can be juxtaposed are irrelevant to this presentation). Cannot consist of or begin with /w/, /j/, /h/ or /ʔ/ and cannot end with a nasal consonant (to distinguish it from Slot VIII).
XII	VC_S	Same infix as Slot IV above	See info for Slot IV above. When used in this slot, the affix is not preceded by /w/ or /j/.
XIII	V_C	A vocalic infix showing the categories of Case (for nouns) or Frame (for verbs)	A single vocalic form (single vowel, diphthong, or disyllabic vocalic conjunct).
XIV	[stress]	Shows a category called Context	There are 4 Contexts, each shown by a different stress pattern: word-initial, ultimate, penultimate, antepenultimate. The latter two patterns may require epenthetic syllables to be added to the word to provide a sufficient number of syllables to distinguish them from word-initial stress (Two- and three-syllable words with initial stress are considered to have word-initial stress, not penultimate or antepenultimate stress).

One additional rule regarding the above structure which affects parsability: If the word has word-initial stress, then Slots I through III may be elided if they have their default values (or empty in the case of Slot III). Such elision is what allows word-initial stress to play a role in the inter-lexical parsing rules previously described.

Binary Versus “Messy”

The intent of this paper has been to illustrate the implementation of a self-segregating morphology via a methodology other than binary exclusivity. This methodology is based on the practical, messy, somewhat arbitrary (but effective) distribution of complex phonotactic constraints, both between words and within a word.

Personally, I find it interesting (especially having studied computer programmer myself in my youth) that human beings, for all of our bilateral symmetry and black-and-white idealism, function better when language structures are based on such messy, multivalent distributions rather than on binary schemes. More evidence, methinks, that human gray matter doesn't operate on binary principles and that attempts to build HAL-9000-level artificial intelligence based on ultimately binary, algorithmic structures are doomed to failure. But that is fodder for another discussion, another time....

Happy 80th Birthday, Ray!

Re-creating Old Europe: Conlanging and European Paleolinguistics

Jörg Rhiemeier

1. Introduction

My main conlang *Old Albic*, the language of my “Elves”, is intended as a language that was once spoken in the British Isles, namely before the Celtic languages, the oldest stratum known to us in this part of the world, spread there. There are a few other languages in this setting, the *League of Lost Languages*, which likewise represent lost linguistic lineages, such as the various *Hesperic* languages on the continent, the *Razaric* languages of the “Dwarves” of Britain, and others.

Making these languages means making languages that can be believed to once having been spoken in the real world, but being lost in time. Thus, I entered research into what (little) is known about languages that once were spoken in Europe, but disappeared before they could be laid down in written form. In my model, there are several linguistic strata that preceded the Indo-European languages that have been spoken in Europe in historical times, and are still spoken today.

What we know about these languages (which shall be referred to as “Paleo-European” here) is understandably very little (apart, of course, those which are still spoken today, namely Basque and the three Caucasian families). In the Mediterranean, some pre-Indo-European languages such as Iberian, Etruscan or Minoan were still used when people adopted the art of writing in these countries, but for the most part, these inscriptions are still highly enigmatic. Etruscan is partly intelligible, but many words and parts of the grammar are mysterious (it tells a lot that the English and the German Wikipedia, for instance, present different case paradigms, following the opinions of different scholars); the Iberian script can be read, but the words not understood; and the Minoan inscriptions (the name “Minoan” is of course a designation imposed on the language by modern scholars; we do not know what these people called themselves, but judging from Egyptian and other sources, it may have been something like ***Kafti**) still await full decipherment, though some progress has been made. Eteocretan, a later non-Greek language of Crete (studied by our jubilar), is written in Greek letters, but not understood, and we do not know whether it is a descendant of Minoan or not; it could have been brought to the island after the fall of the Minoan civilization.

North of the Alps, things look much bleaker still. The art of writing reached these parts of the world only at a time when all the Paleo-European languages had already been ousted by Indo-European and Uralic languages, though there are few inscriptions (in known alphabets: Latin, Cyrillic, Runic, Ogham) which do not seem to make sense in any of the known languages; usually, these inscriptions are considered cryptograms, i.e. enciphered texts in known languages, but there is of course a faint possibility that at least some of them are actual examples of lost languages. Alas, we simply do not know.

Thus, other roads to knowledge have to be sought. One such road are *stratum loanwords*: words in known languages which have no known etymology, and may have originated in languages that have been superseded by the languages in which they are found today. There are, indeed, quite a few such words in the Indo-European languages of Europe, such as Germanic, Celtic and Greek, and much scholarly literature has been written about these words. Of course, it is difficult to decide whether a word that is found, for instance, in Germanic but no other branch of Indo-European, is really a loanword from a lost language of central Europe, or simply a word inherited from Proto-Indo-European (PIE) that for some reason was lost in all other branches. Sometimes, though, such words have structures that make it unlikely that they are inherited from PIE. For instance, we know that the phonemes **a* and **b* were of limited distribution in that language: **a* probably was just an allophone of **e* occurring next to a specific consonant (the “laryngeal”, whatever its precise phonetic value, transcribed as **h₂* – there also were **h₁*, which did not affect vowel qualities, and **h₃*, which turned **e* into **o* – all three lost in all IE languages except the Anatolian ones such as Hittite, and the earliest stages of Vedic and Avestan where seeming metric irregularities point at them still being present when the earliest parts of the Rigveda and the Avesta were composed, see Fortson 2010:212 on Vedic and *ibid.*:232 on Avestan) which only in a late stage acquired the status of an independent phoneme, while **b* is so rare and not found in reliable etymologies that many scholars believe that it was missing in the language. So, if a word in an Indo-European language requires a PIE etymology that contains **b*, or an **a* without a **h₂* anywhere near, it is probably a substratum loanword.

A specific kind of substratum loanwords are semantically opaque *geographic names*. Many geographic objects bear names that do not seem to make sense in any of the known languages that were ever spoken in the relevant location. This is an especially common condition with river names, and enough ink has been shed to give a name to this problem: the German linguist Hans Krahe coined the term “Old European Hydronymy” (German: *alteuropäische Hydronymie*) in the middle of the 20th century (Krahe 1963) for a network of seemingly recurrent river names spanning most of Central and Western Europe. Alas, there is a problem with this approach: we do not know the original meanings of these names. We may know that someone named a river in what is now Germany, **Wisara* (now *Weser*), but which features of the river in question led them to name it that way, and not by one of a few hundred other names that occur in the roster of the Old European Hydronymy?

Accordingly, different hypotheses have been developed here. Krahe assumed that the Old European river names were from a yet undifferentiated dialect of PIE which was ancestral to the IE languages of Western and Central Europe; his colleague Theo Vennemann, in contrast, has proposed that the names are from a language family he names *Vasconic*, i.e. a family of which Basque is the sole surviving member (Vennemann 2003). Of course, they cannot both be right! And finally, the whole “Old European Hydronymy” may just be the linguistic equivalent of constellations or ley lines – a seeming pattern falling out of the sheer amount of data (after all, there are many *thousands* of watercourses in the subject area, so one may expect names to re-occur in places), but not meaning *anything*.

Only in a few cases we can catch a glimpse of meaning when a particular element in geographic names shows a correlation to a property of the relevant features. One nice example, not part of

the “Old European Hydronymy” but perhaps from the same stratum, is the element ***hal-** occurring in names such as **Halle**, **Hallstatt**, **Hallein**, **Schwäbisch Hall** or **Bad Reichenhall** (the spelling **ll** is just a German orthographic convention indicating that the preceding vowel is short, and thus does not require a geminate). What all these places have in common is that they are, or once have been, salt production sites. Not all Central European salt production sites have ***hal-** names, of course, but where there is a ***hal-** name, salt is or has been produced, and the correlation between ***hal-** and salt production is significant enough to conclude that ***hal-**, in whichever language it may come from, apparently meant ‘salt’. We have grasped a word of a vanished language.

Only if we have determined meanings in ways like this, we can attempt to identify the language and its relationships. Our ***hal-** element, for instance, is reminiscent of the ‘salt’ word in Greek (**hals**) and in Welsh (**halen**). Neither of the two, however, were spoken in Central Europe in prehistoric times, and in Welsh, the initial *h* is the result of a rather recent (less than 2,000 years ago) sound shift (thus not yet in place in Proto-Celtic, which was once spoken in some of these places); before that, the initial consonant was ***s**. Yet, it is possible that the word is related to PIE ***sh₂el-** ‘salt’, and that the language is a sister group of PIE. However, as long as we have only one word, such resemblances are meaningless and we may be dealing merely with a chance resemblance. We need to explore more meanings of old names like this. That is a great endeavour which to my knowledge has not been tackled yet.

What we are essentially left with, thus, are loanwords in the lexicons of the attested languages. If a European IE language has a word for a concept which defies all attempts at finding a PIE etymology, we can at least guess that it originated in a Paleo-European language, where it meant something similar as the attested meaning. In the Germanic and Celtic languages of western Europe, such words number in the hundreds – they do not give us a full picture of a language, but at least a sizable basic stock of lexemes on which we can build a conlang.

Of course, a language is more than just words. So what do we know about the lost languages’ *grammars*? Not much, but a bit more than nothing. It has, for instance, been conjectured that the divergent grammatical structure of the Insular Celtic languages, with their radically right-branching word order, their initial mutations and other features alien even to the Continental Celtic languages, reflects the influence of a substratum language which was characterized by a similar syntax. Another example is the phenomenon for which I have coined the term *alpha mobile*: some possible substratum loanwords seem to occur in pairs, one with an initial ***a** (like German **Amsel** ‘blackbird’ < ***a-masl-**), one without (like Latin **merula** ‘blackbird’ < ***masl-**). This seems to point at ***a** having been some sort of prefix or proclitic in the source language, perhaps a definite article (borrowings with article are a thing, as English words of Arabic origin such as **alchemy** or **algebra** show, where **al-** is from the Arabic definite article, **ʔal**).

2. Albic, Hesperic and Para-Indo-European

Old Albic is a member of the *Hesperic* language family, which consists of several languages spread out across western and central Europe. The idea behind Hesperic is that of a sister group of Indo-European that was spoken in Europe before the spread of Indo-European proper, and left traces in geographic names such as the Old European Hydronymy and in loanwords in the Celtic, Italic,

Germanic and Balto-Slavic languages. This hypothesis is of course highly speculative, and would probably not last long if I was to submit it to serious academic discussion. But that is not really my intention; it serves merely as a foundation of a conlang project.

My project of a real-world “Elvish” language did not start here, though. The impetus came from a Tolkien fan fiction story, *Home Eleven* by Martin Baker (Baker s.d., found in early 2000), featuring (Tolkienian) Elves in the modern world. I started working out the language such Elves may speak, starting from Sindarin. That project was named “*Nur-ellen*”, meaning ‘Low Elvish’ in the language itself. In hindsight, the language was not made particularly well, and while I was working on it, my ideas of the nature of the Elves changed. The thing became divorced from Tolkien’s legendarium, and the “Elves” became a human ethnic group in the British Isles of the Bronze Age. This, in turn, led me to abandon the Sindarin-derived language and build something new on my own.

The idea that the language of my “Elves” should be related to Indo-European arose when I read the book *Indo-European and the Indo-Europeans* by T. V. Gamkrelidze and V. V. Ivanov (Gamkrelidze/Ivanov 1995). These two authors are famous mainly for their “glottalic theory”, according to which the class of stops traditionally reconstructed as voiced (e.g., ***d**) actually were glottalized (e.g., ***t̥**). But this, though I worked it into my framework (I do not think this makes sense at the level of PIE proper, but it may have held for an earlier stage of the language, which would be sufficient to explain what this theory was meant to explain), was not the main point of interest to me. More interestingly, Gamkrelidze and Ivanov reconstruct PIE as an active-stative language, i.e. one in which intransitive subjects are grammatically marked like transitive subjects when they are agents (as in *The child sings*), but like transitive objects when they are not (as in *The ball lies in the playground*). This was what I wanted to do in my “Elvish” language. While I feel that some parts of Gamkrelidze’s and Ivanov’s theory are erroneous, I found enough bits in PIE that point at such a direction.

So I decided that my “Elvish” language was a branch of Indo-European which branched off at a very early stage, even before Anatolian, at a stage when PIE still was an active-stative language, and before the emergence of ablaut. The language would have, at an early “Proto-Albic” stage, a vowel system of only three vowel phonemes, ***a**, ***i** and ***u**, of which ***a** was by far the most frequent. (Old Albic proper would have more vowels, due to the workings of an umlaut system.) It fit things that the Old European Hydronymy seems to have the same vowel system. These vowels correspond to the PIE ones as follows:

Proto-Albic	PIE
*a	*e ~ *o ~ ∅
*i	*ei ~ *oi ~ *i
*u	*eu ~ *ou ~ *u

The consonants would be related to the IE ones by a sound correspondence system similar to the Germanic and Armenian stop shifts:

Proto-Albic	PIE
aspirated	voiceless
neutral	voiced
voiced	breathy-voiced

The idea was that Albic had retained the Pre-PIE “glottalic” system but with loss of glottalization, and the old voiceless stops having acquired aspiration instead. (In Old Albic proper, these aspirated stops became fricatives.) This had a nice side effect which I could exploit in building vocabulary. I could use Germanic words with questionable etymology, without having to decide whether they were inherited from PIE and were genuinely cognate in Albic, or borrowed into Germanic from a language related to Old Albic – both pathways would yield the same sound correspondences.

While I was working on this, the question arose whether Albic was to be a part of something bigger (apart from the distant connection to IE) or not. At first, I entertained the notion of a small, self-contained language family, somewhat like Kartvelian. But there were reasons to make it into something larger. It would be plausible if there were a few further languages on the continent, after all the “Elves” came from there, and the Old European Hydronymy was all over the place from the British Isles to the Iberian Peninsula and to the Baltic Sea. Also, the prospect of making a *big* family, something like Geoff Eddy’s Sunovian or Mark Rosenfelder’s Eastern, something with a similar diversity as Indo-European or Uralic, was attractive, so I decided that Albic would be a branch of a larger family, *Hesperic*.

The theory how Proto-Hesperic was related to Indo-European and who spoke it changed while I was progressing. The initial model was inspired by the non-fiction book, *Noah’s Flood* by William Ryan and Walter Pitman (Ryan/Pitman 1998). In this book, the authors, two marine geologists, lay out the hypothesis that the Black Sea Basin suffered a catastrophic inundation about 5500 BC, when the rising ocean burst through the Bosphorus and created the Black Sea as we now know it; before this cataclysm, the basin would have held a freshwater lake at a much lower level. According to the book, both the *Linearbandkeramik* (LBK) culture, the first Neolithic culture of Central Europe, and the Proto-Indo-European language community would have been founded by refugees from the flood originating from the northern shore of the lake. I decided that Hesperic was the LBK language, later spread across Western Europe by the Copper Age Bell-Beaker Culture while IE was encroaching on the LBK homeland from the east.

It soon turned out, though, that there were problems with this idea. First, it seems as if Ryan and Pitman had misinterpreted their data, and there simply was no cataclysm in the Black Sea Basin, at least not at a date convenient to such a hypothesis. Second, it turned out that LBK and the *Yamnaya* culture widely held to be identified with the PIE community, were archaeologically not closely related; instead, LBK originated, via the *Starčevo-Körös-Criş* culture of the Lower Danube region around 6000 BC, from Anatolia, while the *Sredny Stog* culture, which was the immediate precursor of the *Yamnaya*, formed when the autochthonous Dniepr-Donets culture merged with,

or was taken over by, the *Khvalynsk* culture from the Volga region. Also, the two communities were genetically not particularly similar. Third, the Bell-Beaker culture was not a lineal descendant of the LBK culture; rather, it appears to have been related to Yamnaya, though there is also evidence pointing at an origin from the Iberian Peninsula.

So I needed a new framework. My research into prehistoric migrations in Europe led me to assume an intermediate layer between the LBK culture and the spread of PIE proper. In the new model, Proto-Hesperic was the language of a group of very early Indo-Europeans, branching off about 4500 BC and migrating into Central Europe, bringing some innovations such as dairy farming and the ox-drawn plough (but not yet the wheel, nor the domesticated horse, nor metals) into the region (while the Indo-Europeans proper, who came about 1,000 years later, had wheel, horse and copper), which greatly improved their economic base and enabled them to settle on poorer soils which the LBK people had avoided. There is indeed some evidence pointing at such a layer, though it is admittedly weak. A later offshoot of this would have been the Bell-Beaker people. I even found a speculative solution on the Bell-Beaker origin problem. In my model, the *people* bearing this culture originated in Central Europe, but the *material culture* was strongly shaped by innovations from the Iberian Peninsula (most notoriously, copper metallurgy), similar to the way the culture of modern North America is strongly shaped by innovations from California with its Hollywood and Silicon Valley – a region that was added to the Anglo-American cultural sphere late, but attained a strong influence on the whole area nevertheless.

So, as I decided to sever the links to Tolkien's Elvish, I began to sift through the vocabulary of western European languages for words that seem to have been borrowed from a substratum language. At first, the *Duden Herkunftswörterbuch* (Duden-Redaktion 2007), a popular etymological dictionary of German, was my guide, but later, I compiled a list of possible substratum loanwords from the Leiden Celtic (Matasović 2009) and Germanic etymological dictionaries (Kronen 2013) to put the endeavour on more solid ground. The first Old Albic word found this way was *abal* 'apple'. Many more followed, including the self-designation of the Elves, **Elbi** (singular *Alba*; hence the language name 'Albic'), which I re-created from the Germanic ***albaz** ~ ***albiz** 'Elf', a word with uncertain etymology. The aforementioned 'salt'-word, **hal**, also went into the Old Albic lexicon. There are also some words, based on the notion that the Irish Travellers are in part descendants of the pre-Celtic people of Ireland, i.e., of Elves, from *Shelta* (the cant of the Irish Travellers), such as **mynich** 'name'.

Some words have rather complex etymologies. One of my favourite words is *pana* 'monkey'. The starting point was the Germanic word for this animal (English **ape**, German **Affe**) which does not have a satisfactory etymology; it is held to be "from an unknown language" (*aus einer unbekanntem Sprache*; Duden-Redaktion 2007:23). I decided that this unknown language would be Old Albic. But what is the Old Albic word, and where did *they* get the word from, monkeys not being native to the British Isles? The first attempt at the word was **apa**, which did not feel right to me; I was especially bugged by the **p**, which would be the cognate of ***b** which appears to have been missing from PIE, as said above. Also, the Germanic word is an n-stem, so shouldn't the Old Albic word rather be **apana**? That's fine to explain the Germanic n-stem, but the "p-problem" remained.

But eventually, I found a different solution. There is a similar word in Old Albic, **phana**, which means 'goblin'. Now a monkey is not a goblin, but it looks vaguely like what people imagine

goblins to look like, so a semantic shift from ‘goblin’ to ‘monkey’ is quite plausible (compare how the word *lemur*, originally referring to a goblin-like being, came to denote a monkey-like animal of Madagascar). Now the Elves learned of monkeys from someone further south, and in the Iberian Peninsula, there is another branch of Hesperic, *Durian*, of which I hadn’t found out much yet, but I already knew that in that branch, the aspirated stops lost their aspiration. So the Durian cognate of Old Albic **phana** would be **pana**. This word acquired the meaning ‘monkey’, as the Durian speakers found that this exotic animal, which they knew from trade with people in northwestern Africa where monkeys occur natively, looked enough like what they imagined a goblin to look like, and called it that way. So the Durian word **pana** came to mean ‘monkey’, and in this meaning was borrowed into Old Albic. And the missing initial **a** required by the Germanic ***apan-** was no problem, either, since that is nothing else than the definite article in Old Albic, i.e. yet another instance of *alpha mobile*. So everything fell into place.

Thus the lexicon grew. But what about the grammar? Both the phonology and the morphology fell out of the way Old Albic was meant to be related to Indo-European. I have already given the main stop and vowel correspondences above; these were fleshed out with a more detailed sound change list. For my language, I started with a homebrew internal reconstruction of a pre-stage of Proto-Indo-European, drawing on the works of academic scholars like the aforementioned Gamkrelidze and Ivanov, and the late, lamented Danish scholar Jens Elmegård Rasmussen, who had produced a number of interesting articles on this matter (Rasmussen 1999). While this reconstruction probably would not stand the test of academic historical linguistics for long, I think it is good enough to build a conlang on it. The language in question would differ from PIE proper in being first an agglutinating rather than fusional language, and second in being an active-stative language.

In the nominal inflection, I pursued ideas which already had begun to take shape in Nur-ellen. The PIE accusative suffix ***-m** became the marker of the objective case (the case of direct objects and non-agent subjects) on animate nouns, while the agentive case (the case of agent subjects) was unmarked. Inanimate nouns, in contrast, have an unmarked objective case and *no* agentive case; their case paradigms are thus defective. The idea behind this was that the suffix ***-m** was an “animacy-neutralizing” marker in the proto-language, which was added to animate nouns when no animacy was syntactically required. There are three other cases in Old Albic that are limited to animate nouns, the genitive (alienable possessor), the partitive (inalienable possessor) and the dative (mentally affected object). The dative also marks an involuntarily (e.g. accidentally) acting subject, thus forming a part of the “degrees of volition” system of the language. Five further cases exist both in the animate and the inanimate paradigm; they are formed with suffixes added to the objective case form, which with animate nouns means that the suffix follows **-m**. These cases are the instrumental and four local cases – the locative, allative, ablative and perlativ. Adjectives agree with the noun in gender (there are four: masculine, feminine, common animate and inanimate), number and case. And so do nouns modifying nouns, such as genitives – Old Albic has *Suffixaufnahme*. The inanimate instrumental singular of adjectives was also used to form adverbs (like English *-ly*).

The result is a paradigm that looks like this for an animate noun (*Alba* ‘Elf’):

Case	Singular	Dual	Plural
Agentive	Alba	Olbu	Elbi
Genitive	Albas	Olbus	Elbis
Partitive	Albal	Olbul	Elbil
Dative	Alban	Olbun	Elbin
Objective	Albam	Olbum	Elbim
Instrumental	Albemi	Albymi	Elbimi
Locative	Albamal	Olbumul	Elbimil
Allative	Albaman	Olbumun	Elbimin
Ablative	Albamad	Olbumud	Elbimid
Perlative	Albamath	Olbumuth	Elbimith

And for an inanimate noun (*char* ‘stone’):

Case	Singular	Dual	Plural
Objective	char	chorum	cherim
Instrumental	cheri	chorymi	cherimi
Locative	charal	chorumul	cherimil
Allative	charan	chorumun	cherimin
Ablative	charad	chorumud	cherimid
Perlative	charath	chorumuth	cherimith

As with the nominal inflection, I repurposed Indo-European verbal inflection morphemes to build up an active-stative system. The present/aorist endings became agent markers, the perfect/middle ones became patient markers, resulting in a bipersonal paradigm (**sili** ‘to see’):

P ↓ A →	0	1sg	2sg	3sg	1du	2du	3du	1pl	2pl	3pl
0	sili	selma	seltha	selsa	sylmu	sylthu	sylsu	silmi	silthi	silsi
1sg	selha		selhatha	selhasa		silhothu	silhosu		silhethi	silhesi
2sg	selcha	selchama		selchasa	silchomu		silchosu	silchemi		silchesi
3sg	sela	selama	selatha	selasa	silomu	silothu	silosu	silemi	silethi	silesi
1du	solhu		silhothu	silhosu		sylhuthu	sylhusu		silhythi	silhysi
2du	solchu	silchoma		silchosu	sylchumu		sylchusu	silchymi		silchysi
3du	solu	siloma	silothu	silosu	sylumu	syluthu	sylusu	silymi	silythi	silysi
1pl	silhi		silhetha	silhesa		silhythu	silhysu		silhithi	silhisi
2pl	silchi	silchema		silchesa	silchymu		silchysu	silchimi		silchisi
3pl	sili	silema	siletha	silesa	silymu	silythu	silusu	silimi	silithi	silisi

So, 'I see you' is **selchama**, and 'You two see them all' is **silythu**. ('0' in the table means 'no agent'/'no patient'). The missing entries are those where both arguments are 1st or both 2nd person; this is handled with the reflexive marker **-r** instead. Another peculiarity is that inanimate objects always receive singular marking on the verb, so 'I see the stones' is **selama im cherim**, as opposed to 'I see the Elves', **silema im Elbim**.

The syntax of Old Albic was built from another source – the Insular Celtic languages. I have already stated above that the Celtic languages of the British Isles diverge strongly from common Indo-European patterns – even from Continental Celtic, which is in this regard quite conservative – and that this is often ascribed to a substratum language of unknown kind (though some scholars opine that this substratum was Semitic or related to Semitic; but actually, Semitic is not a close match to Insular Celtic syntax beyond being VSO, which is not all that rare, and found in regions not likely to have received Semitic linguistic influences such as Central America or the Philippines). I decided that Old Albic was just this substratum. So Old Albic has a radical head-initial, right-branching word order, and a tendency to phonetically run syntactically closely connected words together, which in combination with a phonological rule that slightly lenits stops after vowels, results in (subphonemic) initial mutations. (As these mutations do not yet change the phonemes, they are not represented in writing – true phonemic initial mutations appear later). There are also verbal noun constructions similar to those in the Insular Celtic.

Old Albic being the language of an urban civilization which I fancy to underlie not only the Germanic and Celtic traditions of elves, but also the Greek traditions of Hyperborea and Atlantis, it is of course also a *written* language. The Old Albic writing system is a featural alphabet, with the letter shapes encoding phonological features. This idea of course goes back to Tolkien's Tengwar, but the Old Albic alphabet is not much like Tengwar apart from being featural. Such scripts are unlikely to evolve "naturally"; they are usually the result of a conscious invention. This is true, for instance, for the Korean Hangeul script (the only featural script in actual use for a major language of our world), which was developed by scholars on behalf of a 15th-century king of Korea. The Old Albic alphabet is likewise intrafictionally attributed to a cultural heroine. The letters are not invented from thin air, though; the basic letter shapes, from which all others are derived, are those of the voiceless stops, which are similar to the Phoenician letters *pe*, *taw* and *kaph*.

The result of all these considerations is the Old Albic language as it is known to the conlanging community today. But, as I have already said in the first sentence of this section, Old Albic is not an isolate. It is a member of a larger family named Hesperic. It has dialects, descendants and relatives. As I am writing this, I do not yet know how many of these there will be in the end, but a picture is emerging.

The Commonwealth of the Elves, the political system in which the Elves lived at the apex of their civilization, is divided into twelve **cvendi**, a term which can be rendered in English as 'tribes' or 'regions'. Accordingly, there are twelve main dialect groups, one for each **cvanda**. Drafts of sound changes for these main dialects already exist in my files, though many things may still change. Of course, the dialects form a dialect continuum where neighbouring dialects are similar to each other, as can be seen in any natlang with a considerable dialectal variation.

When the Commonwealth of the Elves collapsed from internal strife into which foreign powers meddled, and the British Isles became a predominantly Celtic-speaking country, the Albic language was not utterly eradicated, however. Some small pockets of Elvendom survived, and today, there are the “Moonchildren”, Elves that live a Gypsy-like existence all over Europe and the Americas as travelling showmen. All these splinters of the once great Elven nation speak Albic languages, which of course have changed a lot from Old Albic. None of these Neo-Albic languages has been elaborated in detail yet, but basic ideas exist for some of them.

And then there are other Hesperic languages in other parts of the continent. There are nine main branches which can be placed on a 3x3 grid according to their geographical location:

Albic (British Isles)	Viddan (Norway)	Valdiska (Latvia)
Puranian (France)	Hercyno-Alpianic (Germany/Switzerland)	Duniscian (Poland)
Durian (Spain/Portugal)	Padivian (Italy)	Dravinian (Croatia)

Again, there are similarities between nearby branches. These, after all, emerged from a Proto-Hesperic dialect continuum. They are not yet elaborated in detail, though the Hercynian and Alpianic languages are most developed, and clear pictures of Proto-Alpianic and Proto-Hercynian have already emerged. Altogether, there will be at least 20 Hesperic languages, perhaps 30 or even 40, though most will remain sketchy, consisting merely of a comprehensive grammar overview and a word list.

Proto-Hesperic is meant to have been spoken in what is now Hungary around 4000 BC, and accordingly, Dravinian is an early divergent branch, something like the Anatolian branch of Indo-European; like Hittite, Dravinian preserves some archaisms otherwise lost. The remainder of Proto-Hesperic broke up into an eastern branch, to which Valdiska and Duniscian belong, and a western branch which is made up of all the rest; within this, Viddan is in turn divergent, while the rest evenly split into Albic, Hercyno-Alpianic (a “double branch”, consisting of Hercynian in Germany and Alpianic in Switzerland, like Balto-Slavic or Indo-Iranian in Indo-European) and a southwestern group consisting of Puranian, Durian and Padivian. The result is the following family tree:

Hesperic

1. Western
 - 1.1. Narrow Western
 - 1.1.1. Central
 - 1.1.1.1. Albic
 - 1.1.1.2. Hercyno-Alpianic
 - 1.1.1.2.1. Hercynian
 - 1.1.1.2.2. Alpianic
 - 1.1.2. Transrhenan
 - 1.1.2.1. Puranian
 - 1.1.2.2. Mediterranean
 - 1.1.2.2.1. Durian
 - 1.1.2.2.2. Padivian
 - 1.2. Viddan
2. Eastern
 - 2.1. Northeastern
 - 2.1.1. Valdiska
 - 2.1.2. Duniscian
 - 2.2. Dravinian

3. Razaric, Midrean and Para-Kartvelian

But what would the “Dwarves”, an earlier stratum of the British Isles than the “Elves”, speak? And if the LBK people did not speak Proto-Hesperic, what did they speak then? And were these two languages related? I came to the result that *Razaric*, the language of the “Dwarves”, was a descendant of the LBK language, and that the LBK language was related to Kartvelian. I have always been fascinated by the Kartvelian languages such as Georgian since I chanced upon them in my work on the Elvish language and my forays into linguistics. When I found that the LBK people were genetically similar to modern Georgians (in both groups, for instance, the most common Y-DNA haplogroup is G2a), I was amazed by the idea that languages related to Kartvelian were once spoken all over much of Europe!

Also, there is a study by the late Dutch linguist Edzard Furnée according to whom the Pre-Greek substratum – the unknown language that was spoken in Greece before the immigration of Greek and has left substratum loanwords in Greek – may have been related to Kartvelian (Furnée 1979). While most of Furnée’s etymologies are questionable, this fit my plans. In an earlier book (Furnée 1972), Furnée had worked out phonetic vacillations in Pre-Greek words which allow to reconstruct a hypothetical phoneme inventory of Pre-Greek. This inventory is quite similar to the Proto-Kartvelian one but misses quite a few phonemes – the uvulars, the ejectives, the mid and back sibilants, the lateral fricative, the vowel **o** – which are, with the exception of **o**, exactly those which Greek could not render and would have merged with other sounds; as for **o**, there is evidence that it is secondary in Kartvelian in a similar way as ***a** is secondary in PIE (Gamkrelidze/Mačavariani 1982:93).

I already had a phoneme inventory for Proto-Razaric, which was not designed to be compatible with Kartvelian, but I did not want to abandon, and could be reconciled with it. The most unusual trait of this inventory was the existence of three alveolar affricate/fricative series, one sibilant,

one flat (rhotic) and one lateral. I had the idea that these would correspond to the three sibilant series of Proto-Kartvelian:

Razarcic	Kartvelian
sibilant	front
flat	back
lateral	middle

I am not sure of this, though, and consider the idea that instead, the flat and lateral affricates originated in Kartvelian “harmonic clusters”, namely the flat affricates from clusters with uvulars (e.g. **tq > tr*), and the lateral ones from ones with velars (e.g. **dg > dl*), while the three sibilant series fell together into one in Razarcic. I have not decided about this matter yet.

The method of building my family, which I provisionally named “*Midrean*” (from the Proto-Midrean word for ‘human being’), is essentially the same as that I use for Hesperic. The Midrean languages are related to Kartvelian in the same way as Hesperic is to Indo-European, except that the time depth is far greater. Proto-Midrean was spoken about 8,000 years ago, and the common ancestor of Midrean and Kartvelian would have been at least 1,000 years earlier still. There are thus fewer common points between Midrean and Kartvelian than between Hesperic and Indo-European; the relationship was about as old as the hypothetical relationship between Indo-European and Uralic (which is not established yet, but I am in favour of this hypothesis). I was thus much freer in designing Midrean than I was in designing Hesperic; also, many Hesperic words could be loanwords from Midrean.

Midrean of course has its own branch structure. First, there is a more conservative eastern group, spoken in the Pannonian Basin and the Lower Danube area, which keeps, among others, the three sibilant series of Proto-Midrean intact; and a western group, spoken in Central Europe and the British Isles, where the mid and back sibilants had shifted to laterals and rhotics as laid out above for Razarcic.

As I am writing this, all this, however, is in a preliminary stage; no grammar sketch of Proto-Midrean or any of its daughter languages exists yet, and many things are likely to change in the future.

4. Eteonoric, Mesolithic languages, Para-Uralic and others

And there is even more yet to explore! What did the people of Europe speak *before* the Neolithic? The Europe of hunter-gatherers probably was linguistically highly diverse when the first farmers came in. Each of the three great Mediterranean peninsulas may have had one to three separate families of its own (in the Iberian Peninsula, we have three apparently mutually unrelated languages – Basque, Iberian and Tartessian – in the Iron Age, though one, two or all three may have entered in the Neolithic or later). North of the Alps, the diversity probably was less, but I expect one family in the west (I call this *Paleo-Atlantic*) and one in the east (I call this *Paleo-Pontic*); maybe a third one in the Pannonian Basin (*Paleo-Pannonian*). It is possible, but uncertain, that

Basque, as Vennemann (2003) has it, is the last remnant of Paleo-Atlantic (but Basque may just as well be a Neolithic arrival instead, and Vennemann's etymologies, often attempting to account for geographic names by means of Basque which could be easily accounted for by means of known languages such as Celtic or Old High German, are hard to take seriously at any rate).

What were these languages like? The few surviving and attested languages give some ideas. In the Mediterranean, there seems to have been a tendency towards small consonant inventories in the west and large ones in the east. Proto-Basque, the early form of Basque which can be reconstructed to have been spoken in Roman times, had 16 consonants, coming in strong/weak pairs contrasting only in medial position, which may point at a yet earlier stage with just eight consonants (Trask 1997:125-127). Iberian and Tartessian also had fairly small inventories, if the writing systems they used are not underspecifying. Etruscan, whether it is native to Italy or originates further east, has 17 consonants, and the Caucasian languages are famed for their rich systems (Georgian with its 28 consonants is the *smallest* – it goes all the way up to 80 consonants in Ubykh).

For the northern area, I conjectured, based on the kinds of inventories we find in the Indo-European languages of this area, that the western languages were sibilant-poor and the eastern ones were sibilant-rich.

There already is a language made by a group effort in the League of Lost Languages, *Proto-Eteonoric*. This is meant to be a Paleo-European language in the eastern Alps, and there is no way I could integrate it into either Hesperic or Midrean. Hence, I decided that it was a surviving Mesolithic language. With its rich sibilant system, Eteonoric probably is a Paleo-Pontic or Paleo-Pannonian language.

Finally, if there are Para-Indo-European and Para-Kartvelian conlangs, why not also Para-Uralic ones? The idea that such languages could have existed arose in me when I read a paper by Petri Kallio according to which the long held connection between Proto-Uralic and the Pit-Comb Ware culture was problematic because Proto-Uralic was much younger than the latter (Kallio 2015). So there may have been a language family related to Uralic like Hesperic is to Indo-European. Another idea for future projects!

And then, one could ask, “What if Iberian and Etruscan hadn't died out completely, but left modern descendants?” Even that is an attractive challenge to the conlanger re-creating Old Europe.

5. Conclusions

So the prehistoric linguistic landscapes of Europe offer a great playground for a conlanger! We will probably never be able to *reconstruct* these languages, but we sure can *re-create* them with our conlanging skills.

6. Bibliography

Baker, Martin (s.d.). *Home Eleven*. <http://www.alt-tolkien.com/r13home11.html> (last access: 05/12/2018).

- Duden-Redaktion (2007). *Das Herkunftswörterbuch*. Etymologie der deutschen Sprache. Mannheim.
- Fortson, Benjamin W. (2010). *Indo-European Language and Culture*. Chichester: Wiley-Blackwell.
- Gamkrelidze, Tamaz; Ivanov, Vyacheslav (1995). *Indo-European and the Indo-Europeans*. Berlin: Mouton de Gruyter.
- Gamkrelidze, Tamaz; Mačavariani, Givi (1982): *Sonantensystem und Ablaut in den Kartwelsprachen. Eine Typologie der Struktur des Gemeinkartwelischen*. Tübingen: Narr.
- Kallio, Petri (2015). The language contact situation in prehistoric northeastern Europe. In: Robert Mailhammer, Theo Vennemann, Birgit Anette Olsen (eds.), *The Linguistic Roots of Europe*, pp. 77-102.
- Krahe, Hans (1963). Die Struktur der alteuropäischen Hydronymie. *Abhandlungen der geistes- und sozialwissenschaftlichen Klasse, Akademie der Wissenschaften und der Literatur Mainz*. Stuttgart: Steiner.
- Rasmussen, Jens Elmegård (1999): *Selected Papers on Indo-European Linguistics. With a section on comparative Eskimo linguistics*. 2 volumes. Copenhagen: Museum Tusulanum Press.
- Ryan, William; Pitman, Walter (1998): *Noah's Flood*. New York: Simon & Schuster.
- Trask, R. L. (1997): *The History of Basque*. London: Routledge.
- Vennemann, Theo (2003): *Europa Vasconica – Europa Semitica*. Berlin: Mouton de Gruyter.