Univerza v Ljubljani
FILOZOFSKA FAKULTETA

Masaryk University
Faculty of Arts

Arna Nováka 1/1
602 00 Brno
Czech Republic

Asst. Prof. Dr. Luka Repanšek

Chair of Indo-European Comparative Linguistics
Department of Comparative and General Linguistics
Faculty of Arts, University of Ljubljana

Aškerčeva 2, SI-1000 Ljubljana, Slovenija

Opponent's review of the habilitation thesis

The development of the IE clusters *obstruent + t/s(d^h)*

by PhDr. Ondřej Šefčík, PhD

The habilitation thesis by Dr. Šefčík deals with an important, problematic, and therefore often discussed topic of Indo-European comparative linguistics, namely that of the development of consonantal clusters of plosives (voiced or voiceless) followed by plosives or the PIE sibilant **s*. The thesis is true to its title in as far it deals with all such combinations, although it quickly becomes clear that the focus of the work lies on a) the development of PIE *TT*- and *Ts*-clusters as far as the *centum*-languages are concerned (which regarding any other possible combinations are not really that problematic) and b) the development of *TT*- and *Ts*-clusters as well as *ḰP* (especially *ḰT*) and *K*^{j}*s* clusters in the *satəm*-group, where the latter are particularly relevant. He does not deal with the so-called *thorn*-clusters (**TḰ*), which he justifies through scope limitations. Admittedly, given the actual focus of this thesis, the treatment of *thorn*-clusters would naturally fall out of its scope as there the problems involved are of an entirely different nature if one accepts the idea that they *did* develop differently in core Indo-European, i.e. that their outcomes in the individual branches/languages are the result of a shared phonetic modification of the starting point. The thesis is designed to be a comprehensive account of the behaviour of **Pt* and **Ps* clusters across the entire IE language family, which it does succesfully by providing numerous examples for each of the combinations at interest, branch after branch. In this respect it undoubtedly represents the first systematic treatment of such PIE sequences and their development towards the individual archaic IE languages to-date and will be useful as a starting point for further research on the topic. Other works on a similar topic are significantly less exhaustive by dealing only with a fraction of what this thesis aspires to

systematically assess. The mentioned aspect of the thesis is of course purely classifiying and results in a well-designed catalogue of the relevant questions and examples to back them up. The actual novelty that it brings, however, is the approach with which the development of the reconstructed PIE sequences is tackled. The author criticises what he calls the "affricativisation" process, with which the evolution of, say, PIE *TsT*- and *Ts*-clusters or **ḰT*-clusters (in *satəm*-languages) is usually explained, and proposes to build on an old idea of "spirantisation" that has, according to the author, a greater explanatory power and at the same time takes care of the problematic cases that the "affricativisation"-type of development leaves unaccounted for. The basic idea is this: in every such cluster the plosive developed into a spirant rather than an affricate. Such spirants were later (depending on the environment) assimilated, turned into sibilants, were subject to fortition etc. The author shows very capable command of the data and is comfortable navigating the historical phonologies of all the main branches of PIE, using the methodology that is generally used in the field of historical linguistics/comparative linguistics. The results of the thesis must therefore be taken seriously, given that they are arrived at within the framework of both comparative linguistics and general linguistics, the latter in as far as the principles/findings of the phonological theory are being taken advantage of in assessing the probability of the reconstructed middle-stages in the development of a particular cluster. The linguistic analysis is well planned-out and structured: the author analyses all instances of the clusters in question in all the relevant combinations (excluding but the **TḰ*-clusters, as explained), following their reflexes in the individual branches or sub-branches, backing up each reflex with several examples, upon which follows a careful analysis of the state of affairs: first, the author gives the traditioanl account, identifies the issues that this approach brings and then tries to see whether the novel theory might better account fot the attested reflexes. The overal conclusion of the thesis is that there indeed exist several reasons why the "spirantisation" model yields a more convincing trajectory than starting from an affricate would. The language of the thesis is English, but there are several instances of unidiomatic use, some even of ungrammatical structures (especially concerning the use of the definite article and in a few cases sentence structure) and unfortunate choice of wording. These do not, however, generally make the point the author is trying to make unclear, since to an Indo-Europeanist it is mostly unambiguously clear what the point being made is anyway. I do suggest, nevetheless, that the text be looked through and corrected by a native speaker before its publication.

The objections that as the formal opponent I must voice at this point are as follows:

a) I am not very impressed with the Old Church Slavonic being used as the *sole* source of examples for the presentation of the Slavic data. Indeed, this is the oldest attested Slavic language, but it is *exactly* the clusters in question that Eastern South Slavic (which OCS is genetically) has very uncharacteristic reflexes for, so I would propose that the individual reflexes be presented in their reconstructed Proto-Slavic stage.

b) I think it is misfortunate that the term Old Indo-Aryan is being used for the Vedic data as if Vedic were the only form of Old Indo-Aryan – in several respects, which are absolutely crucial to the topic at hand since they involve *exactly* the discussed sequences, the Prakrit dialects (Middle Indo-Aryan) are a very important witness. Without the Prakrit data the image one gets is completely distorted. Consider, e.g., the voiceless *kṣ*-reflex of the Vedic/Sanskrit idiom for PIE **ǵʰt* et sim., which is a defining feature of *that* particular dialect of Old Indo-Aryan only.

c) The section on Celtic does not pay sufficient attention to the Lepontic and Celtiberian data. Especially as regards the author's conclusions on the behaviour of Continental Celtic, the data is absolutely biased, since they are formed on the basis of Gaulish alone. To claim, then, that for Cont. Celtic a development along the lines of *$s + t$ > *ϑs > *$\vartheta\vartheta$ (> *đđ*/*ss*) might be proposed really runs counter to the Celtiberian data at least, which unambiguously point to *st*-clusters being preserved in both Anlaut and Inlaut. The entire dataset of Cont. Celtic must therefore be taken into account otherwise the results of the analysis will necessarily be faulty.

d) I object to the idea that there is *any* kind of regularity in the *ḍḍh* type of reflexes of PIE *$\acute{k}/\acute{g} + d^h$. It is absoutely clear that every instance of *ḍḍh* in such cases is *purely* analogical, the *ḍ* (actually a voiced allophone of *ṭ*) having been introduced before the ending (be it verbal or nominal) to prevent allomorphy that would have otherwise arisen after the deletion of *$\underline{z}$. Only such a scenario explains why there is a difference between, say, *rīḍhi* and *dviḍḍhi* in Vedic/Sanskrit. The proportion being *dveṣ-ṭi*: *$dviz$-ḍhi > *dvīḍhi → *dviḍ-ḍhi*, the synchronic equation *ṣ* = *ṭ* being won in the 2./3. sg. act. impf. of the type *a-dveṭ* (vs. *reḍhi* : *dvīḍhi*).

e) I am not convinced that the spirantisation principle accounts better for the data as far as the *ḰT*/*Ḱs*-clusters are concerned. There is no avoidance of an affricate stage at any rate even if spirant articulation of *Ḱ* in front of *t* etc. is assumed. Such a stage could only be secondarily achieved *through* an affricate. At the same time it must also be stressed that the secondary fortition processes that many of those theoretical chains would in the end demand is not very probable phonetically and no real phonological/phonetic reasons for these proposed developments are laid out.

f) I am not convinced that cases such as *voḍhum* etc. point to the development of the voiced dental into an approximant. Rather what we have here is the *a being rounded by the preceding *w and then lengthened by compensatory lengthening just as any other *a would be, so *$waz\underline{ḍ}^hum$ > *$woz\underline{ḍ}^hum$ > *$wō\underline{ḍ}^hum$ = *voḍhum* etc. In the case of *dhehi* ~ *dehi* < *dhazdhi* ~ *dazdhi*, at least in my view, the *j as a replacement for *d is only apparent, the *z having been lost and the resulting hiatus having been filled with *j just as it sporadically happens in external sandhi, cf. *sūre* < *sūraz*. Meanwhile, *yodhi* and *bodhi* are (as per Jasanoff) most unambiguously analogical formations after *joṣi*, given that these two roots are the only ones that behave this way in the aorist imperative and are at the same time the only ones that end in the same consonant as the 2. sg. act. impv. ending starts in. These two at least cannot and should not be used as examples of a voiced dental developing into an approximant.

g) *TS* (meaning *kṣ*) does not analogically replace the expected voiced reflex (*DZ*) in Vedic/Sanskrit – the voiceless reflex is simply the result of regressive voicing assimilation, which is a reaction to *Z > *S*, the defining feature of Vedic/Sanskrit being systematic elimination of voiced (+/– aspirated) sibilants, i.e. the voiced-sibilant filter (as opposed to Middle Indo-Aryan dialects).

h) That Proto-Indo-Aryan possessed a voiced aspirated sibilant is practically a given, seeing that *$\underline{ʒ}^h$ is more or less directly attested in Hitt. *$\underline{u}$ašanna* = Proto-Indo-Aryan *$w\acute{ā}\underline{ʒ}^hana$- and given the fact that otherwise there is no convincing way of explaining why PIE *$\acute{g}^h$ and *g /__E,j yielded *ĥ* in that system in the end: only an *aspirated* voiced sibilant would upon the operation of the voiced-sibilant filter be debuccalised to a voiced glottal *spirant*.

i) I am not at all convinced by the argument that *tar-* from etymological **ptar-* is more convincingly explained by assuming **ptar- > *ftar- > *htar- > tar-*. There is no example to support such a development for Avestan and the by-form *ptar-* strongly suggest that *tar-* is a younger and sporadic development reflecting simple cluster-simplification.

j) The main argument in favour of the spirantisation model in the case of **ḰT* in Vedic/Sanskrit is that PIE **tst < *tt* sequences and **tʃt < *ḱt* sequences (if following the affricativisation principle) should have behaved the same. This argument is not cogent, if admitted that that there is absolutely no proof that inherited **tst* sequences behaved as **t͡st* (i.e. sequences of an actual *secondary* affricate and *t*) rather than **tst*, while **tʃt < *ḱt* was surely just such a sequence, the different treatment of **tst* vs. **tʃt* being then expected and easily accounted for under any model. That **tst* yields **st* in Iranian is not a good argument in favour of Proto-Indo-Iranian **t͡st*, given that Proto-Iranian monophthongises *all* primary and secondary *Ts* sequences into secondary affricates, which are then systematically simplified in front of plosives, so **sts > *t͡st > *st* in Proto-Iranian is an unproblematic development, which is abundantly paralleled within that system.

k) I do not see any reason why following the affricativisation model one would need to propose a chain like **tšš* > (**tš* >) **šš > *š* for Iranian. This is a oversimplification in my view, which then does not really do justice to the standard proposal that one should in fact start from original affricates. For Iranian I would rather propose the following set of changes (chronologically):
1) **tʃ, *dʒ > *ʃ, *ʒ / __n*
2) monophthongisation:
**t + *ʃ > *tʃ, *d + *ʒ > *dʒ, *t + *s > *ts / __*
**t + *ɕ > *tɕ, *d + *ʑ > *dʑ / __* (coallescence with **tɕ, *dʑ*)
3) **tʃ, *dʒ > *ts* (coallescence with the new **ts*), **[dz] / __* by push-chain
4) **s, *z > ∅ / #,$__*ts, *dz*
5) **ts, *dz, (tʃ), *dʒ > *s, *z, (*ʃ), *ʒ / __T(#), T__*
6) **t > ∅ / s,ʃ__#*
7) **s, *z > *ʃ, *ʒ / [+ bilabial]__[– alveolar]*
8) **ts > *s / *s$__*
9) **ts > *s̱, *dz > *ẕ*
10) **tʃ* (< **tḱ, *ḱs*), **dʒ* (< **ǵʰs*) *> *ʃ, *ʒ / __*

l) In ft. 54 there is a statement that the apprently short *r̥* in *tr̥ḍhá-* & co. for **tr̥z̥ḍhá-* is probably analogical for *ri < *r̥̄*, but this view cannot be upheld. The compensatory lengthening of vowels in **Vǵʰt* type of sequences, which is otherwise regular, is *systematically* absent if that vocalic element is *r̥*, simply for the fact that *r̥ could not accept* secondary length (mind that there is *no* long syllabic *r̥* in Proto-Indo-Aryan, all the cases involving the Apl. and Gpl. of *r*-stems are purely analogical).

m) I would challenge the idea that "the Old Indo-Aryan data leads us towards **kš* uniformly for verbs and nouns, i.e., towards the original neutralization of palatovelars to a plain velar before original **s*" (p. 63). This cannot at all be true as otherwise the *ṭ* reflex of the same sequences that otherwise yield *kṣ* in absolute auslaut could be explained. Words like *spaṭ* & co. < **-ḱs#* unambiguously prove that there was a middle stage **-ṭṣ-*, later dissimilated to *kṣ* in Inlaut, but kept as *ṭ* since the relative chrnonological ordering of **CC(C) > C / __##* is one of precedence over the **ṭ > k / __ṣ* rule.

n) There is no strength in the claim that the *tau Gallicum* was phonetically a spirant rather than an affricate *ʦ*. The author claims that the latter value is arrived at by what is predicted by the theory (**tst* > **ʦt* > **ʦ* vel sim.), but his own account actually does no better. He simply assumes that the graphematic representation of the phoneme at question was or must have been a voiceless dental spirant to accommodate the predicion of his own theory. The proposal that the *zraif* character of the *m*-Ogam series was originally actually a *ʦ* is not seriously considered, only briefly mentioned in a footnote. It is indeed more likely that *zraif* stood for **s*^{w} but that question is far from being settled, so a discussion is of course in order, especially given that this has significant bearing on determining the original value of the *tau Gallicum*. I, for one, see no problem in assuming that **tst* > **ʦ* in Proto-Celtic, given that the exact same thing has happened in Proto-Italic as well as Proto-Germanic, where there is even indirect proof of such an intermediate stage if one considers that any old **ʦ* sequence inherited nto Proto-Germanic eventually merged with the reflex of **tst*.

Disregarding the above remarks, the framework within which this thesis is carried out is methodologically convincing and consistent, so that it **meets the standard requirements placed on habilitation thesis in the field of General and diachronic linguistics**.

Ljubljana, 9. 12. 2021 Luka Repanšek