Spring 6-1-2010

Reconstructing Indo-European Syllabification

Andrew M. Byrd
University of Kentucky, andrewbyrd@uky.edu

Reconstructing Indo-European Syllabification

UNIVERSITY OF CALIFORNIA
Los Angeles

Reconstructing Indo-European Syllabification

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Indo-European Studies

by

Andrew Miles Byrd

2010

The dissertation of Andrew Miles Byrd is approved.

Christopher M. Stevens

Brent Vine

Calvert Watkins

Kie Zuraw

H. Craig Melchert, Committee Chair

University of California, Los Angeles

2010

To my parents, Jack and Vicky,

who have always encouraged me to look for four-leaf clovers.

TABLE OF CONTENTS

Symbols

C	consonant	U̯	glide
C̥	syllabified consonant	V	vowel
F	fricative	V̄	long vowel
H or h_x	laryngeal	*	reconstructed as, violation
K	tectal	**	non-occurring
L	liquid	>	diachronically develops into
N	nasal	<	diachronically derives from
O	obstruent	→	becomes by analogy or sound law
P	stop	↣	becomes by analogy & sound law
R	sonorant	μ	mora
R̥	syllabic sonorant	σ	syllable
T	dental stop	$\$ \text{ or }]_{\sigma} \text{ or } .$	syllable boundary
U	high vowel	#	word boundary

Abbreviations

acc.	accusative	Alb.	Albanian
aor.	aorist	Arm.	Armenian
Av.	Avestan	Bal.	Balochi
Boeot.	Boeotian	Bret.	Breton
Corn.	Cornish	Cret.	Cretan
Cyp.	Cypriot	dat.	dative
Dor.	Doric	du.	dual
Eng.	English	Gaul.	Gaulish
GAv.	Gathic Avestan	gen.	genitive

Germ.	German	Gk.	Greek
Gmc.	Germanic	Goth.	Gothic
EWAia	Mayrhofer 1985-2001	Hitt.	Hittite
HLuv.	Hieroglyphic Luvian	IE	Indo-European
IEW	Pokorny 2005	Ion.	Ionic
impfct.	imperfect	instr.	instrumental
Ital.	Italian	iter.	iterative
Jap.	Japanese	Khot.	Khotanese
Lac.	Laconian	Lat.	Latin
Latv.	Latvian	Lith.	Lithuanian
LIV	Rix et al. 2001	loc.	locative
Luv.	Luvian	Lyc.	Lycian
masc.	masculine	MBret.	Middle Breton
MIr.	Middle Irish	M/P	mediopassive
MWel.	Middle Welsh	NIL	Wodtko et al. 2008
nom.	nominative	NPers.	Modern Persian
nt.	neuter	OCorn.	Old Cornish
OCS	Old Church Slavic	OE	Old English
OHG	Old High German	OIce.	Old Icelandic
OIr.	Old Irish	OLat.	Old Latin
OLith.	Old Lithuanian	OPhryg.	Old Phrygian
OPruss.	Old Prussian	OSax.	Old Saxon
Osc.	Oscan	Osc.-Umb.	Osco-Umbrian
OT	Optimality Theory	Parth.	Parthian

PGmc.	Proto-Germanic	PIE	Proto-Indo-European
PIIr.	Proto-Indo-Iranian	PInd.	Proto-Indic
PIran.	Proto-Iranian	PItal.	Proto-Italic
pl.	plural	pres.	present
PSlav.	Proto-Slavic	RCS	Russian Church Slavic
Russ.	Russian	TA	Tocharian A
TB	Tocharian B	Toch.	Tocharian
RV	Rig Vedic	sg.	singular
Skt.	Sanskrit	Slav.	Slavic
subj.	subjunctive	Umbr.	Umbrian
Ved.	Vedic	Wel.	Welsh
Yagh.	Yaghnobi	YAv.	Young Avestan

Acknowledgments

The following dissertation would not be nearly as good – and probably never even completed – if I had not been helped by so many along the way.

I would first like to thank my advisor, Craig Melchert, for giving an effective mix of skepticism and support throughout this process, and for putting a bug in my ear that the topic of Indo-European syllabification was one that so sorely needed revisiting. I have been blessed to have had a committee of excellent scholars at my disposal, with Kie Zuraw aiding me through the intricacies of generative phonology and Brent Vine, Cal Watkins and Chris Stevens through the mysteries of Indo-European linguistics.

I am indebted to the many other teachers that I have had throughout my academic career, beginning with Jared Klein at the University of Georgia, followed by Onofrio Carruba in Pavia, Jeremy Rau and Jay Jasanoff at Harvard, and Vyacheslav Ivanov, Raimo Anttila and Stephanie Jamison at UCLA. I am most especially grateful to Stephanie during my time in L.A., whose insightful work, though rarely dealing with consonant clusters per se, has taught me the value of effective presentation and of common-sense scholarship. I have also been fortunate to have studied generative phonology with Bruce Hayes and Kie Zuraw during my stay at UCLA, and to have been introduced to the wonderful world of Georgian (whose consonant clusters still make me giddy) with Pam Munro and Manana Batashvili as my guides.

I would also like to thank all of my fellow UCLA PIES graduate students, especially Jessica DeLisi and Tim Dempsey, who have been most kind to have read through multiple incarnations of what follows, as well as the sole other member of my 2004 cohort, Dieter Gunkel, for six years of challenging discussions and ideas. I would also like to show my appreciation to Masato Kobayashi and Götz Keydana, who have been kind enough to give me advice throughout this entire process and whose work the reader will no doubt recognize to be especially influential on my own.

I would last like to thank my family and friends for listening to me prattle on about Indo-European linguistics and consonant clusters throughout the years, especially my wife Brenna, who keeps me sane and smiling, and my parents, who have kept me driving on.

VITA

Sept. 25, 1979	Born, San Diego, California.
2002	A.B., Linguistics University of Georgia, Athens Athens, Georgia
2002–2003	Fulbright Fellowship Università degli Studi di Pavia Pavia, Italy
2003-2004	Visiting Student in Linguistics Harvard University Cambridge, Massachusetts
2005-2006	Graduate Research Mentorship University of California, Los Angeles
2007-2009	Teaching Assistant Linguistics Department University of California, Los Angeles
2008	Candidate of Philosophy Program in Indo-European Studies University of California, Los Angeles
2008	Lecturer in Linguistics California State University, Fullerton Fullerton, California

Publications and Presentations

Byrd, Andrew Miles (In press a). Motivating Sievers' Law. *Proceedings of the 21st Annual UCLA Indo-European Conference.*

——— (In press b). Deriving Dreams from the Divine: Hittite *tesḫa-/zasḫ(a)i-*. *Historische Sprachforschung.*

——— (In press c). Predicting Indo-European Syllabification through Phonotactic Analysis. *The Sound of Indo-European - selected papers from the conference, held in Copenhagen 16–19 April 2009.*

——— (October 2009). Motivating Sievers' Law in a Stratal OT Framework. Paper presented at the 21st Annual UCLA Indo-European Conference, Los Angeles, CA.

——— (April 2009). Predicting PIE Syllabification through Phonotactic Analysis. Paper presented at the Sound of Indo-European Conference, Copenhagen, Denmark.

——— (March 2009). Deriving Dreams from the Divine: Hittite *tesḫa-/zasḫ(a)i-* Paper presented at the 219th Meeting of the American Oriental Society, Albuquerque, NM.

——— (2008). Review of Kobayashi, Masato, *Historical Phonology of Old Indo-Aryan Consonants (Study of Languages and Cultures of Asia and Africa Mono-*

graph Series, No. 42), in the *Journal of the American Oriental Society* 128 (4):793-4.

——— (July 2008). Towards a Predictive and Explanatory Theory of Laryngeal Deletion in Proto-Indo-European. Paper presented at the 27th East Coast Indo-European Conference, Athens, GA.

Munro, Pamela, Virgil Lewis, and Nicole Gfroerer (2007a). *Shaap Kaij?*, with Marcus Smith, Andrew Byrd, Kevin Ryan, Kyle Wanberg, Heather Willson, and Jennifer Fischer. Academic Publishing, UCLA.

——— (2007b). Book Notice of Clackson, James and Birgit Anette Olsen (eds.), *Indo-European Word Formation: Proceedings of the Conference held at the University of Copenhagen October 20th - 22nd 2000*, in *Language* 83 (1):217.

——— (2006a). Return to Dative *anmaimm*. *Èriu* 56:145-55.

——— (2006b). Book Notice of Blake, Barry J. and Kate Burridge (eds.), *Historical Linguistics 2001: selected papers from the 15th International Conference on Historical Linguistics*, Melbourne, 13-17 August 2001, in *Language* 82 (3):671-2.

——— (2005). Book Notice of Laugesen, Amanda, *Convict words: language in early colonial Australia*, in *Language* 81 (3):782.

——— (November 2004). The Long Dative Singular of Neuter n-Stems in Old Irish. Paper presented at the 16th Annual UCLA Indo-European Conference, Los

Angeles, CA.

——— (2003). Book Notice of Vanseveren, Sylvie, ≪*Prodige à voir*≫ *Recherches comparatives sur l'origine casuelle de l'infinitif en grec ancien*, in *Language* 79 (4):823.

Abstract of the Dissertation

Reconstructing Indo-European Syllabification

by

Andrew Miles Byrd

Doctor of Philosophy in Indo-European Studies

University of California, Los Angeles, 2010

Professor H. Craig Melchert, Chair

The chief concern of this dissertation is to investigate a fundamental, yet unsolved problem within the phonology of Proto-Indo-European (PIE): the process of syllabification. I show that by analyzing the much more easily reconstructable word-edge clusters we may predict which types of consonant clusters can occur word-medially, provided that we assume a special status for certain consonants at word's edge. Having thus analyzed the entire PIE phonological system, I believe I have developed the first working hypothesis of Indo-European syllabification, which we may now use to predict which types of syllable-driven rules of consonant deletion and vowel epenthesis occurred within PIE. My dissertation argues that there existed at least five phonological processes of this type. The second half of the dissertation focuses on the problem of Sievers' Law, through which I argue for the tendency in PIE to keep morphemes syllabically distinct, in accordance with a high-ranking constraint ALIGN. I conclude by proposing that the assumption of morphological relevance in the syllabic derivation provides us with a mechanism to reconcile the well-established principle of ONSET MAXIMIZATION with the reconstructable parsing of VCCV sequences as VC.CV.

CHAPTER 1

Assumptions, Previous Scholarship & Goals

1.1 Overview.

The primary focus of this dissertation is the reconstruction of syllabification in Proto-Indo-European (PIE). PIE was a language that was never recorded, whose form may be inferred (reconstructed) through the comparison of attestations of its descendant languages, which include Latin, Russian, Hindi and English, among others.[1] Syllabification, as is well known, is a crucial component within the phonologies of all of the world's languages. It can drive phonological rules, dictate what are possible sequences in languages and direct the shape of a language throughout its history. Within Indo-European linguistics, the reconstruction of PIE syllabification is a long-standing problem, traditionally viewed either as a minor phonological phenomenon or as an unanswerable conundrum. This dissertation aims to dispel both of these notions, demonstrating that we may reconstruct syllabification in a non-circular fashion and that there were a number of phonological processes within PIE driven by violations of syllable structure.

Of course, before we investigate the problem at hand, we must first understand what methodological tools are at our disposal and review the past scholarship devoted to the study of IE syllabification. In this preliminary chapter we will begin by ex-

[1] See Watkins 2000:vii-xxxv, Meier-Brügger 2003 and Fortson 2010 for excellent overviews of Indo-European linguistics and the Indo-European languages.

amining the phonology of PIE assumed in this dissertation's analysis, with discussion restricted to key issues relevant for the reconstruction of PIE syllabification. We will then proceed to a brief discussion of the premises and benefits of Optimality Theory, the primary phonological framework used in this dissertation. Next, we will examine the main characteristics of syllable structure from a typological point of view, concluding with a summary of past views of syllabification in the IE literature.

1.2 Assumptions of Proto-Indo-European Phonology.

The segmental phonemic inventory of PIE assumed in this dissertation is, for the most part, uncontroversial. It closely resembles the views presented in Mayrhofer 1986 (unless explicitly stated otherwise), to which I refer the reader for all phonological matters not discussed here in depth.

To start, I assume the standard PIE inventory of vowels, as set forth in Mayrhofer 1986:90.

(1) PROTO-INDO-EUROPEAN VOWELS.

*i, (*ī)				*u, (*ū)
	*e, *ē	*ə	*o, *ō	
		*a, *ā		

One important difference with Mayrhofer 1986 is that I will use the vowel *ə in this dissertation to indicate both epenthesis in a consonant sequence with laryngeal (schwa primum)[2] and schwa secundum, as I assume *ə to be the go-to 'fix-it' vowel in PIE for any illicit consonant cluster that requires repair. For example, in my reconstructions *ə is given both in the oblique stem of 'father' (**pəh*$_2$*tr-*, with schwa primum), as well

[2]Also known as a 'vocalized' or 'syllabic' laryngeal.

as in forms such as *k^wtu̯ṓr 'four' → *k^wətu̯ṓr > Lat. *quattuor* 'four' and *$d^h\hat{g}^h$més 'earth, ground (gen.sg.)' → *d^hə$\hat{g}^h$més ↣ Hitt. *taknaš* 'id.' (with schwa secundum).

The rich array of PIE consonants presented in Mayrhofer 1986:91ff. will also be reconstructed in this dissertation. I assume the traditional threefold series of stops (voiceless, voiced and voiced aspirate),[3] three distinct sets of tectal consonants (palatal, velar and labiovelar)[4] and reconstruct three "laryngeal" consonants (*h_1, *h_2, *h_3).

(2) PROTO-INDO-EUROPEAN CONSONANTS.

	labial	dental	palatal	velar	labiovelar	postvelar
voiceless stop	*p	*t	*k̂	*k	*k^w	
voiced stop	*b	*d	*ĝ	*g	*g^w	
voiced aspirate	*b^h	*d^h	*$\hat{g}^h$	*g^h	*g^{wh}	
fricative		*s				*h_1, *h_2, *h_3
nasal	*m	*n				
liquid		*r, *l				
glide	*u̯		*i̯			

The phonetic status of the laryngeals is very important for a study of IE syllabification: whether these segments were syllabic or non-syllabic determines how many syllables are to be posited for a particular word that contains them. Although scholars are not entirely sure (and likely never will be) what the precise phonetic values of the laryngeals were, we can assume with some certainty that these sounds were phonemically consonants produced in the back of the mouth and throat, given the vowel coloring effects by *h_2 and *h_3 and their partial continuation within Anatolian as velar or pharyngeal

[3]In other words, I do not follow the glottalic theory. For a thoroughly convincing argument in favor of the traditional reconstruction, see Garrett 1998.

[4]See Melchert 1987 and Melchert 1989:23-32, who convincingly demonstrates that Luvian continued three distinct reflexes of the three tectal series: *zú-wa/i-ni-* 'dog' (PIE *k̂u̯on-), *karš-* 'cut' (PIE *kers-) and *kui-* 'who' (PIE *k^wi-).

obstruents. Due to the uncertainty of their place of articulation I have simply labeled them as ‘postvelar’ consonants, a term that encompasses the uvular, pharyngeal and glottal points of articulation (see Bessell 1993). In this dissertation I will not follow the arguments of Reynolds, West and Coleman 2000, who propose that the laryngeals were prosodically weak vowels.

As suggested by Kessler (n.d.), Weiss (2009:50), Keydana (forthcoming a) and many others,[5] it is very likely that the laryngeals were all fricatives of some sort. Kessler (ibid.) provides a number of arguments in favor of such a view. First, in the one branch where they are partially preserved, Anatolian, the laryngeals were continued as velar or pharyngeal fricatives.[6] Second, the laryngeals pattern with the one assured fricative in PIE, **s*, in the general root template $*\{s, h_x\}PRVR\{s, h_x\}P\{s, h_x\}$.[7] Should we assume the laryngeals to be fricatives, we may simply posit a template *FPRVRFPF. Third, the laryngeals were clearly less sonorous than the reconstructed sonorants,[8] given their resistance to “nucleification” ($*h_x \nrightarrow *\mathring{h}_x$) in sequences of the shape RH{C, #}, despite the very strong tendency within PIE to maximize onsets. Lastly, through the examination of possible phonotactic structures in PIE, we find that word-medially, there are only two attested medial consonant sequences of the shape *-PCO-: *-PsO-[9] and *-PHO-.[10] Since there are no reconstructable sequences of the shape *-PPO- (****-ek̂pto-*), nor of the shape *-PRO- (****-ek̂rto-*), the existence of *-PHO- as a possible medial cluster suggests that the PIE laryngeals were neither

[5] See Kessler, n.d. for references.

[6] Continued by Hittite, Luvian and Palaic *ḫ(ḫ)* (Melchert 1994: 21-22).

[7] See page viii for a complete list of symbols and abbreviations used in this dissertation.

[8] For discussion of sonority, see section 1.4 below.

[9] For example, *$*h_1$etsk̂é/ó-* ‘eat (iter.)’, **su̯ek̂sto-* ‘sixth’.

[10] Cf. virtual **u̯ed*$^h h_1$*sm̥* > Ved. *ávadhiṣam* ‘I slew’, **d*^{h}*ugh*$_2$*ter-* > Ved. *duhitár-* ‘daughter’, **h*$_2$*ak̂h*$_3$*tu-* > Skt. *aśitum* ‘to eat’.

stops nor sonorants — hence they were fricatives.[11]

1.2.1 Allophonic Variation.

Throughout this dissertation I will attempt to be as consistent as possible in denoting allophonic and morphophonemic variation, when such variation is known. This includes the following:

1. Sonorant Syllabification.

   PIE liquids (*r, *l*) and nasals (*m, *n*), which occupy the nucleus of a syllable will be indicated by the traditional mark of syllabification in IE studies, a circle under the consonant (*r̥, *l̥, *m̥, *n̥*). PIE high vowels *i, *u* which are situated outside of the nucleus will be marked with an 'arch' under the glide (*i̯, u̯*).

2. Laryngeal Coloring.

   If *h_2 or *h_3 stands adjacent to tautomorphemic *e*, that *e* is colored to *a* and *o*, respectively.[12] For example, I will use *$h_2$ak̑ros* 'sharp' (Gk. *ákros*, OIr. *ér* 'high, nobel', OLith. *ašras* 'sharp') 'at the top, end, edge' for *$h_2$ek̑ros*;[13] *doh₃*- (Gk. *dídōsi* 'gives', Lat. *dōnum* 'gift') for *deh₃*- 'give'.[14]

3. Laryngeal Feature Assimilation.

   In PIE, laryngeal features (voicing and aspiration) were licensed (contrastive) only before vowels and sonorants. If an underlying voiced or voiced aspirate obstruent precedes another obstruent, it assimilates its laryngeal features to the

[11] On the other hand, if Jasanoff 2003:77[37] is correct in assuming loss of *h_1 in the cluster *-Oh_1O-, then perhaps this gives weight to the idea that *h_1 was a glottal stop (see Kessler, n.d. for references), since the sequence *-PPO- was blocked in PIE.

[12] Mayrhofer 1986:132, 142.

[13] NIL 288.

[14] Mayrhofer 1986:101.

following consonant. Examples include: **ni-sd-ós* ‘nest’ → PIE **nizdós* (Skt. *nīḍá-*, Lat. *nīdus*, Eng. *nest*); **neg*^{w}- ‘become dark’ (LIV 449) + **-t-* + **-s* → **nék*^{w}*ts* ‘evening (gen.sg.)’; **nig*^{w}- ‘wash’ + *-tó-* → **nik*^{w}*tó-* > Skt. *niktá-* ‘washed’, Gk. *á-niptos* ‘unwashed’; **u̯eĝ*^{h}- + **-s-* → **u̯eḱs-* (Skt. *vakṣ-*, Cyp. *éwekse* ‘brought’, Lat. *vēxī* ‘I carried’).[15]

4. Final Voicing.

   In PIE, all final stops were realized as voiced in word-final position, as they are in Hittite (Melchert 1994:85) and Old Latin (Meillet-Vendryes 1968:146ff.). Though at first glance typologically bizarre, such a process is not entirely unheard of cross-linguistically: see Yu 2004 for synchronic and diachronic discussion of the same phenomenon in Lezgian, a Nakh-Daghestanian language. Examples include PIE */b^{h}eret/ → *[b^{h}ered] ‘carried (3sg. impfct)’ and PIE */ad/ → *[ad] ‘at’.

5. Laryngeal ‘Vocalization’.

   In PIE, there were ‘vocalized’ variants of the laryngeals, contrasting with their true consonantal allophones. As to exactly what is meant by laryngeal ‘vocalization’, I assume the epenthesis of a reduced vowel adjacent to the laryngeal consonant, with the schwa being preposed before the laryngeal in initial syllables (PIE **ph*$_{2}$*trés* ‘father (gen.sg.)’ → **pəh*$_{2}$*trés* > Lat. *patris*) but postposed elsewhere (PIE *d*^{h}*ugh*$_{2}$*tér-* ‘daughter’ > PInd. **d*^{h}*ug*^{h}*Hətár-* > **d*^{h}*ug*^{h}*ətár-* > **d*^{h}*uǰ*^{h}*itár-*> Ved. *duhitár-*; PIE **méĝh*$_{2}$ > PIIr. **máǰ*^{h}*Hə* > Ved. *máhi*).[16]

---

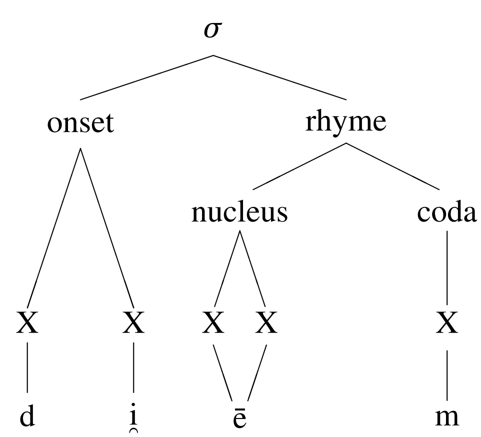
[15]Cf. Mayrhofer 1986:110. Thus, I will not assume that Bartholomae’s Law applied within PIE, though its status will not affect any of the matters discussed in this dissertation.

[16]Mayrhofer 1986:138. I admit that the lack of compensatory lengthening in the case of initial syllables poses a serious challenge for these assumptions and is a problem I hope to address in the near future. For a plausible alternative to schwa epenthesis in Proto-Indic, see Kobayashi 2004:132-4. Note that while I do agree with Mayrhofer about *where* schwa was epenthesized in PIE laryngeal clusters, I do not always follow his reconstructions of *when* schwa epenthesis occurred. For example, Mayrhofer assumes schwa epenthesis in the strong form of ‘daughter’ (**d*^{h}*ugh*$_{2}$*ətér-*), whereas I do not (**d*^{h}*ugh*$_{2}$*tér-*).

1.2.2 Reconstructing Consonant Clusters.

The final matter to be discussed regarding the reconstruction of PIE segmental phonology is the reconstruction of consonant clusters. In order to parse the syllable structure of any human language, we must first identify two factors: 1) what can occupy the syllable nucleus and 2) what can occupy the syllable margins. The syllable-initial margin is more commonly known as the onset and the syllable-final margin as the coda.[17] If the onset and coda are occupied by more than one consonant, then they are known as complex onsets and codas, respectively.

(3) SYLLABLE-INTERNAL STRUCTURE (PIE **di̯ḗm* 'sky (acc.sg.)').[18]



In order for us to reconstruct syllabification for PIE in a secure fashion, it is crucial that we have a clear conception of what is and what is not a consonant cluster in the proto-language. We may distinguish three types of reconstructed consonant clusters: those directly attested, those indirectly attested and those reconstructed for structural and/or etymological reasons.

[17] Blevins 1995:212ff.

[18] Based on Blevins 1995:213.

1. Directly attested.

   Language α (β, γ, etc.) has a cluster XY(Z), which we reconstruct as *XY(Z) (vel sim.) in the proto-language. This cluster need not be continued in all languages.

   This is the most secure type of reconstructable consonant cluster. Examples of directly attested clusters include:

WORD-INITIAL

_*tr-_: _*tréi̯es_ ‘three’ > Ved. _tráyaḥ_, Gk. _treĩs_, Eng. _three_

_*pt-_: _*ptero-_ ‘feather’ > Gk. _pterón_ ‘id.’, Arm. _tʻertʻ_ ‘leaf’

_*st-_: _*stah_$_2$_-_ ‘stand’ > Ved. _sthitá-_, Gk. _statós_, Eng. _stand_

*str-: _*streu̯-_ ‘spread’ > Lat. _struō_ ‘build’, OCS _-strujǫ_ ‘destroy’, Eng. _strew_

*pst-: _*psten-_ ‘breast, nipple’ > Av. _fštāna-_, Skt. _stána-_, OIce. _speni_[19]

WORD-MEDIAL

*-rsn-: _*persnV-_ ‘heel’ > Ved. _pā́rṣṇi-_, Gk. _ptérnē_, Lat. _perna_, Goth. _faírsna_

*-kst-: _*s(u̯)eksto-_ ‘sixth’ > Ved. _ṣaṣṭhá-_, Goth. _saihsta_, Lith. _šẽštas_

*-tsḱ-: _*h_$_1$_etsḱé/ó-_ ‘eat (iterative)’ > Hitt. _azzik-, azzak-_ /at͡sk-/

*-i̯str-: _*(h_$_1$_)oi̯stro/ah_$_2$ > Gk. _oístros_ ‘rage’, Lith. _aistrà_ ‘vehement passion’

WORD-FINAL

_*-ns_: _*-ons_ ‘_o_-stem acc.pl.’ > Goth. _-ans_ ‘id.’, Cret. _-ons_[20]

_*-ḱs_: _*h_$_3$_rēḱs_ ‘ruler’ > Ved. _rā́ṭ_, Lat. _rēks_, Gaul. _-rīx_ (_Dumnorix_, etc.)

---

[19] Cf. PIE _*pster-_ ‘sneeze’ (IEW 846-7).

[20] Cf. Cret. _tons eleut_^{h}_erons_ ‘the free men’

-k^wts: **né<u>k^wts</u>* 'evening (gen.sg.)' > Hitt. *nekuz* /nek^{w}t͡s/

**-rst*: *Virtual **d^hḗ<u>rst</u>* 'fastened (3sg. aor.)' > GAv. *dārəšt* 'held (3sg. aor.)'[21]

**-kst*: **h_2u̯é<u>kst</u>* > GAv. *vaxšt̰* 'made grow (3sg. aor.)'

2. Indirectly attested.

   If language α has the sequence X(Z) and language β has the sequence Y(Z), we may infer the existence of the cluster *XY(Z) or *YX(Z), if X does not originate from Y and vice versa.

   For example, we may reconstruct the onset **ksn-* for PIE in **ksneu̯s-* 'sneeze', which according to Pokorny (IEW 953) is a contamination of **pneu̯-* 'breathe' and **ksēu̯-* 'sneeze'. The root **ksneu̯s-* is reconstructable through comparison of PIran. **(k)sn-* in NPers. *išnōša, ašnōša* and Gmc. **hn-* (< **kn-*), continued by OIce. *hnjōsa*, OHG *niosan* (IEW 953).[22] A reconstructed ***skneu̯s-* may be ruled out because there are no reconstructable words of the shape **sPN-* in PIE, whereas the onset **ksn-* is clearly found in **ksneu̯-* 'sharpen' (LIV 373).

   In reconstructing PIE, a significant number of indirectly attested consonant clusters contain at least one laryngeal. If one language has VC or CV and another only C, the V is typically reconstructed as going back to an original laryngeal *h_x, part of an earlier PIE sequence *h_xC. For example, word-initially one finds such alternation with the 'prothetic vowel' of Greek, Armenian and Phrygian:[23]

---

[21] According to LIV 145, this *s*-aorist is an Avestan innovation, but I see no reason to doubt this cluster's existence within PIE.

[22] Eng. *sneeze* comes from OE *fnēosan*, from PIE **pneu̯-* 'breathe, gasp' (Southern 1999:71).

[23] These examples (and countless others) are found in Beekes 1969:18ff.

#C(C)-	#VC(C)-
Lat. *stella*, Skt. *str̥bhis*	Gk. *aster-*, Arm. *astł* ‘star’
Lat. *nōmen*, Ved. *nā́ma*	Lac. *enuma*, Arm. *anown*, OPhryg. *onoman* ‘name’
Goth. *riqis*	Gk. *érebos*, Arm. *erek* ‘darkness’

The contrast between vowel-initial forms in Greek, Armenian and Phrygian and vowel-less forms elsewhere has led scholars to reconstruct the forms as *$h_2$ster-* ‘star’, *$h_1$nóh$_3$mn̥*[24] ‘name’ and *$h_1$reg^{w}os* ‘darkness’,[25] respectively.

Were initial sequences of the shape *h_xC(C)-* true consonant clusters in PIE? They need not have been so: it is entirely conceivable that the *h_2 was ‘vocalized’ in *$h_2$ster-*, *[h_2əster-] or the like, whose vowel was continued in Greek, Armenian and Phrygian but deleted elsewhere in IE:[26]

$$\text{PIE *\#h}_x\text{əC(C)-} \begin{cases} \text{\#VC(C)- in Greek, Armenian and Phrygian} \\ \text{\#C(C)- elsewhere.} \end{cases}$$

This is not the most economical explanation, however. Should we view *$h_2$ster-* as having been phonetically *[h_2əster-] in PIE, we must assume independent laryngeal deletion and aphaeresis in seven of the ten main Indo-European Branches (Albanian, Balto-Slavic, Celtic, Germanic, Indo-Iranian, Italic and Tocharian), with Greek, Armenian and Phrygian deleting the initial laryngeal but retaining the inherited vocalized laryngeal as a short vowel.[27]

[24] Or similar. The true shape of this word is actually highly contentious, though few would dispute the presence of an initial laryngeal, which is the point here.

[25] See Olsen 1999:764 for an alternative explanation of Arm. *erek*.

[26] Cf. Keydana, forthcoming a.

[27] Anatolian, of course, retains the laryngeal (*ḫaštēr-*), likely with vowel epenthesis [ḫaštēr-] (Melchert 1994:111).

The second option is to assume that clusters of the shape *$\#h_x C(C)$- were originally bi- and tripartite consonant clusters in PIE.

$$\text{PIE } *\#\text{h}_x\text{C(C)-} \begin{cases} \text{\#VC(C)- in Greek, Armenian and Phrygian} \\ \text{\#C(C)- elsewhere.} \end{cases}$$

This is the simplest explanation, as it requires the fewest number of steps in the daughter languages. In Greek, Armenian and Phrygian, a schwa was epenthesized in the initial cluster *$\#h_x C(C)$-, with later deletion of pre-vocalic laryngeal: PIE *$\#h_x C(C)$- > *$\#h_x əC(C)$- > *$\#ə_x C(C)$-.[28] Whether this was an innovation shared between the three languages (not an unreasonable scenario, given their proximity to one another) or an independent innovation is immaterial for the topic at hand; the crucial point is that there was NO laryngeal vocalization in PIE in this particular consonant sequence.[29] In the other IE languages, initial *$\#h_x C(C)$- > *$C(C)$-, with simple deletion of the initial consonantal laryngeal.

In short, if a laryngeal is to be reconstructed in PIE due to the continuation of a vowel in one language but ø in another, and there is no independent evidence for a process of vowel deletion in the second language (or any other language with laryngeal deletion), then a consonantal *h_x should be reconstructed for PIE.[30]

To some readers, this observation might seem trivial. It does, however, hold important consequences for laryngeal clusters in other positions of the PIE word.

---

[28]Deletion of pre-vocalic laryngeal also occurred before 'true' vowels inherited from PIE: cf. **h_2aĝ-* 'drive' > Gk. *ágō*, Arm. *acem* 'I drive'. For Phrygian, cf. **$h_1$ei̯tōd* > *eitou* 'let it be(come)'.

[29]Laryngeal 'vocalization', as we will see in chapter 2, *did* exist as a synchronic phonological process in PIE – just not here.

[30]See Kessler, n.d. and Kobayashi 2004:132-3 for similar assumptions regarding PIE laryngeal 'vocalization'. Ringe (2006:137-8) finds evidence of deleted schwa primum (vocalized laryngeal) in PGmc., which would require us to assume laryngeal vocalization in early Proto-Germanic, followed by a later process of syncope.

For example, the PIE word ‘great (nt.nom./acc.sg.)’ is uncontroversially reconstructed as **méĝ*h_2 (NIL 468). This form is continued by Gk. *méga* and Skt. *máhi*, with laryngeal ‘vocalization’, and by Hitt. *mēg*, with laryngeal deletion. The lack of a vowel in the Hittite form points to an original consonant cluster *-*ĝ*h_2. Similarly, PIE *h_2anh_1mV- was pronounced with a tripartite medial cluster *-nh_1m-, given the deletion of laryngeal in GAv. *ą̇nman-* ‘soul, breath’ vs. laryngeal ‘vocalization’ in Gk. *ánemos*, Lat. *animus*, OIr. *anaimm*, etc.[31]

3. Reconstructed for structural or etymological reasons.

   The last, and least certain, type of reconstructed consonant cluster is that which is reconstructed purely for paradigmatic or etymological reasons. This type may be defined as follows:

   All languages show a particular word as having a consonant sequence Y(Z). However, the existence of a related form, either etymological or paradigmatic, contains at least one additional consonant, X. The presence of this consonant in this related form suggests its presence in the original form, *XY(Z).

   For example, Hitt. *išpant-* ‘night’ is related to Ved. *kṣáp-, kṣapā́* ‘night’, Av. *xšap-* ‘darkness’ and Gk. *pséphas*, *pséphos* ‘dark’ (IEW 649). This connection presumes that Hitt. *išpant-* was at some point *k^w*spent-* ‘night’. Our only evidence for this cluster is its etymological connection. However, a case could be made for it being a legal PIE onset, given this cluster’s structural similarity with the onset **pst-*, as found in PIE **pster-* ‘sneeze’ and **psten-* ‘breast, nipple’. Similarly, Gk. *bdéō* ‘fart’, which is related to Lat. *pēdō* ‘fart’ (< **pezd-*), Russ. *bzdětъ*[32] ‘fart quietly’ and Lith. *bìzdas* ‘anus’,[33] suggests an original PIE clus-

[31] See NIL 307 for discussion, with references.

[32] < PSlav. **pъzd-*.

[33] < **pizd-*.

ter *<u>bzd</u>-. However, since no language directly or indirectly attests the cluster *bzd*-,[34] it is not inconceivable that the onset **bzd*- was already simplified to **bd*- at a point within PIE. Again, independent evidence of the existence of its voiceless counterpart **pst*- renders it plausible that **bzd*- was a legal cluster in PIE.

On the other hand, the PIE word for 'comb', **pek̂-ten*- (Lat. *pecten*), a derivative of **pek̂*- 'comb' (Gk. *pékō*, Lith. *pešù* 'pluck, pull at the hair'), is frequently reconstructed as beginning with a tripartite onset **pk̂t*-: PIE **pk̂ten*- > Gk. *kteís*, *ktenós* 'comb'.[35] However, this reconstruction is highly unlikely for two reasons. First, if **pk̂t*- had been a legal PIE onset, it would be the only onset of the shape *PPP- reconstructable. Second, the cluster **pk̂t*- is not directly or indirectly attested in any IE language: there exists no attested ***pten*-, ***pk̂en*- or ***pək̂ten*- in another language alongside Gk. *kten*-.

In this dissertation I will consider all directly and indirectly attested consonant clusters (types 1 and 2) to be securely reconstructable for PIE. However, due to the frequently subjective and uncertain status of type 3 clusters, I will reconstruct these clusters only if they are extremely similar in shape with type 1 and 2 clusters.[36] For a complete list of reconstructed word-initial and word-final consonant clusters, I refer the reader to Appendix B.

---

[34] Seeing as the outcome of PIE **pst*- in Baltic is *sp*- (Lith. *spenỹs*, OPruss. *spenis*), one might expect the outcome of PIE **bzd*- to have been similar (Proto-Baltic **zb*-?). It is imaginable, of course, that PIE **bzd*- > **bizd*- in Proto-Baltic-Slavic, thereby making this sequence indirectly attested.

[35] Mayrhofer 1986:117. The derivation of Pashto *ẓ̌manj* 'comb' from **fšančī* (Morgenstierne 1929:199) is not assured; see Charpentier (1929:197) for an alternative explanation.

[36] By 'extremely similar', I mean that all features are identical save one minor difference, such as a difference in voicing and in place. For example, I find the reconstruction of **k^w spent*- 'night' and **bzd*- 'fart softly' as fairly plausible, as there exists only one difference in place and voicing with the reconstructable onset **pst*-. However, we must tread carefully here, as the onset ***mst*-, which also differs in one feature with **pst*- ([-nasal] → [+nasal]) is clearly impossible.

1.3 Optimality Theory.

This dissertation will employ the most widely used constraint-based phonological framework, Optimality Theory (OT), which proposes that grammars arise from the interaction of conflicting constraints (Prince & Smolensky 1993). OT formalizes the concept of 'conspiracies', or the triggering of one or more phonological rules by the avoidance of a single phonological structure.[37]

There are two basic types of constraints within OT: faithfulness constraints and markedness constraints. Faithfulness constraints require that the surface form (the output) be identical to the underlying form (the input) in some fashion. Markedness constraints place requirements on the structural well-formedness of the output. The interaction of these two types of constraints results in the winning candidate, or the most 'optimal' form. For example, let's assume two constraints within the grammars of German and English:

(4) SAMPLE CONSTRAINTS.

a. *FINALVOICE: No surface form may contain a final voiced obstruent. Assign one $*$ for each instance.

b. IDENT(VOICE): Corresponding input and output segments have identical values for the feature [voice]. Assign one $*$ for every change.

The first constraint, *FINALVOICE, is a markedness constraint that blocks voiced obstruents from occurring in absolute word-final position in the surface form. The second constraint, IDENT(VOICE), is a faithfulness constraint, which requires that the voicing of the surface form be identical (faithful) to the underlying form.

[37] Kisseberth 1970. For an introduction to Optimality Theory, see Kager 1999 and McCarthy 2008.

The ranking of these two constraints differs in German and English. In German, it is more important to avoid word-final voiced obstruents in the output than it is to be faithful to the voicing of the input, hence the ranking *FINALVOICE ≫ IDENT(VOICE). On the other hand, in English it is more important to be faithful to the voicing of the input than it is to avoid word-final voiced obstruents, hence the ranking IDENT(VOICE) ≫ *FINALVOICE. This is illustrated in derivations of the word for 'hand' in German and English.[38]

(5) GERMAN AND ENGLISH 'HAND'.

a. German:

/hand/	*FINALVOICE	IDENT(VOICE)
a. [hand]	*!	
b. ☞ [hant]		*

b. English:

/hænd/	IDENT(VOICE)	*FINALVOICE
a. ☞ [hænd]		*
b. [hænt]	*!	

1.3.1 Relevance to IE Phonology.

Let's now apply the OT framework to a fairly well understood phenomenon in PIE, the avoidance of geminates.[39] This avoidance is reflected by a number of phonological

[38] Note that ☞ indicates a winning candidate, ∗ a constraint violation, and ! follows a constraint violation that rules out a particular candidate as the most optimal.

[39] Cf. Meillet 1903 (1934):131, Kent 1953:26, Mayrhofer 1986:110-2 & 120, Szemerényi 1996:109-10, Kobayashi 2004:38. For typological comparison see Itô 1988:139ff.

processes either reconstructed for PIE or attested in one of the most archaic IE languages, Sanskrit. The advantage of utilizing an Optimality Theoretical framework in our reconstructions of PIE is that it provides us with tools that allow us to go beyond merely describing a process, by connecting specific phonological phenomena with the underlying tendencies of the PIE phonology as a whole.

1. ø → [s] / T_TV.[40]

   In PIE, when two adjacent dental stops immediately preceded a vowel, an *[s] was inserted between the two. Cf. *u̯id-tó- 'seen, known' → *u̯itsto- (> Germ. *ge-wiss* 'certain', Gk. *á-ïstos* 'unknown', etc.); *h_1ĕ̄d-ti* 'eats' → *h_1ĕ̄tsti* (> Hitt. *ēzzazzi* /ētstsi/, Welsh *ys*, etc.)

2. *T → ø / VT_RV or V_TRV.[41]

   In PIE, when two adjacent dental stops were followed by a sonorant, one dental was deleted with no compensatory lengthening.

   The attested evidence presents a conflicting picture of which dental was lost and whether voicing assimilation preceded or followed the loss of dental. For example, PGmc. **setlo-*[42], Lat. *sella* and Gaul. *sedlon*, all require that *-*t*- of the suffix was lost with the preceding stop having undergone no voicing assimilation: PIE **sed-tlo-* 'instrument of sitting' > **sedlo-*.[43] On the other hand, Gk. *métron* 'measure' (< PIE **métro-* < PIE **méd-tro-* 'instrument of measuring'),[44] Skt. *átra-* 'nourishment' (< **$h_1$ed-tro-*) and PGmc. **χerθra-* 'entrails, heart'

[40] The "double dental rule". Mayrhofer 1986:110-11; Hill 2003:3-7.

[41] The "*métron* rule". Saussure 1885:246ff.; Schindler apud Mayrhofer 1986:111.

[42] Goth. *sitls* 'bench', OE *setl*, OHG *sezzal*.

[43] Olsen 1988: 13. Skt. *sattra-* 'sacrifice with 12 or more pressing days in the Soma cult' and Av. *hastra-* 'a gathering where the sacrifice hymns were recited' (< **sed-tló-*) must be considered as secondary formations.

[44] If not from **méh$_1$-tro-* with loss of **h_1* by the 'Weather Rule'; see Peters 1999.

(< *$\hat{k}$erd-tro-)[45] imply the opposite. The *métron* rule will be discussed in more detail in section 3.6.

3. $*VR_1 \rightarrow [\bar{V}] / \_R_1\#$.

   In PIE, when two identical sonorants were adjacent at the end of a word, one sonorant was deleted with compensatory lengthening of the immediately preceding vowel.

   Cf. PIE **dṓm* 'house (acc.sg.)' < **dóm-m* (Arm. *tun* 'id.' < **dōm*), **dʰéǵʰōm* 'earth (acc.sg.)' < **dʰéǵʰom-m* (Schindler 1977b:31). Though only attested for **-mm*, I see no reason not to expect this simplification for all word-final geminate sonorants.

4. $*R_1 \rightarrow ø / V\_R_1V$.

   In PIE, when two identical sonorants were adjacent and immediately preceded a vowel, one sonorant was deleted with no compensatory lengthening.

   Cf. **ném-mn̥* 'gift' > **ném-n̥* > OIr. *neim* 'poison' (Rasmussen 1999:647) and perhaps **stomh₁mn̥* > **stommn̥* > **stómn̥* > Gk. *stóma* 'mouth' (Melchert, p.c.).[46] Just as with word-final **-mm* deletion, I see no reason not to expect this simplification for all word-medial geminate sonorants.

5. $*Vs \rightarrow [\bar{V}] / \_\ s\#$.[47] (possibly)

   In PIE, when two /s/ were adjacent at the end of a word, one /s/ was deleted with compensatory lengthening of the immediately preceding vowel.

---

[45] Feist 1939:235; Olsen 1988:21. Cf. Goth. *hairþram* 'entrails, heart', OHG *herdar* 'entrails', OE *hreþer* 'breast, stomach, heart'; cf. Lith. *kartóklys* 'omasum'.

[46] With loss of laryngeal by the Saussure-Hirt effect (Nussbaum 1997).

[47] First pointed out to me by Ron Kim, in an e-mail dating May 19, 2009, though already suggested by Szemerényi 1970:109. Though less common than in instances of sonorant loss, compensatory lengthening upon fricative loss is attested, likely via an intermediate stage of an approximant (Kavitskaya 2002:66-75). For example, cf. **izdem, *nizdos* > Lat. *īdem, nīdus*.

Cf. *$h_2$áu̯s-os-s* 'dawn (nom.sg.)' > **$h_2$áu̯sōs*. Perhaps also **mus-s* > **mūs* 'mouse (nom.sg.)' and **nas-s* > **nās* 'nose (nom.sg.)' (Szemerényi 1996:117; cf. IEW 755). The lengthened vowel in the suffix of **$h_2$áu̯sōs* is usually taken to be analogical to forms such as *d^héĝ^{h}ōm* (< *d^héĝ^{h}oms*), though this may be an unnecessary assumption. If loss of **s* in this sequence turns out not to invoke compensatory lengthening, rules (5) and (6) should be collapsed together.

6. *s → ø / V__ sV.[48]

   In PIE, when two /s/ were adjacent and immediately preceded a vowel, one /s/ was deleted with no compensatory lengthening.[49] Cf. **$h_1$és-si* 'you are' > **$h_1$ési* (Skt. *ási*, Gk. *éi*, etc.); **$h_2$us-s-és* 'dawn (gen.sg.)'> Skt. *uṣás*; **$h_2$us-s-ih_1* 'ear (nom. du.)' > **$h_2$usih_1* > Av. *uši*.[50]

7. Skt. /b/ → [d] / __ b^h.

   In Sanskrit, when a root-final labial stop preceded the instrumental plural case ending *-bhis*, that stop is realized as a dental. For instance, *ap-* 'water' + *-bhis* (instr.pl.) → *abbhis* → *adbhis*.

8. Skt. /s/ → [t] / __ + *-s-* (in certain morphological categories).

   In Sanskrit, when a root-final /s/ preceded the /s/ suffix of the future and aorist, the /s/ of the root became a [t]. For example, *vas-* 'dress oneself' + *-sya-* (future) → *vat-sya-*; *vas-* 'id.' + *-s-* (aorist) → *a-vā́t-s-* (Narten 1964:239-240).[51]

---

[48] Mayrhofer 1986:120-1.

[49] Note that in each case of word-medial geminate simplification, no compensatory lengthening occurs.

[50] Nussbaum 1986:132. Cf. OCS *uxo* (nom.sg.), *ušese* (gen.sg.) < **$h_2$us-es-* (IEW 785).

[51] Although processes (7) and (8) are not attested elsewhere in IE, the ban of heteromorphemic geminates is assured as a PIE phenomenon. Did PIE speakers utter forms, which were direct ancestors of the attested Skt. forms (**$h_2$adb^{h}is* 'waters (instr.pl)', **(e)u̯ḗtsm̥* 'I got dressed') or was there another repair strategy at work: **s*-epenthesis (**$h_2$abzb^{h}is*); deletion (**(e)u̯ḗsm̥*)?

Although each of the rules above differs in process, each is identical in its goal – to eliminate a sequence consisting of two of the same segment. We may hypothesize that an undominated constraint within the PIE grammar is the driving force behind each of the rules given: the OCP, or the avoidance of adjacent identical segments (McCarthy 1986).

(6) THE OBLIGATORY CONTOUR PRINCIPLE (OCP).
Two identical segments may not be adjacent to each other.[52]

The power of reconstructing the OCP for PIE is that this constraint allows us to explain each of the rules above as well as predict that heteromorphemic words such as *seh_1-h_1e*, **sek-kos* or **ser-ros* cannot occur in PIE, even though we currently have no evidence to prove this directly.

Typologically, it is very common for the OCP to be highly ranked in a language's grammar. For example, the OCP blocks geminate sequences in English, except across certain prosodic boundaries, such as in compounds like *penknife* [pɛn:aɪf] and heteromorphemic formations such as *solely* [sowl:i]. In PIE, we find the opposite situation. Whereas geminates were blocked across morpheme boundaries, tautomorphemic geminates were permitted in PIE: **atta* 'daddy',[53] **kakka* 'poo-poo',[54] **akka* 'momma'[55] and **anna* 'momma'.[56] Meillet (1934:132) suggests that these cases of gemination show "valeur expressive", a feature he convincingly argues to be a pan-IE

[52]More precisely formulated as: "At the melodic level, adjacent identical elements are prohibited" (McCarthy 1986:208).

[53]Gk. *átta*, Lat. *atta* ("both used as respectful forms of address for old men" [Ringe 2006:71]), Goth. *atta* 'daddy', Hitt. *attaš* 'father'.

[54]Gk. *kakkáō*, MIr. *caccaim*, Russ. *kákatъ*.

[55]Gk. *Akkṓ*, Lat. *Acca (Lārentia)*, Skt. *akkā* 'momma'.

[56]PAnat. **ánna* (Melchert 1994:147).

phenomenon.[57]

A number of scholars, beginning with Saussure (1885),[58] have contended that the avoidance of geminates in forms such as **u̯id-tó-* 'known', *méd-tro-* 'measurement' and **$h_1$és-si* 'you are' should be attributed to the fact that in PIE postvocalic consonants were pronounced ambisyllabically. Under this assumption, there actually was no deletion in words such as **méd-tro-*; rather, all words of the shape */VC_1C_2V/ were pronounced as *[VC_1.C_1C_2V]. The sequence **méttrom* was the same in pronunciation as **métrom*: both were realized as *[met.trom] with an ambisyllabic */t/.

This theory, however, creates more problems than it solves. First, it does not explain why we find deletion of a dental stop in words of the shape VTTRV (the *métron* rule) but not in the double dental rule, where there is **s*-epenthesis. Second, while the assumption of ambisyllabic consonants in PIE might explain the simplification of geminates in the *métron* rule, **$h_1$éssi* 'you are' and **némmn̥* 'gift', it does not demonstrate why there is geminate avoidance in other environments, such as the simplification of word-final geminates in **dṓmm* 'house (acc.sg.)' and **$h_2$áusoss* 'dawn (nom.sg.)'. Lastly, should we follow Saussure in assuming that postvocalic consonants were ambisyllabic in PIE, then underlying sequences of the shape *VC_1V and *VC_1C_1V would have been phonetically equal. Just as words of the shape */V$C_1C_1C_2$V/ must be reconstructed as *[VC_1C_2V] (*métro-*, *$h_1$etro-*), so should words of the shape */VC_1C_1V/ be reconstructed as *[VC_1V]. This is clearly false, since there was a distinction between

[57] As Craig Melchert points out to me, given the considerable amount of evidence presented in this section for a highly ranked OCP constraint in PIE, should the connection of Lat. *immō* with Hitt. *imma*, CLuv. *imma* and HLuv. *ima* 'indeed' be true (PIE **immoh_2* vel sim.), the PIE form would have to have been synchronically monomorphemic or a shared innovation. See Kloekhorst 2008:384 and de Vaan 2008:300 for recent discussion, with references.

[58] Followed by Meillet 1934:129-30, Hermann 1923:351ff. Schindler (apud Mayrhofer 1986:111-2) considered loss of dental in **méd-tro* as due to the typologically bizarre syllabification $[\text{mett}]_\sigma[\text{ro}]_\sigma$, with subsequent simplification of tautosyllabic geminate (likewise Keydana 2004:171). See chapters 3 and 4 for arguments against such an analysis.

monomorphemic words containing geminates, such as the well-attested **atta* 'daddy' (not ***ata*), and those words with a single intervocalic consonant, such as **éti* 'still'[59] and **ápo* 'away'.[60]

However, assuming a high-ranking OCP constraint within the PIE grammar that targets only heteromorphemic geminates presents no such problems. Moreover, we are now provided with the motivation for each instance of geminate avoidance above. In the OT derivation of each of these forms, we will need to assume three additional faithfulness constraints: MAX-T, DEP-/s/ and MAX-/s, R/, as defined in (7) below.

(7) OCP CONSTRAINTS.

a. MAX-T: Every dental stop in the input has a correspondent in the output. Assign one $*$ for each instance of deletion.

b. DEP-[s]: Every **s* in the output has a correspondent in the input. Assign one $*$ for each instance of epenthesis.

c. MAX-/s, R/: Every **s* and sonorant in the input has a correspondent in the output. Assign one $*$ for each instance of deletion.

The relative ranking of each constraint within the PIE grammar will determine which output form is most optimal. In each instance we of course find that the most optimal form avoids a geminate sequence. This is due to the undominated ranking of the OCP constraint. Each underlying form shows simplification of a geminate sequence, except in the case of the 'double dental rule', where words of the shape *VTTV undergo **s*-epenthesis. For this reason, we will need the constraint ranking MAX-T $\gg$ DEP-[s] $\gg$ MAX-/s, R/. The ranking MAX-T $\gg$ DEP-[s] ensures that dental stops (T) are preserved, though with necessary **s*-epenthesis in the output form to avoid violation of

[59] Ved. *áti*, Gk. *éti*, Lat. *et*, etc. (IEW 344).

[60] Ved. *ápa*, Gk. *ápo*, Lat. *ab*, etc. (IEW 53).

the undominated OCP constraint. In cases where there are underlying geminates that are not dental stops, the ranking Dep-[s] ≫ Max-/s, R/ is required, which results in simplification of the geminate sequence instead of *s-epenthesis. It is for this reason that the candidate ***némn̥*, and not *****némsmn̥***, is chosen. In short, in order for all of the correct forms to be chosen in the derivation, we require the constraint ranking to be: OCP ≫ Max-T ≫ Dep-[s] ≫ Max-/s, R/.

(8) The OCP.

			OCP	Max-T	Dep-[s]	Max-/s, R/
a.		u̯ittó-	∗!			
b.		u̯itó-		∗!		
c.	☞	u̯itstó-			∗	
d.		dómm	∗!			
e.	☞	dṓm				∗
f.		némmn̥	∗!			
g.		némsmn̥			∗!	
h.	☞	némn̥				∗
i.		h$_2$áu̯soss	∗!			
j.	☞	h$_2$áu̯sōs				∗
k.		h$_1$éssi	∗!			
l.	☞	h$_1$ési				∗

1.3.2 PIE Conspiracies.

To conclude this section, the assumption of a high-ranking OCP constraint within the PIE grammar explains why change is required in reconstructable (underlying) heteromorphemic geminate sequences and also predicts the ban of other, unattested geminate sequences. In this dissertation's study of IE syllabification, our goals will be similar, with our reconstructions aiming to explain the phonological process in question as well as to provide predictions for other phonological phenomena in PIE.

1.4 Theoretical Assumptions of the Syllable.

Let us now turn to the main focus of the dissertation: the reconstruction of syllabification. Broadly speaking, a syllable may be defined as an abstract mental construct through which speech segments are organized. Though it has been difficult to identify any acoustic or articulatory correlates in what speakers and linguists tend to think of as syllables, it has been claimed that many (but not all) syllables are accompanied by a chest pulse ("an individual burst of action by the expiratory muscles").[61] However, a number of phoneticians, beginning with Ladefoged 1967, have argued this account to be incorrect. In the absence of such a theory, no good acoustic or articulatory definition of the syllable exists, a fact which leads Ladefoged and Maddieson (1990:94) to suggest that perhaps the syllable should be viewed strictly as a phonological unit. In fact, almost all of the evidence cited in favor of the syllable as a true constituent is phonological in nature, found in cases where phonological rules and constraints may be more succinctly expressed through the assumption of an underlying syllable structure than without one. Blevins (1995:207) cites four such cases.[62]

First, the syllable may function as a domain for phonological rules and constraints,

[61] Devine & Stephens 1994:9-10, to which I refer the reader for discussion, with references.

[62] The theoretical assumptions of the syllable in this dissertation will follow those views of Blevins 1995, unless noted otherwise. Though it lies beyond the scope of this dissertation to examine the problem of PIE syllabification in light of cue-based phonological theory (as given in Steriade 1997, Côté 2000, etc.), I do hope to examine PIE syllabification within such a framework (and others) in the near future.

a domain which is "larger than the segment, smaller than the word, and contains exactly one sonority peak." For example in Classical Latin, stress is assigned based on the number of syllables in a word and the weight of the penult: to the first syllable if the word is monosyllabic or disyllabic; to the penultimate syllable if the word is trisyllabic or longer and the penult is heavy; and to the antepenultimate syllable in trisyllabic words or longer if the the penult is light.[63] In Cairene Arabic we find that emphasis (pharyngealization) may spread only tautosyllabically, and a consonant may lose its emphasis if it is resyllabified within another syllable. For example, there is emphasis spread within both the first and second syllables of *ṣ̌ạṇṭịṭ* 'purse' (underlying /šạntịt/), though the final *-t* loses its emphasis in *ṣ̌ạṇṭịt-i* 'my purse' and *ṣ̌ạṇṭịṭ is-sitt* 'purse of the lady', when it is realized as the onset of the following syllable. Since emphasis will spread only to segments within the same syllable, the final *-t* of *ṣ̌ạṇṭịṭ* 'purse' will only be realized as emphatic if cosyllabified with the preceding /ị/: /šạntịt/ → [ṣ̌ạṇ]$_\sigma$[ṭịṭ]$_\sigma$ but /šạntịt-i/ → [ṣ̌ạṇ]$_\sigma$[ṭị]$_\sigma$[ti]$_\sigma$.[64]

Second, certain phonological rules and constraints may target the edges of syllables (margins). One such example frequently cited is the process of aspiration; for example, in English, aspiration targets syllable-initial voiceless stops (Kahn 1980:73).[65]

Third, there are cases where the syllables themselves are targets of morphological rules and language games. One common morphological process that targets syllables is reduplication, a process which we will discuss in greater detail in section 3.3.6. As for

---

[63] Weiss 2009:110.

[64] Kenstowicz & Kisseberth 1979:260.

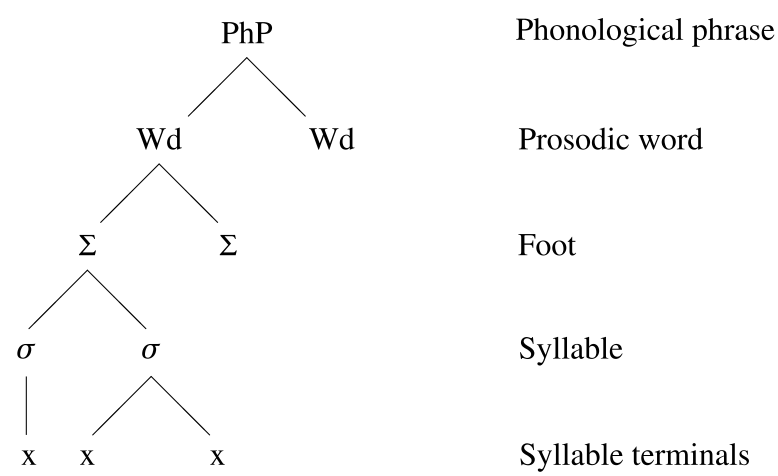
[65] One might claim that laryngeal feature neutralization (voicing and aspiration; see 1.2.1) was a process in PIE that targeted syllable codas, as reflected in **nig^{w}tós* → *[nik^w]$_\sigma$[tós]$_\sigma$ 'washed' and the generalization of voiced stops in word-final position (**b^{h}eret* → *[b^{h}e]$_\sigma$[red]$_\sigma$). However, this could not have been the case, since words such as **h$_2$aĝros* 'field (nom.sg.)' and **h$_2$ak̂ros* 'sharp (masc.nom.sg.)' were both syllabified as VC.CV (*[h$_2$aĝ]$_\sigma$[ros]$_\sigma$ and *[h$_2$ak̂]$_\sigma$[ros]$_\sigma$, respectively). If laryngeal neutralization had targeted all obstruents in coda-position, then *[h$_2$aĝ]$_\sigma$[ros]$_\sigma$ would have been automatically realized as *[h$_2$ak̂]$_\sigma$[ros]$_\sigma$.

language games, or ludlings, there are quite a few in English. The most well-known game is Pig Latin, where – depending on which ‘dialect’ of Pig Latin one speaks – the speaker will either fully or partially transpose a word-initial onset to the end of a word, adding *-ay*: *sleep* → *eep-slay* or *leep-say*; *strict* → *ict-stray*, etc.[66] Another such game is [ɪz] infixation, made popular by West Coast Gangsta Rap, in which [ɪz] is inserted in between the onset and rhyme: *sleep* → *slizzeep* (***sizzleep*); *strict* → *strizzict* (***stizzrict*, ***sizztrict*).

Lastly, native speakers have clear intuitions regarding syllables within their language. This fact is reflected by the very existence of syllabaries as writing systems in many of the world’s languages, such as Hittite, and the use of syllable counting and syllable weight in poetic meter, as seen in many of the oldest IE languages.[67]

In section 1.2.2 above I presented a hierarchical structure of the syllable, where the syllable node dominates the onset and rhyme, which in turns dominates the nucleus and coda. I will also assume that syllables are part of a fixed prosodic hierarchy.

(9) UNIVERSAL PROSODIC HIERARCHY.



[66]See Barlow 2001 for a recent discussion.

[67]Cf. Meillet 1923.

This hierarchy[68] will become important in our discussion of extrasyllabicity in the following chapters, where we will see that segments within certain PIE edge consonant clusters may be syllabified at a prosodic level higher than the syllable (most likely at the Wd level).

A distinguishing characteristic of the syllable is that it is very frequently organized with regards to the sonority of the segments within. Sonority is an important concept in the study of syllabification, as the more sonorous a segment is the more likely it is to function as a syllable peak, or the nucleus. Sonority may be defined as "[the] loudness relative to other sounds produced with the same input energy (i.e., with the same length, stress, pitch, velocity of airflow, muscular tension, etc.)".[69] Therefore in order for us to be able to identify individual syllables in PIE we must be able to identify syllable peaks, the most sonorous segments within the syllable.

Though sonority scales tend to differ slightly from language to language, certain hierarchies may be viewed as strong tendencies, if not universals. For example, we find voiced segments to be more sonorous than voiceless ones, vowels to be more sonorous than consonants and sonorants to be more sonorous than obstruents. This has led many to assume a universal sonority hierarchy, such as in Blevins 1995:211: low vowels > mid vowels > high vowels > glides > liquids > nasals > voiced fricatives > voiceless fricatives > voiced stops > voiceless stops. In this dissertation I will follow Kobayashi (2004:23) and Keydana (forthcoming a) in assuming a similar hierarchy to have been present in PIE.[70]

[68] Based on Blevins 1995:210.

[69] Blevins 1995:207.

[70] E = **ă*, **ĕ*, **ŏ*. Kobayashi (2004:23) is correct in observing that the reason the rightmost sonorant within a sequence of two adjacent unsyllabified sonorants is syllabified is due to a preference for onset maximization, not because glides, liquids and nasals were of equal sonority.

(10) PIE SONORITY HIERARCHY.

E > U̯ > L > N > F > P

As early as Sievers 1881, scholars have noted that cross-linguistically, sonority tends to decrease within a syllable when moving away from the syllable nucleus, a phenomenon typically referred to as the SONORITY SEQUENCING PRINCIPLE.[71]

(11) SONORITY SEQUENCING PRINCIPLE (SSP).

"Between any member of a syllable and the syllable peak, only sounds of higher sonority rank are permitted."[72]

To take some examples from PIE, a word such as **séms* 'one (masc.nom.sg.)' does not show an SSP violation, since the onset **s* rises in sonority to the syllable peak (the nucleus **é*) and then decreases in sonority into the coda (N > F). On the other hand, words such as PIE **stə*h_2*tós* 'having stood (masc.nom.sg.)' and **méĝ*h_2 'great (nt.nom./acc.sg.)' each illustrate one SSP violation. In **stə*h_2*tós*, there is an SSP violation in the onset, since **s* is of higher sonority than the following **t* (F > P); in **méĝ*h_2 there is also one SSP violation, this time in the coda, since laryngeals were of higher sonority than stops (F > P).

One last point to make regarding the syllable pertains to phonotactics. As Kobayashi 2004:17 rightly points out, "[U]nderstanding the synchronic restrictions on the syllable, i.e. figuring out what kind of syllable is well-formed or ill-formed for the speakers of a language in question, is a prerequisite for describing alternation patterns of segmental duration." By identifying what constituted a possible syllable within various stages of Sanskrit, Kobayashi has been able not only to parse word-medial syllable divisions but also to pinpoint what were possible nuclei and consonant

[71] Note that in this dissertation I will consider both sonority rises (**ste-*, **-ets*) and sonority plateaus (**pte-*, **-ept*) to be violations of the SSP.

[72] Clements 1990:284ff. Cf. Keydana 2004:164ff.; forthcoming a.

clusters, thereby giving us a more precise knowledge of Sanskrit phonotactics as a whole. In this dissertation I will apply Kobayashi's methodologies to my study of PIE, a methodology which I believe can give us a better understanding of its phonotactics.

1.5 Past Views of PIE Syllabification.

As we have seen in the previous section, syllabification may play a significant role in a language's phonology. In this section we will revisit the most notable treatments of Indo-European syllabification in past scholarship, addressing which views should be continued and which ones should be improved upon.

The first, and only, comprehensive comparative treatment of IE syllabification was published nearly a hundred years ago by Eduard Hermann (1923). The bulk of this book is devoted to a survey of the synchronic evidence for syllable structure in all of the major IE branches but Anatolian (Albanian, Armenian, Balto-Slavic, Celtic, Greek, Indo-Iranian, Italic, Germanic and Tocharian), concluding with a brief discussion of Hermann's views of PIE syllabification. However, the scope of his investigation of PIE syllabification was fairly narrow, focusing solely on the problem of syllable division. Undoubtedly his most significant finding is showing that every sequence of the shape *VCCV in PIE was syllabified as *[VC]$_\sigma$[CV]$_\sigma$, even those where *CCV formed a legal syllable onset: PIE *[put]$_\sigma$[los]$_\sigma$ 'son' (not **[pu]$_\sigma$[tlos]$_\sigma$), *[h$_1$es]$_\sigma$[ti]$_\sigma$ 'is' (not **[h$_1$e]$_\sigma$[sti]$_\sigma$), *[h$_2$r̥t]$_\sigma$[kos]$_\sigma$ 'bear' (not **[h$_2$r̥]$_\sigma$[tkos]$_\sigma$). Hermann (1923:351) backs up his hypothesis with a considerable amount of evidence from Greek, Italic, Celtic, Germanic and Indo-Iranian, and points to a tendency for closed syllables to become open within the history of the the majority of the Indo-European languages.

Many of Hermann's findings regarding PIE syllable division are upheld most recently in an important contribution by Keydana (2004), who gives one of the first up-to-date analyses of PIE phonology done within an OT framework. In this paper Keydana assumes the following syllable divisions in PIE:[73]

1. *VCCV → *$[\text{VC}]_\sigma[\text{CV}]_\sigma$
   PIE **u̯itto*- 'known' → *$[\text{u̯it}]_\sigma[\text{to}]_\sigma$ → *$[\text{u̯it}]_\sigma[\text{sto}]_\sigma$

2. *VCRV → *$[\text{VC}]_\sigma[\text{RV}]_\sigma$
   Gk. *métro*- 'measure' → $[\text{mét}]_\sigma[\text{ro}]_\sigma$

3. *VCCRV → *$[\text{VCC}]_\sigma[\text{RV}]_\sigma$
   Skt. *matsya*- 'fish' → $[\text{mats}]_\sigma[\text{ya}]_\sigma$; PIE **méttro*- → *$[\text{mett}]_\sigma[\text{ro}]_\sigma$ → *$[\text{met}]_\sigma[\text{ro}]_\sigma$

According to Keydana, these syllable divisions may be generated through the interaction of the following constraints:

(12) KEYDANA'S SYLLABIFICATION CONSTRAINTS.

a. *COMPLEXONSET: Onsets may not contain more than one consonant in the output. Assign one ∗ for each instance.

b. NOCODA: No syllable may have any consonants in the coda. Assign one ∗ for each instance.

c. *COMPLEXCODA: Codas may not contain more than one consonant in the output. Assign one ∗ for each instance.

In order to arrive at the syllable divisions he reconstructs for PIE, he assumes the following constraint ranking: *COMPLEXONSET ≫ NOCODA ≫ *COMPLEXCODA.

[73] See Keydana 2004:171ff. for discussion of *V{R,V}CCV sequences.

(13) PIE Syllabification (Keydana 2004)

/VCCV/, /VCRV/, /VCCRV/		*CompOns	NoCoda	*CompCoda
a.	V.CCV, V.CRV, VC.CRV	*!		
b. ☞	VC.CV, VC.RV, VCC.RV		*	(*)

Though Keydana's analysis nicely explains Hermann's reconstructions within an OT framework, it unfortunately encounters a number of exceptions. For instance, the parade example of PIE syllabification, */k̂u̯nbʰi̯s/ 'dogs (instr.pl.)', which is syllabified as $*[\text{k̂u̯n̥}]_{\sigma}[\text{bʰi̯s}]_{\sigma}$ (> Skt. *śvabhis*) fails his rankings.[74]

(14) PIE */k̂u̯nbʰi̯s/ (Keydana 2004)

*/k̂u̯nbʰi̯s/		*CompOns	NoCoda	*CompCoda
a. ☹	$*[\text{k̂u̯n̥}]_{\sigma}[\text{bʰis}]_{\sigma}$	*!		
b. ☞	$*[\text{k̂un}]_{\sigma}[\text{bʰis}]_{\sigma}$		*	

One may avoid these results by following Keydana's own suggestion (forthcoming, b) that there was a general tendency within PIE for coronal codas to be avoided in PIE, formalized through the reconstruction of a constraint *R/C 'Coronal sonorants are blocked in coda position'.[75] While the ranking *R/C ≫ *ComplexOnset will certainly produce the expected form $*[\text{k̂u̯n̥}]_{\sigma}[\text{bʰis}]_{\sigma}$, it still leaves instances such as **srutós* 'flowed',[76] **k̂litós* 'leaned'[77] and **k̂lutós* 'heard; famous'[78] unexplained.

[74]The ☹ indicates a candidate that doesn't win, but should.

[75]For further discussion, see section 5.2.1.2.

[76]Ved. *srutá-*, Gk. *rhutós*, Lith. *srutà* and Latv. *strauts* (IEW 1003).

[77]Ved. *śritá-*, Av. *srita-*, Gk. *klitós* (IEW 601).

[78]Ved. *śrutá-*, Gk. *klutós*, Lat. *inclutus*, OIr. *cloth* (IEW 605).

If simple codas were preferred over complex onsets (as given in Keydana's ranking *COMPLEXONSET ≫ NOCODA ≫ *COMPLEXCODA), then these forms should have been realized as ***sr̥u̯tós*, ***k̂l̥i̯tós* and ***k̂l̥u̯tós*, respectively.

Cluster division, of course, is not the only aspect of syllabification, nor is it the only one that has been investigated by scholars thus far. For example, we have known for quite some time what may act as a possible syllable nucleus in PIE — low vowels (**ĕ̄*, **ŏ̄*, **ă̄*), glides/high vowels (**ĭ̄*, **ŭ̄*), liquids (**r̥*, **l̥*) and nasals (**m̥*, **n̥*). We are fairly certain that fricatives and stops never behaved as sonority peaks,[79] unlike, for example, in Imdlawn Tashlhiyt Berber.[80] This is grounded in the fact that no IE language suggests a reconstruction with a syllabic obstruent, as there are no words of the shape *CO̥C inherited in any IE language. For instance, there is no suggestion that the initial sequence of **psten-* 'breast' was syllabified as **$[\text{ps̥}]_\sigma[\text{ten}]_\sigma$, the medial sequence of **s(u̯)ek̂sto-* 'sixth' as **$[\text{s(u̯)e}]_\sigma[\hat{\text{k}}\text{s̥}]_\sigma[\text{to}]_\sigma$ or the final sequence of **u̯ḗk̂st* 'carried' as **$[\text{u̯ḗ}\hat{\text{k}}]_\sigma[\text{st̥}]_\sigma$. All evidence suggests that these sequences were bona fide consonant clusters.

An important advance in our understanding of PIE syllabification was made by Meillet (1934:134-6), who first discussed the direction in which the rules of syllabification operate in PIE. In his *Introduction*, Meillet gives four distinct rules of the syllabification of sonorants in PIE, each based on the segments' surrounding environments:[81]

[79]The possibility of direct laryngeal vocalization aside.

[80]See Dell & Elmedlawi 1985.

[81]Meillet gives a fifth 'rule' of syllabification (or lack thereof): "A l'initiale : il n'y a pas de règle générale" ("Word initially, there is no general rule."). I will address his observation later on in this section.

1. First: “If a sequence of two sonorants follows a vowel or is in the first syllable of a word : the first is consonantal, the second syllabic.”[82]

   Examples

   a. PIE _*sru̯-tó-s_ ‘flowed (masc.nom.sg.)’ → _*srutós_ > Skt. _*srutáḥ_, Gk. _rhutós_

   b. PIE _*k̂u̯n-b_^{h}_i̯s_ ‘dog (instr.pl)’ → _*k̂u̯n̥b_^{h}_is_ > Skt. _śvábhiḥ_

   c. PIE _*g_^{wh}_rn-su̯_ ‘mind (loc.pl)’ → _*g_^{wh}_rn̥-sú_ ↣ Gk. _p_^{h}_rasí_ ‘mind; heart; diaphragm (dat.pl.)’ (IEW 496)

   d. PIE _*k_^{w}_etu̯r-to-s_ ‘fourth (masc.nom.sg.)’ → _*k_^{w}_etu̯r̥tos_ > OCS _četvrъtъ_

2. Second: “If a sequence of two sonorants follows a single consonant and precedes a vowel : the first is syllabic, the second consonantal.”[83]

   Examples

   a. PIE _*k̂u̯nés_ ‘dog (gen.sg.)’ → _*k̂unés_ ↣ Skt. _śúnaḥ_, Gk. _kunós_

   b. PIE _*k_^{w}_etu̯r-es_ ‘four (nom.pl.)’ → _*k_^{w}_etures_ > Lith. _keturì_ (cf. Skt. _catúraḥ_ ‘four (acc.pl.)’)

   c. PIE _*di̯u̯és_ ‘sky, sky god (gen.sg.)’ → _*diu̯és_ > Skt. _diváḥ_, Gk. _Di(w)ós_

   d. PIE _*ĝ_^{h}_i̯més_ ‘winter (gen.sg.)’ → _*ĝ_^{h}_imés_ > Av. _zimō_, Gk. _-k_^{h}_imos_, Skt. _himáḥ_

3. Third: “If a sequence of two sonorants follows a vowel and precedes either a consonant or the end of the word: the first is consonantal, the second syllabic."[84]

[82] “Entre deux consonnes après syllable brève ou dans la syllabe initiale du mot : la première sonante est consonne, la seconde voyelle.”

[83] “Entre consonne précédée de syllabe brève et voyelle : la première sonante est voyelle, la seconde consonne.”

[84] “Après voyelle, devant consonne ou à la fin du mot: la première sonante est consonne, la seconde voyelle.”

Examples

PIE **néu̯n* 'nine' → **néu̯n̥* > Skt. *náva*, Lat. *novem*, Gk. *enné(w)a*;

PIE **neu̯nti̯-* → **neu̯n̥ti-* > Skt. *navatí-*, Av. *navaiti-* 'ninety'.[85]

4. Fourth: "If a sequence of two sonorants stands between two vowels : the first forms the second half of a diphthong, the second is consonantal."[86]

   Examples

   a. PIE **oi̯u̯os* 'one' → **oi̯u̯os* > Av. *aiva*, Cyp. *oiwos* 'only' (cf. OLat. *oinos*, Goth. *ains*, etc.)

   b. PIE **de/oru̯V* 'oak (tree)' → **deru̯V* > Lith. *dervà*, OCS *drěvo*, Gaul. *derwen* 'oak', Gk. *dourós* 'tree; stick (gen.sg.)' (< **doru̯ós*)

Schindler (1977a:56) recognized a broader pattern within Meillet's rules of PIE syllabification and very elegantly collapsed all four observations into one phonological rule. He understood that in PIE, if given two adjacent sonorants that are potential syllable nuclei, the rightmost was always chosen, if it was not adjacent to a 'true' vowel (**e, *a, *o*, etc.). Schindler formulated his rule as follows:

(15) RULE OF PIE SYLLABIFICATION (SCHINDLER 1977A)

$$\begin{bmatrix} +son \\ -syll \end{bmatrix} \rightarrow [+syll] \,/ \left\{ \begin{matrix} -syll, \\ \# \end{matrix} \right\} \_\_ \left\{ \begin{matrix} -syll, \\ \# \end{matrix} \right\}$$

(iterative from right to left)

Since publication, Schindler's 'right-to-left' formulation has been widely accepted by nearly all scholars to date and has become the standard view of PIE syllabification.

[85] IEW 318.

[86] "Entre deux voyelles : la première sonante est second élément de diphtongue et l'autre est consonne."

However, Schindler himself (1977a:56-7) recognized there to be five instances where this rule does not correctly predict the syllabification reconstructable for PIE.

(16) EXCEPTIONS TO (15).

a. Roots of the shape *#R̯R̯-.

We find a number of roots with onsets of the shape *#R̯R̯- (*u̯i̯V-, *u̯lV-, *mn-, etc.)[87] which should be syllabified as *#R̥R̯- (*ui̯V-, *ulV-, *m̥n-, respectively).

b. Nasal-infixed presents.

The weak stems of nasal-infixed verbs also do not conform to (15).[88] For example, PIE */i̯u̯ngénti/ 'they yoke' →*[i̯un]$_\sigma$[gén]$_\sigma$[ti]$_\sigma$, not expected **[i]$_\sigma$[u̯n̥]$_\sigma$[gén]$_\sigma$[ti]$_\sigma$.

c. Accusatives in *-im, *-um(s), *-r̥m(s).

Here we must reconstruct PIE *méntim* 'mind (acc.sg.)', *s(e)uh$_x$num(s)* 'son(s) (acc.)' and *p(əh$_2$)tr̥ms* 'fathers (acc.)', not expected ***ménti̯m̥* 'mind', ***s(e)uh$_x$nu̯m̥(s)* 'son(s)' and ***p(əh$_2$)trm̥s*, respectively.

d. Word-medial *$_\sigma$[mn-.

Medial sequences of the shape *mn-* syllabify as such, with subsequent reduction to *m-* or *n-* by the *asno* law.[89] For example, *ĝ^{h}ei̯mnah$_2$* 'winter' → *[ĝ^{h}ei̯]$_\sigma$[mno]$_\sigma$ → *[ĝ^{h}ei̯]$_\sigma$[mo]$_\sigma$ > Lith. *žiemìnis*; *h$_2$ak̂mnés* 'stone (gen.sg.)' → *[h$_2$ak̂]$_\sigma$[mnés]$_\sigma$ → *[h$_2$ak̂]$_\sigma$[nés]$_\sigma$ > Skt. *aśnaḥ* 'id.'.[90], not expected **[ĝ^{h}e][i̯m̥]$_\sigma$[no]$_\sigma$, **[h$_2$a][k̂m̥]$_\sigma$[nés]$_\sigma$, respectively.

---

[87] See appendix B for a full list, with examples.

[88] See 5.2.1.1 for further discussion.

[89] Schmidt 1895:159.

[90] Mayrhofer 1986:159.

e. The sequence /CR_1R_2V/ is realized as [CR_1R̥$_2$]$_\sigma$[V]$_\sigma$, if within the same paradigm we also find the sequence /CR_1R_2C/ (→ [CR_1R̥$_2$]$_\sigma$[C-]$_\sigma$). Examples include **triōm*, not ***tr̥i̯ōm* and Skt. **-v(i)yās*, not ***-uyās*.

While the exceptional syllabifications found in (16e) may easily be attributed to analogy with other members of the paradigm, four exceptions still remain.

The last advance in our understanding of PIE syllabification is found in Kobayashi (2004:22), who correctly recognizes that the expression "right to left" in Schindler's formulation leads to overgeneration of nucleus placement. He suggests the following: "[i]f we can code the principle of minimizing the syllable coda in the procedure of nucleus placement itself, the use of such a directional expression will become unnecessary." Kobayashi's suggestion of coda minimization will be given in this dissertation in an opposite fashion: onset maximization.

(17) ONSET MAXIMIZATION (OM).

Syllabify as many consonants as possible within the onset.

This principle is neatly expressed through three constraints that Kobayashi assumes in the PIE grammar, HNUC, ALIGNNUC and ONSET.[91]

(18) KOBAYASHI'S SYLLABIFICATION CONSTRAINTS.[92]

a. "HNUC: When there is more than one segment which can become the nucleus of a syllable, the nucleus is assigned to the one with the highest sonority. In the case of PIE */k̂u̯nb^{h}is/ inst.pl. 'dog,' this constraint requires *u̯ to be the nucleus (> **k̂unb^{h}is); when, on the other hand, *n

[91]Frazier 2006:21 also assumes the principle of OM for PIE, through reconstruction of a high-ranking constraint NOCODA instead of ALIGNNUC. For our purposes, they are equivalent.

[92]Taken from Kobayashi 2004:23.

becomes the nucleus (> *k̂u̯n̥b^{h}is), it is counted as a violation of this constraint.

b. ALIGNNUC: ALIGN(Nucleus, R, σ, R): Align the right edge of a syllable nucleus with the right edge of a syllable, i.e. minimize codas.

c. ONSET: A segment to the left of a syllable nucleus is an onset; in other words, diereses are not allowed. The candidate *k̂u.n̥.b^{h}is (> **śuabhis), in which both the adjoining sonorants become the nuclei of two separate syllables to better satisfy ALIGNNUC, is ruled out by the constraint."

In order to produce the desired syllabification of the parade example */k̂u̯nb^{h}i̯s/ as *[k̂u̯n̥]$_{\sigma}$[b^{h}is]$_{\sigma}$, Kobayashi ranks these three constraints as follows: ONSET ≫ ALIGNNUC ≫ HNUC.

(19) PIE **k̂u̯nb^{h}i̯s* → *[k̂u̯n̥]$_{\sigma}$[b^{h}is]$_{\sigma}$

*/k̂u̯nb^{h}i̯s/		ONSET	ALIGNNUC	HNUC
a.	*[k̂un]$_{\sigma}$[b^{h}is]$_{\sigma}$		*!	
b. ☞	*[k̂u̯n̥]$_{\sigma}$[b^{h}is]$_{\sigma}$			*
c.	*[k̂u]$_{\sigma}$[n̥]$_{\sigma}$[b^{h}is]$_{\sigma}$	*!		

The advantage that Kobayashi's analysis has over previous analyses is that it explains two of the exceptions given in (16), complex *#RR- onsets such as **u̯i̯eh$_1$t* 'turned' (16a) and medial onsets of the shape **mn-* (16d).[93]

[93]Though syllabifications of the type *h$_2$ak̂mnés* 'stone (gen.sg.)' → *[h$_2$ak̂]$_{\sigma}$[mnés]$_{\sigma}$, with subsequent deletion to *[h$_2$ak̂]$_{\sigma}$[nés]$_{\sigma}$ (likely a postlexical rule; see section 4.3.1), are not generated by the constraints given in Kobayashi's analysis, we may safely view this type of syllabification as an instance of ONSET MAXIMIZATION.

(20) PIE *u̯i̯eh₁t → *[u̯i̯eh₁t]$_σ$

*/u̯i̯eh₁t/	ONSET	ALIGNNUC	HNUC
a. *[u]$_σ$[i̯eh₁t]$_σ$	*!		
b. ☞ *[u̯i̯eh₁t]$_σ$			

Because Kobayashi's ONSET MAXIMIZATION principle produces the same results as Schindler's 'right to left' formulation given in (15) and also provides a straightforward explanation for exceptions (16a) and (16d) above, this dissertation will follow the OM principle as the starting hypothesis for PIE syllabification.

1.6 Goals of Dissertation

Having now surveyed a number of past views of IE syllabification, it is clear that many problems are still waiting to be solved. Nevertheless, there remain scholars, especially within the "Leiden School", whose practice is not to indicate the syllabification (or more specifically, nucleus placement) of reconstructed PIE forms, arguing it to be "superfluous".[94] However, given the existence of the aforementioned exceptions to our understanding of Indo-European syllabification as well as reconstructable instances of lexicalized syllabifications (cf. PIE **kur-*, not expected ***kwr̥-*, perhaps seen in Hitt. *kūrkaš* 'foal' and Gk. *kúrnos* 'bastard'),[95] it is evident that the process of IE syllabification is NOT as straightforward as these scholars would lead us to believe. As Kobayashi (2004:21[5]) rightly points out, "Proto-Indo-European forms are not just a string of mechanically reconstructed symbols but are subject to phonological restrictions and well-formedness conditions, just like forms in ancient and modern lan-

[94] Beekes 1995:125.

[95] See Melchert 1994:132, with references.

guages." His insight into the problem is correct: we should always remember that PIE was a human language, and as such its speech sounds were organized into syllables.

In this dissertation I hope to answer a number of questions regarding Indo-European syllabification by examining each through a theoretical lens, incorporating research done on syllable structure in generative linguistics. The main problems I will address are the following:

- First, in addition to understanding *how* syllabification worked in PIE, is it possible to know *what* could be syllabified in PIE? This amounts to a study of margin (onset and coda) phonotactics: how many consonants can be placed in the onset and coda; is there a maximum?

- Are restrictions on all syllables the same in PIE, or does it matter where the syllable is situated within the word?

- Can we identify any phonological rules as being driven by violations of syllable structure?

- In the preceding section we saw contradictory views in scholars: there are those who believe that complex onsets were disfavored (Hermann, Keydana) and there are those who posit onset maximization (Kobayashi, Frazier). How may we reconcile Kobayashi's principle of ONSET MAXIMIZATION with the substantial amount of data in Hermann that points to the syllable division *VC.CV?

- Lastly, how can we explain the two remaining exceptions to Schindler's 'right-to-left' formulation: nasal-infixed presents and accusatives in *-*im,* *-*um(s),* *-*r̥ms*?

CHAPTER 2

Predicting Indo-European Syllabification through Phonotactic Analysis.

2.1 Overview.

In this chapter we will begin by reexamining the PIE laryngeal loss rule CHCC > CCC, accepting Hackstein's 2002 formulation.[1] This will lead into a general discussion of the methodology one needs to deduce the syllabification of a dead language and what such rules of syllabification may be able tell us about the PIE phonology in general. The methodology used in our reconstruction of PIE syllabification is called the DECOMPOSITION THEOREM (DT), the inference of medial syllabification based on possible edge clusters in a language. Following certain key insights of Steriade (1999), we will see that the DT is motivated by the universal tendency for speakers to construct medial syllable structure based largely (but not solely) on a speaker's knowledge of a language's word-edge phonotactics.

[1]Chapters 2 and 3 are revised and extended versions of Byrd, forthcoming c.

2.2 CHCC > CCC Revisited.

There are two variants of the word for 'daughter' reconstructable for PIE: one with a laryngeal, *d^hugh_2ter-, and one without, *d^hukter-.[2] The former is continued in numerous languages (Skt. *duhitár-*, GAv. *dugᵊdar-*, Gk. *$t^hugáter$-*, TA *ckācer*, TB *tkācer*), where we find the expected outcome of an interconsonantal *h_2, namely a vocalic reflex (Skt. *i*, Gk. *a* and Toch. *ā*)[3] or deletion (Av. *dugᵊdar-*). The latter form, *d^hukter-, is only unambiguously continued by Iranian **duxθrī-* (OPers. **duhçī-*) and **duxtar-* (NPers. *duxtar*),[4] Gaul. *duxtir*[5], Goth. *dáuhtar*[6] and possibly Osc. **fuutreí** 'girl (dat.sg.)',[7] Arm. *dowstr*[8] and HLuv. *t(u)watra/i-*, Lyc. *kbatra-* 'daughter',[9] since

---

[2]*/dʰugter/; */g/ is devoiced via laryngeal feature neutralization, as discussed in section 1.2.1.

[3]Mayrhofer 1986:136ff.

[4]On April 17, 2009 at the Sound of Indo-European conference in Copenhagen, Denmark, Agnes Korn kindly suggested to me that *duxt(a)r-* should rather be viewed as a reflex of YAv. *duɣδ(a)r-* by a process of fricative devoicing, as is seen in GAv. *aogᵊdā* 'speak (3rd sg. act. impfct.)' > **aoγδa* > YAv. *aoxta*. This is a possibility only if fricative devoicing may be demonstrated to have existed in all Iranian languages that attest to **duxt(a)r-*. However, according to Hoffmann-Forssman 2004:95, the replacement of expected **aoγδa* (< *aogᵊdā*) with *aoxta* is not phonological, but rather strictly analogical (cf. YAv. *saēta*, *staota* 'praised' with *-ta*; Hoffmann-Forssman 2004:204). Note that YAv. shows not *duxt(a)r-*, but rather *duɣδ(a)r-* (< *dugᵊdar-*), with expected fricativization of obstruent cluster. If *aoxta* were the result of a phonological rule of fricative devoicing, why does this not also occur in 'daughter'?

[5]If Zair 2009 is correct in postulating that laryngeals were regularly lost in the environment VC__ T{V/#} in Proto-Celtic, Gaul. *duxtir* would not provide conclusive evidence of CHCC > CCC in Celtic.

[6]Ringe 2006:138.

[7]< PItal. **fuχtrei̯* (de Vaan 2008:253). The normal outcome of a cluster ***-kt-** is **-ht-** in Oscan, not **-t-** with compensatory lengthening. For discussion of this form with references see Untermann 2000:306-7 and de Vaan 2008:253.

[8]See Hamp 1970, Clackson 1994:166-7, Olsen 1999:148,768 and Martirosyan 2010:244-5. CHCC > CCC may be assumed if one follows Muller 1984 and Beekes 1988:77, who propose that the normal development of medial laryngeals in Armenian is loss before a single consonant, vocalization before a cluster.

[9]Melchert 1994:69 assumes loss of *h_2 with later anaptyxis in the cluster **-gtr-*: *d^hugh_2tr- > **dugtr-* > *dug_etr- > **dugatr-*. Kloekhorst 2008:902-4 assumes a highly archaic PIE *d^hu̯egh$_2$tr-* > *d^hu̯egtr-* > *d^hu̯atr-*, with expected **g* deletion in Luvian and Lycian. In both scenarios, CHCC > CCC, though it remains unclear to me why */g/ would have not been realized as **k* in this position in Anatolian, as we see elsewhere in IE.

Slav. *dъ$\check{s}ti$,[10] and Lith. *duktẽ*[11] may be derived from either form.

2.2.1 Evidence and Past Scholarship.

Gernot Schmidt (1973) was the first to examine these two variants in detail. The focal point of his discussion deals with sorting out the multiplicity of forms in the Iranian languages. While Sanskrit only shows one variant (*duhitár-*), there are three in Iranian (Schmidt 1973:38): **dugdar-* (GAv. *dugᵊdar-*), **duxθrī* (Old Persian **duhçī-*, Modern Persian *došī-zah* 'young girl') and **duxtar-* (Modern Persion *duxtar*). Schmidt convincingly shows that the latter two forms must derive from a sequence ***d^h*ukt(e)r-*, since Bartholomae's Law does not occur in these forms (Schmidt 1973:54). Bartholomae's Law, which is often called the 'Buddha rule'[12] and is to be reconstructed back at least as early as Proto-Indo-Iranian (PIIr.), involves the progressive transfer of laryngeal features (voicing and aspiration) from a voiced aspirate onto the following voiceless segment (Kobayashi 2004:115ff.): PIE ***d^hugh_2ter*-* > PIIr. ***$d^hug^hh_2tar$*-* > PIr. ***$d^hug^hd^har$*-*> PIr. ***d^hugdar*-* > Av. *dugᵊdar-*.

To explain the two variants of 'daughter' reconstructed, Schmidt (1973) sets up the following linear phonological rule for PIE.

(21) CHCC > CCC (SCHMIDT)

$*H \rightarrow \emptyset / C \_\_ CC$

A laryngeal is lost in the second position of a sequence of four consonants.

Deletion presumably occurred in the oblique stem, which contained the sequence

---

[10]Derksen 2008:129.

[11]ibid.

[12]The form *buddha* derives from *budh-* 'be awake' + *-ta-*, a past (passive) participial suffix.

CHCC: *$d^h ugh_2 trés$* 'daughter (gen.sg.)' > *$d^h uktrés$*.[13] The oblique (or weak) stem differs from the strong stem in where the accent was located and which part of the stem had a vowel. For example, in the strong stem *$d^h ugh_2$-tér-*, accented **é* was found in the suffix, while in the oblique stem the accent and vowel were located in the ending (**$d^h ug(h_2)$-tr-és* 'daughter (gen.sg.)').[14]

Further instances of the CHCC > CCC rule may be seen in the following:[15]

1. Alternation between Skt. strong stem *janima* 'birth' < **ĝenh_1-mn̥* and oblique *janman-* < **ĝenh_1-mn-*, to which may be compared Dor. Gk. *génnā* 'descent' < **ĝén$h_1$mnah_2*, Skt. *jantú-* 'person' (< oblique **ĝenh_1-tu̯-*).[16] Root in question: PIE ***ĝenh_1-*** (IEW 373-5; LIV 163-5).

2. Lat. *verbum*, Hesych. *ért^{h}ei · p^{h}t^{h}éngetai* 'speaks' from **u̯erh_1-$d^h h_1$-o-*; cf. Gk. *rhẽma*, TA *wram* 'thing'. Lith. *vañdas* 'name' is ambiguous, as laryngeal loss in **u̯orh_1-$d^h h_1$-o-* may also be attributed to the Saussure-Hirt Effect (see Nussbaum 1997 and Yamazaki 2009, with references). Root in question: ***u̯erh_1-*** (IEW 1162-3; LIV 689-90).

[13]Peters (apud Mayrhofer 1981:436, 1986:138[172]) restricts Schmidt's laryngeal deletion rule to the sequence CHCCV$_{[+stress]}$:

(22) CHCC > CCC (Peters).
$H \rightarrow \emptyset$ / C __ CCV$_{[+stress]}$
A laryngeal is lost if it is in second position in a sequence of four consonants if the accent follows the cluster.

However, neither discussion provides any instances of laryngeal retention in the sequence V́CHCC-, and therefore I fail to see any benefit to such a modification of Schmidt's original rule.

[14]For more details on these alternations see Fortson 2010:79ff.

[15]Items (1), (2), (3), (4), (7) and (9) have been taken from Hackstein 2002, to which I refer the reader for additional (and at times less certain) examples.

[16]It seems unlikely that CHCC > CCC is reflected in Lat. *germen* 'seed' (***genimen*), though this possibility should not be ruled out entirely. See de 2008:Vaan 261 for discussion.

3. OIr. *fo·ceird* ‘places’, from **ker-d*^{h}*h*$_{1}$*-o-*; cf. Ved. *kiráti* ‘spreads, pours out’. Root in question: PIE ****kerH-*** ‘spread, pour out’ (IEW 933-5; LIV 353-4).

4. Hitt. *paltsḫa-* ‘pediment’, from PIE **pl̥th*$_{2}$*-s-h*$_{2}$*-ó-* (Melchert 1994:69) and TB *plätk-* ‘step forward’, from PIE *pl̥th*$_{2}$*-sḱé/ó-* ‘stretch out’ (via **pl̥tsḱé/ó-*); cf. Ved. *prathāná-* ‘spreading out’. Root in question: PIE ****pleth***$_{2}$**-** (IEW 833; LIV 486-7).

5. PIE **ĝ*^{h}*d*^{h}*i̯és* ‘yesterday’, from **ĝ*^{h}*h*$_{1}$*-d*^{h}*i̯-és*;[17] cf. Ved. *hyás*, Gk. *(e)k*^{h}*t*^{h}*és*, Lat. *herī*, OIr. *in-dé*, Wel. *doe*, Alb. *dje*, PGmc. **gestra-* (Goth. *gistra-dagis*, OE *geostra*, etc.). Root in question: ****ĝ***^{h}***(o)h***$_{1}$ ‘back, beyond’ (IEW 416).

6. Perhaps Proto-Celtic **sexskā/i-* ‘rushes, sedge’ (Matasović 2009:331), from **sekh*$_{x}$*skV-* (via **seksk-*). Cf. OIr. *seisc*, MWel. *hescenn*, MBret. *hesq*, etc. Root in question: PIE ****sekh***$_{x}$**-** ‘cut’ (IEW 895-6; LIV 524; Jasanoff 2003:80).

7. RV *mahnā́* ‘size, power (instr.sg.)’, perhaps from **meĝh*$_{2}$*mnéh*$_{1}$, with later **m* deletion by the *asno* law (NIL 473).[18] Stem in question: PIE ****meĝh***$_{2}$**-** (IEW 708ff.; NIL 468ff.).

8. MWel. *kyscaf* ‘sleeps’, from **kufské/ó-* < **kupské/ó-* < **kubh*$_{2}$*ské/ó-* (Schumacher 2000:87[74]); cf. Lat. *-cumbō* ‘lie down’. Root in question: PIE ****keu̯bh***$_{2}$**-** (IEW 590; LIV 357-8).

9. Lastly, perhaps the most compelling evidence for the rule in question comes primarily (but not exclusively) from Tocharian, through the continuation of the sequence **-d*^{h}*h*$_{1}$*sḱé/ó-* in certain words.[19] Root in question: PIE ****d***^{h}***eh***$_{1}$**-** (IEW 235-9; LIV 136-8):

---

[17] Proposed by Vine (2008:17).

[18] Schmidt 1895:159.

[19] Also from Hackstein 2002, who builds upon work done by Melchert 1977 and Jasanoff 1978:38.

a. TB *kätk-* 'arrange' < **k̂é tsk̂e/o-* < **k̂é* $d^h h_1$*sk̂e/o-*; cf. **k̂e/*k̂o-* 'here' (IEW 609-10) in Arm. *sa* 'this', Lat. *ce-do* 'gimme, hand it over', Eng. *he*, etc.

b. TB *wätk*^{a}- 'decide, command' < *u̯i tsk̂e/o-* < *u̯i* $d^h h_1$*-sk̂e/o-*; cf. Skt. *vidh-* 'allot, satisfy'.[20]

c. Toch. *kātk-* 'be happy' < **ga(*h_2*) tsk̂é/ó-* < **gah*$_2$ + $d^h h_1$*-sk̂é/ó-*; cf. Gk. *gēt*^{h}*éō* and Lat. *gaudēre* (< **gáh*$_2$ *u̯i* $d^h eh_1$-).

d. Lat. *suēscō* 'am accustomed' < **su̯é tsk̂e/o-* < **su̯é* + $d^h h_1$*-sk̂e/o-*; cf. Ved. *svadhā́* 'habitual state'.

Keydana (2004:172) has argued against the assumption of the CHCC > CCC rule on theoretical grounds, since, according to him, it would be impossible for both vowel epenthesis and consonant deletion to have occurred in such similar sequences in the same synchronic system using an OT framework (which, of course, we will also be using in this dissertation). His scenario, however, may be disputed on at least two grounds. First, as argued in section 1.2.1, it is not assured that the laryngeals were 'vocalized' in the sequence *VCHCV in PIE — in fact, it is simplest to assume that those language (sub-)families that vocalized, did so independently, as did those languages which deleted laryngeals in this sequence.

However, even if we were to assume that laryngeal vocalization (or epenthesis) 'fixed' intervocalic consonant clusters of the shape *-CHC- in PIE, it would still be unproblematic to set up a phonological rule *CHCC > *CCC within an OT framework, provided that the CHCC > CCC rule occurred at a different point in time than did vowel epenthesis in the sequence *-CHC-. We could accept this on blind faith, since there's no particular reason to believe that both rules occurred at the same point in time. Or, should we accept that the vocalization of *H in intervocalic *CHC sequences

[20] Hackstein 2002:8 also entertains the possibility that TB *wätk*^{a}- derives from PIE **uth*$_2$*-sk̂é/ó-*, a *-sk̂é/ó-* formation to **u̯eth*$_2$- 'say' (LIV 694-5), seen in OLat. *votāre* 'prohibit', OIr. *as:pena* 'witnesses'.

is einzelsprachlich, the archaicness of the *CHCC > *CCC rule relative thereto may be established, since the CHCC > CCC rule can be shown to have occurred both in dialects that 'vocalize' laryngeals in intervocalic *CHC clusters as well as in those dialects that do not. For example, we find that *CHCC > *CCC occurs in Proto-Italic (_*fuχtrei̯_ < _*d^{h}ug(h$_2$)tréi̯_; cf. de Vaan 2008:253), a language which vocalizes *VCHCV to *VCaCV: Lat. _animus_ < PItal. _*anamos_ < PIE _*h$_2$anh$_1$mos_. On the other side, we find evidence of *CHCC > *CCC in Anatolian (cf. Hitt. _palzaḫḫa-_ 'pediment' < _*pl̥tsh$_2$ó-_ < _*pl̥th$_2$sh$_2$ó-_[21] 'broad area'; Melchert 1994:69), which does not vocalize laryngeals in the intervocalic sequence *CHC (Melchert 1994:65). The sequence *-Ch$_{1/3}$C- undergoes deletion, not epenthesis, in _*u̯orh$_1$ĝent-_ > _úarkant-_ 'fat',[22] _*h$_1$ómh$_1$sei_ > Hitt. _ānši_ 'he wipes'[23] and likely _*h$_2$arh$_3$sei_ > _*ḫarši_ 'he furrows',[24] and *-Ch$_2$C- either retains its laryngeal or deletes it, depending on the cluster in question; cf. _*b^{h}érh$_2$ti_ > _parḫzi_ 'chases' with retention in the sequence *-Rh$_2$C-,[25] but with deletion in *-Th$_2$C-: _h$_2$u̯edh$_2$-i̯e/o-_ > Hitt. _ḫu̯etti(i)ya-_ 'pull, draw'.[26]

In order to accept Schmidt's CHCC > CCC rule for the protolanguage, we would therefore need to postulate that at the time the sequence CHCC was disfavored, the relevant constraint ranking was DEP-V ≫ MAX-C;[27] in other words, it was more im-

---

[21]This form is a 'transponat', one that likely never existed as such in PIE, but will be presented as a PIE form for expositional purposes.

[22]Kloekhorst 2008:81.

[23]cf. Kloekhorst 2008:72.

[24]Analogically replaced by _ḫaršzi_ according to Kloekhorst 2008:313. See also LIV 272-3, with references.

[25]Kloekhorst 2008:634-5 reconstructs _*b^{h}érh$_{2/3}$ti_. Setting the ambiguity of the laryngeal aside, for our purposes the reader should simply note that Hittite illustrates undeniable cases of laryngeal retention and deletion with NO inherited vowel epenthesis.

[26]Melchert 1994:65ff.

[27]Should one prefer to work with true 'vocalized' laryngeals (i.e. _*h̥$_x$_), one could substitute DEP-V with IDENT(σ) 'Do not alter the value of the feature [syllabic]. Assign one $*$ for every instance in the output where this feature has been altered.' See section 4.3.4.

portant for IE speakers not to epenthesize a vowel to fix the sequence CHCC than it was to delete a consonant. At a later date, however, be it in late PIE or within the individual languages, this constraint ranking was reversed, which allowed for vowel epenthesis (vocalization) to occur in complex sequences containing laryngeals in word-medial position.[28]

*CHCC ≫ Dep-V ≫ Max-C ≫ *VCHCV

(*[d^{h}uk]$_{\sigma}$[trés]$_{\sigma}$, *[d^{h}ug]$_{\sigma}$[h$_2$tḗr]$_{\sigma}$)

↓

*CHCC, *VCHCV ≫ Max-C ≫ Dep-V

(*[d^{h}u]$_{\sigma}$[gh$_2$ə]$_{\sigma}$[trés]$_{\sigma}$,[29] *[d^{h}u]$_{\sigma}$[gh$_2$ə]$_{\sigma}$[tḗr]$_{\sigma}$)

If, as I have suggested above, laryngeal vocalization did not occur in the PIE sequence *VCHCV, then only those languages which show vocalization, and not deletion, necessarily reversed the inherited PIE constraint ranking DEP-V ≫ MAX-C; language groups such as Balto-Slavic and Iranian retain the older configuration.[30]

2.2.2 Counterexamples.

Despite the many attractive examples of laryngeal loss in the environment CHCC, there are numerous counterexamples that contradict both Schmidt's (and Peters') CHCC >

[28]The constraints *CHCC and *VCHCV stand for 'The sequence *CHCC is prohibited in the output.' and 'The sequence *VCHCV is prohibited in the output.', respectively.

[29]Cf. Skt. *duhitúr*, Gk. *t^{h}ugatrós*. Of course, if a language continues the more archaic oblique form with laryngeal deletion, the form would remain as *[d^{h}uk]$_{\sigma}$[trés]$_{\sigma}$.

[30]As we will see in chapter 4, there is reason to believe that in PIE DEP-V ≫ MAX-C at the stem level of the grammar, but this constraint ranking was reversed at the postlexical level. It is therefore conceivable that the rule *H → *H̥ / VC __ CV originated as a postlexical process in PIE, which was only continued by certain languages. This would entail that the CHCC > CCC rule and the vocalization rule could actually have existed within the same time frame, though at different lexical levels, thereby bringing the number of possible scenarios under which the CHCC > CCC rule may have existed in the PIE grammar to three.

CCC rule, presented in (23) and (24), respectively (Hackstein 2002:10-11).[31] The counterexamples in (24) will be discussed in the next chapter in section 3.3.3.

(23) Counterexamples to CHCC > CCC: Word-Initial Position

a. *$d^h h_1$s-nó- > *fasno- > Lat. *fānum* 'temple'; cf. Skt. *dhiṣṇya-* 'pious; mindful'

b. *ph_2-sḱé/ó- > TB *pāsk-* 'protect, observe, retain'

c. *dah_2-swé → *dh_2-swé > Skt. *dīṣva* 'dole out!'

d. *sth_2-mn-ó- > OIr. *taman* 'treetrunk',[32] Gk. *stamnós* 'big drinking mug', TB *stām* 'tree', pl. *stāna*

e. *dh_1-mn-ó- > Gk. *démnia* 'bed, repository', Ion. *krḗdemnon* 'headband'

f. *$dh_3\hat{g}^h$mó- 'askew' > Gk. *dokʰmós*, Skt. *jihmá-* 'id.'[33]

g. *ph_2trés 'father (gen.sg.)' > Lat. *patris*, Skt. *pitúr*, etc.

(24) Counterexamples to CHCC > CCC: Word-Internal Position

a. *$\hat{k}erh_2$srom > Lat. *cerebrum* 'brain'

b. *$temh_x srah_2$ > Skt. *tamisrā*, Lat. *tenebrae* 'darkness'[34]

c. *$\hat{g}enh_1 trih_2$ > Lat. *genitrix*, Ved. *jánitrī-* 'bearer, mother'

d. *$\hat{g}enh_1 d^h$lo- > Gk. *génetʰlon* 'relative'

---

[31] Hackstein provides two additional examples, Toch. *nāsk-* 'spin, yarn' and Toch. *nāsk-* 'bathe, wash', which he proposes are derived from *snh_1sḱé/ó-, with a secondarily introduced consonant cluster CCHCC-. I have omitted both of these forms from my study, since neither form was syllabified as *snHsḱé/ó- (vs. *sn̥Hsḱé/ó-) in PIE and therefore did not possess a quadripartite consonant cluster CHCC, the structural description for Schmidt's laryngeal loss rule, as given in (21).

[32] For a different etymology, cf. Joseph 1982:36ff.

[33] < *ẑ́iẑ́ʰmá- < PInd. *diẑ́ʰmá-. See Beekes 1969:183, EWAia I:591 and IEW 181, with references.

[34] Cf. OHG *dinstar* 'dark', Lith. *tim̃sras* 'dark red', with expected (likely inner-dialectal) deletion (IEW 1064).

e. *$h_2$ar$h_3$trom > Gk. *arotron*, OIr. *arathar*,[35] Arm. *arawr* ‘plow’[36]

f. *(h_x)i̯en$h_2$trih_2- > Lat. *ianitrīcēs* ‘brothers’ wives’[37]

g. *tér$h_1$trom ‘auger’ > Gk. *téretron*, OIr. *tarathar*[38]

If CHCC > CCC is strictly a linear rule as per Schmidt, we would expect laryngeal loss in each of the examples in (23) and (24) above. While some of the examples may be explained by analogy, such as Skt. *dīṣva*[39] or Lat. *fānum*[40], it would be difficult to account for vowel epenthesis in an example such as *dh_3ĝ^{h}mo- ‘askew’ through a non-phonological process, since there is no other evidence for a root *doh_3ĝ- attested in IE.

2.2.3 Hackstein’s Syllable-Based Treatment of CHCC > CCC.

Hackstein (2002) noted that all of the exceptions in (23) are easily explained if one assumes a syllable boundary in Schmidt’s laryngeal loss rule, as given in (25).

(25) CH.CC > C.CC (HACKSTEIN)

*H → ø / C __ $CC

A post-consonantal laryngeal is lost at a syllable boundary before two consonants.

---

[35]Cf. OWel. *arater*, Wel. *aradr*, OCorn. *aradar*, MBret. *arazr*, Bret. *arar* ‘plow’.

[36]Lat. *arātrum* may or may not belong here, as this form could be a remodeling of an inherited **aratrom* or simply derive from the verb *arāre* ‘to plow’ (Weiss 2009:283). It is impossible to tell.

[37]Cf. Gk. *einatéres*, Arm. *nēr*, Ved. *yātar-*, etc. See NIL 204-7 for discussion.

[38]Also Wel. *taradr*, Corn. *tardar*, MBret. *tarazr* (Joseph 1982:41-2). Cf. Lat. *terebrā* ‘auger’ < *ter$h_1$srah_2 (Schrijver 1991:210).

[39]PIE *d$h_2$sw- > *tsu̯- (with voicing assimilation) → *d(ə)$h_2$su̯-, by analogy with such forms as *dah_2t ‘(s)he shared (aor.)’ (> Skt. *áva adāt* ‘(s)he divided’; LIV 103-4).

[40]One could conceivably argue for the following scenario: d^h$h_1$sno- > *tsno- → *d^h(ə)$h_1$sno-, with reinsertion of laryngeal from the full-grade stem *d^{h}eh_1s-, as continued by Osc.-Umb. **fíísnú** ‘temple (nom.sg.)’ (< *d^{h}eh_1s-; cf. Lat. *fēriae*, Osc. **fíísíaís** ‘holiday’); Untermann 2000:281-3.

Hackstein's reformulation nicely handles all of the counterexamples given in (23). In words such as *$d^h h_1$*sn*ó*- and *ph_2tr*é*s*, *H did not immediately precede a syllable boundary since it was a member of a tautosyllabic quadripartite consonant cluster.

Hackstein, however, does not discuss the conditions for syllabification in PIE or how one determines the location of the syllable boundary in the sequence CHCC. Moreover, why does laryngeal deletion only occur at a syllable boundary, whereas elsewhere we find vowel epenthesis (vocalization)? And lastly, and in my opinion most crucially, does the rule CH.CC > C.CC result from something inherently "bad" about the sequence CH.CC or is it the product of a more general phonological tendency within the PIE grammar? In the remainder of this chapter, we will develop a methodology that will provide us with a means to answer these questions.

2.3 General Guidelines to Reconstructing Indo-European Syllabification.

In Section 1.3.1, I argued that the assumption of ranked constraints in our analysis of the PIE phonology has immediate ramifications for our understanding of the PIE grammar, both in the explanation and connection of currently understood phonological processes as well as in the prediction of what was a possible word shape in the proto-language. For example, reconstructing a highly ranked OCP constraint for PIE explains not only the 'double dental rule' and the simplification of geminate */s/ in word-medial position, it also predicts that no word in PIE could contain a geminate sequence across a morpheme boundary, regardless of segment. Even though we have no direct evidence for a ban on words like **sel+los* or *tah_2+h_2os*,[41] we can infer

[41] This raises the question: what *did* speakers do when an underlying heteromorphemic geminate *h_2h_2 occurred in PIE, such as is reconstructed in **mólh₂h₂a*' 'I grinded' (Jasanoff 2003:89)? If one follows Jasanoff 1988:73ff. in assuming that PIE *-$óh_2h_2e$# > *-$óh_2u$#, then it appears that there was

their prohibition through the existence of a large number of phonological rules in PIE, which are most likely driven by a highly ranked OCP constraint.

Just as the reconstruction of a high-ranking OCP constraint in the PIE grammar provides us with both explanatory and predictive power, so should an optimal theory of PIE syllabification be able to explain (parse) the syllable structure of known PIE words and predict which sequences may be syllabified. As we have seen in section 1.4, syllabification plays an integral role in every phonological system of the world. Though we clearly understand what could function as a possible syllable nucleus in PIE, we currently have no way of broadly gauging what constituted a possible onset or coda in PIE. In order to predict accurately how many and what kinds of consonants were allowed in multipartite onsets and codas in PIE – and more generally how the process of syllabification worked in PIE – we must develop a systematic methodology. It should first and foremost observe universal (or near-universal) characteristics of the syllable as set forth in section 1.4, since Proto-Indo-European was at one point in time a living human language and behaved as such. This methodology should also be non-circular. We should not devise a theory of syllabification based solely on one's explanation of certain phonological rules, as many have done in the past.

2.3.1 Deducing Indo-European Syllabification.

Although we presently lack a means of determining PIE syllabification, we have seen that laryngeal deletion in the sequence CHCC *is* better explained if a syllable boundary is introduced, as per Hackstein. So why do we find deletion in the sequence CH.CC in medial position? Let us tentatively posit that the consonant sequence did not consist of a legal coda plus a legal onset. If *h_2 was a consonant and not [+syllabic]

---

a general simplification of (at least heteromorphemic) medial sequences *-Vh_2h_2V- in PIE, once again without subsequent compensatory lengthening (cf. section 1.3.1 above).

(cf. Mayrhofer 1986:121ff.) in the medial sequence CHCC, *d^h*ugh*$_2$*trés* 'daughter (gen.sg.)' must have been divided into two syllables, since there were only two possible syllable nuclei within the word.

We may immediately rule out **gh*$_2$*tr-* and **h*$_2$*tr-* as onsets of the second syllable, as they are not onsets reconstructable for PIE. **tr-* and **r-* were acceptable onsets in PIE (cf. *tréyes* 'three (nom.pl.)', **h*$_2$*n̥.rés* 'man (gen.sg.)'),[42] though it is debatable whether **r-* appears word-initially.[43] The most natural syllable division would have occurred after *-*gh*$_2$ and before **tr-*, since the onset of the second syllable would have been maximized and the syllable boundary would also coincide with the morpheme boundary (if not *d^h*ug-h*$_2$*tr-*).[44] Thus it is likely that 'daughter' was syllabified as *d^h*ugh*$_2$*.tr-*. A similar coda is also seen in **méĝh*$_2$ 'great' (Gk. *méga*, Skt. *máhi*, Hitt. *mēg*), where we find either vowel epenthesis (**méĝh*$_2$*ə* in Gk., Skt.) or laryngeal deletion (**mēg* in Hitt.) in the daughter languages. Holding off discussion of **méĝh*$_2$ until the next chapter, for the time being I'll assume that *-*gh*$_2$ was not a legal coda in PIE, and this is what prompts deletion in the form in question: *d^h*ugh*$_2$*trés* → *d^h*uktrés*.

Examples of *CHCC in (23) above undergo vowel epenthesis (vocalization): Lat. *fānum* 'temple' < **fasnom* < *d^h*əh*$_1$*snó-* ← *d^h*h*$_1$*snó-*. If deletion, and not epenthesis, had taken place in order to fix this 'bad' cluster *CHCC, the result would have been **tsnó-*, which would not have been a legal word, as this sequence in word-initial position is not reconstructable for PIE. Similarly, the reason deletion does not occur in

[42] **h*$_2$*n̥rés* (and not **h*$_2$*nrés*) should be reconstructed as the original syllabification for two reasons. First, no daughter language indicates that *#*h*$_x$*nr-* was a legal onset in PIE. Second, the less complex onset *#*nr-* is also not reconstructable for PIE, a prerequisite for a more complex *#*h*$_x$*nr-*, following the Substring Generalization, which states that "all substrings of a well-formed onset or coda should themselves be well-formed"; see Hammond 1999:54, following Greenberg 1978:250. Gk. *andrós* 'man (gen.sg.)' therefore should be reconstructed as **h*$_2$*n̥.rós* (← **h*$_2$*n̥.rés*).

[43] Weiss 2009:38.

[44] Cf. Pinault 2005 [2006].

the oblique stem of 'father' is because if *#ph_2tr- had reduced to *#ptr-, it would have resulted in the sequence *#PPR-.[45] This was the classic environment for the epenthesis of schwa secundum, a process of vowel epenthesis well-attested in many IE languages; cf. *mĝnó- > *məĝnó- > Lat. *magnus* 'great', *$d^h ĝ^h$més > *$d^h əĝ^h$més > Hitt. *taknas* 'earth (gen.sg.)',[46] *k^wtu̯ōr > *k^wətu̯ōr > Lat. *quattuor* 'four' and *ptnah$_2$- > *pət-na-h_2- > Gk. *pítnēmi*, Lat. *pandō* (< *patnō*) 'fly' (Mayrhofer 1986:118). We may therefore say that epenthesis occurs in the sequence *CHC(C)- if deletion would have resulted in a bad cluster.[47]

(26) Conditions for PIE Laryngeal Cluster Repair:
Delete a laryngeal in a bad cluster if the result would produce a legal consonant sequence; otherwise, insert a schwa.

2.3.2 Proto-Indo-European 'father'.

It has been claimed by Schmidt (1973) that laryngeal deletion occurred in the oblique stem of 'father', *ph_2tr-, to produce *ptr-, which is reflected by the Avestan oblique stem *fəδr-* 'father'.[48] However, as Insler (1971:573[2]) and Beekes (1988:86-7) contend,

[45] Where R = /*m, *n, *l, *r, *u̯/. Cf. Mayrhofer 1986:175-7.

[46] Schindler 1977b:31.

[47] Likewise the onsets **psk̂-, **tsu̯-, **stmn- are not reconstructable for PIE, providing an explanation for epenthesis in *ph_2-sk̂é/ó-, *d̂h_2-su̯é and *sth_2-mn-ó-. On the other hand, *$ĝ^h d^h$i̯- was a legal word-initial onset in PIE (we know this since it's reconstructable), and so in the form *$ĝ^h h_1 d^h$i̯és 'yesterday', deletion of the laryngeal was permitted to occur (following Vine 2008).

[48] Schmidt also posits laryngeal deletion in PIE 'father's brother': *ph_2tru̯i̯o- > *ptru̯i̯o-. He proposes that *ptru̯i̯o- is continued by YAv. *tūriia-* 'father's brother' and Proto-Slavic *stru̯i̯o- 'id.', while Skt. *pitr̥vyà-*, Lat. *patruus*, OHG *fatureo* all go back to secondary *ph_2tru̯i̯o-/*ph_2tr̥u̯i̯o-, with *h_2 analogically restored into the root. Kortlandt (1982) has been severely (and, in my opinion, convincingly) critical of the derivation of Slavic *str- from *ptr-, thus making the Avestan form the only possible evidence for a cluster *ptr- (< *ph_2tr-) in 'father's brother'. Note, however, that there is no trace of *p in YAv. *tūriia-* 'father's brother'; this suggests that the form should be derived directly from *(p)tr̥u̯i̯o- (cf. Hoffmann & Forssman 2004:94), with the expected simplification of a bipartite obstruent onset before a syllabic resonant (*#PPR̥ > *#PR̥; see Schindler 1977b:31f.). Perhaps *h_2 was deleted by rule (26),

the Avestan form is more likely to have been produced by analogy with the strong stem *ptar-*. This is confirmed by rule (26) above, which allows for the sequence **pt-* to have originated only in the strong stem, where there had been a sequence **ph$_2$tV-*. The original paradigm for 'father' was therefore: strong stem **ptér-*; weak stem **pəh$_2$tr-́*.

The cluster *#CHC- had two possible fates in PIE: *#CC- or *#CəHC-. The former, *#CC, is the expected regular outcome of the phonological rule given in (26), if the resulting cluster was legal: **ph$_2$tér-* > **ptér-* > Av. *(p)tar-*.[49] We know the onset **pt-* to have been legal in PIE through the reconstruction of words such as **ptero-* 'wing' (Gk. *pterón* 'wing', Arm. *t'ert'* 'leaf, foliage'). The latter outcome, *#CəHC-, arose in two ways, either by sound law or by analogy. If the deletion of *H in *#CHC- resulted in an illegal cluster, then schwa epenthesis occurs as the expected regular outcome of (26): **d^{h}h$_1$só-* 'possessing the divine' > PIE **d^{h}əh$_1$só-*, not ***tsó-*. The onset **ts-* is not reconstructable for any PIE word and it is for this reason that laryngeal deletion was blocked. PIE **d^{h}əh$_1$só-* produces Gk. *t^{h}eós* 'god' as well as Proto-Anatolian **dasó-* 'votive offering', continued by HLuv. *tasan-za* 'votive stele', Lyc. *θθẽ* 'altar' and Lyd. *taśẽṽ* 'votive object' (Melchert 1997:49-50; Watkins 2008:139-40.).[50] These

producing a legal onset **pt-*, which was subsequentally reduced by the rule **#TTR̥* > **#TR̥*: **ph$_2$tr̥u̯i̯o-* > **ptr̥u̯i̯o-* > **tr̥u̯i̯o-* > *tūriia-* (cf. Mayrhofer 1986:138[172]). I am indebted to Marek Majer for references and helpful discussion on the Slavic material.

[49] Av. *(p)tar-* 'father' shows the only instance of laryngeal deletion in a word-initial *#CHC- sequence in Iranian (Hoffmann & Forssman 2004:81-2), which perhaps may be considered a sign of its antiquity. Contrast GAv. *sīšā* 'show! teach!' (~ Ved. *śiṣánt-* 'instructing') < **#k̂h$_x$s-* (LIV 318-9).

[50] In Byrd, forthcoming a, I propose that the Hittite words for 'dream' (*tesh̯a-/zash̯(a)i-*) ultimately derive from the PIE root/stem **d^{h}eh$_1$s-* 'divine, divinity', formed to an earlier *o*-stem **d^{h}h$_1$s-h$_2$-ó-* '(possessing) the divine; divination (of the night)'. The original oxytone *i*-stem noun, *zash̯(a)i-* (phonetically *tsh̯í-*; Rieken apud Hoffner & Melchert 2008:47), illustrates the deletion of **h$_1$* in the sequence CHCC. Of course, as we have seen, laryngeals were lost in the sequence CHCC only if the resulting outcome had been a legal cluster in PIE. The question, then, is whether **Th$_1$sh$_2$í-* would have produced a legal sequence **tsh$_2$í-* in PIE (thence Hitt. *tsh̯i-*), or, to avoid an illicit onset, would rather have undergone schwa insertion (**Təh$_1$sh$_2$í-*).

The answer to this is by no means straightforward, as one finds conflicting forms in the attested IE languages. For example, both Hitt. *zikke-/zaske-* /tske-/ 'put (iter.)' and Toch. *tāskmāṃ* 'similar to' (Hackstein 1995:189) are derived from **d^{h}h$_1$sk̂é/ó-*, the former with laryngeal deletion and the latter epenthesis. Nevertheless, following the well-established Substring Generalization (Greenberg particular Anatolian forms assure vowel epenthesis in this sequence as a PIE process, since as Melchert (1994:65) argues, *h_1 invariably disappears without a trace in Anatolian. Secondly, *#CəHC arises when the laryngeal was reinserted into the cluster via analogy: *ph$_2$ter- > *pter- → *pəh$_2$ter- > Gk. *patér-*. Crucially, when a laryngeal was restored analogically it was also accompanied by schwa epenthesis – otherwise it could not have been syllabified.

In short, the onset cluster *#CHC- was not tolerated in any fashion in PIE and had to be 'fixed', either through a rule of laryngeal deletion or through a rule of schwa epenthesis.[51] These conditions, as they apply to the oblique forms of our two familial words in question, *d^{h}ugh$_2$trés 'daughter (gen.sg.)' and *ph$_2$trés 'father (gen.sg.)', may be formulated through the following constraints:

(27) CONSTRAINTS FOR LARYNGEAL CLUSTER CONDITIONS

a. PARSE: Syllabify all segments. Assign a $*$ for every violation.

1978:250), if *tsk̂- and *tsh$_2$- had been legal onsets in PIE, we would expect *ts-, sk̂-, and *sh$_2$- to have been legal onsets as well. The latter two can be established as onsets (cf. *sk̂eh$_x$(i̯)- 'shine' and *sh$_2$au̯- 'rain'), but *ts-* cannot. Thus, if *ts- is not an onset reconstructable for PIE, we would predict that *tsk̂- and *tsh$_2$- are not legal PIE onsets, either. In short, */Th$_1$sh$_2$í-/ was most likely realized as *[Təh$_1$sh$_2$í-]. Of course, a PIE form with schwa epenthesis does not account for either Hitt. *zikke-/zaske-* or *zasḫ(a)i-*. But *zasḫ(a)i-* is not a form reconstructable back to PIE – it is only attested in Hittite – and so we must assume that rules within a proto-Hittite grammar produced *zasḫ(a)i-*, not rules within PIE. Note that while *ts- was not a legal onset in PIE, it was in Hittite, as an affricate /ts/. This, I believe, provided Hittite speakers with the option of deleting *H in a sequence *THsC, be it one inherited (though morphologically renewed) as in the case of *d^{h}h$_1$sh$_2$é/ó- 'put (iterative)' or one newly formed within the prehistory of Hittite, *Th$_1$sh$_2$í- 'dream'.

On the other hand, if Melchert (2003) is correct in following Merlingen 1957:51 in the assumption that thorn clusters, phonemically *TK-, should undergo a process of affrication or s-epenthesis (*TK > *TsK) as continued by CLuv. *īnzagan=za* 'inhumations' < *en d^{h}z(^{h})ĝ^{h}ōm, then it is conceivable that PIE *d^{h}h$_1$sk̂é/ó- did simplify to *tské/ó-, which was inherited by Hittite but morphologically renewed in Tocharian.

[51] If Melchert (1994:175) is correct in deriving Hitt. *išḫunawar* 'sinew' from PIE *sh$_2$nóh$_1$u̯r̥ (with a tripartite onset), perhaps the ban of *#CHC should be more narrowly viewed as a ban of *#PHC (vel sim.). Kloekhorst (2008:395-6), however, denies the existence of any such form meaning 'sinew' and contests that *išḫunau-* only means 'upper arm'. Following Weitenberg 1984:224-5, he reconstructs *sh$_2$u-neu-, originally 'throwing arm', which would therefore provide no evidence for a PIE tripartite onset *sh$_2$n-. For now I will leave this question open for future investigation.

b. *CHCC: The sequence *CHCC is prohibited in the output. Assign a $*$ for every violation.

c. $*_\sigma$[PPR: The sequence two stops + sonorant (*r, *l, *m, *n, *u̯) is prohibited in the onset of the output. Assign a $*$ for every violation.

d. DEP-V: Every vowel in the output has a correspondent in the input. Assign a $*$ for each instance.

e. MAX-C: Every consonant in the input has a correspondent in the output. Assign a $*$ for each instance.

f. INTACT: Do not delete the root in its entirety. Assign a $*$ for every violation.

The first constraint, PARSE, demands that every segment be syllabified in the output. The next two, *CHCC and $*_\sigma$[PPR, are PIE-specific markedness constraints, which require that the sequences CHCC and $_\sigma$[PPR be avoided in the output. DEP-V and MAX-C are faithfulness constraints, which require that no vowel be epenthesized and no consonant be deleted, respectively. Because we find laryngeal deletion in *$d^h ugh_2$trés* and not epenthesis, the constraint ranking must be DEP-V $\gg$ MAX-C. Lastly, the constraint INTACT, though ad hoc for my analysis, very understandably requires that the root be 'intact' in the winning candidate in some fashion, with there being at least one phonological segment of the root remaining in the output.[52]

[52]Cf. REALIZEMORPH 'A morpheme must have some phonological exponent in the output', proposed by Walker 2000.

(28) PIE *$d^h ugh_2 trés$ → [d^huk]$_\sigma$[trés]$_\sigma$

/$d^h ugh_2$ + trés/			PARSE	*CHCC	DEP-V	MAX-C
a.		[$d^h ugh_2$]$_\sigma$[trés]$_\sigma$		*!		
b.		[$d^h ug$]$_\sigma h_2$[trés]$_\sigma$	*!			
c.		[$d^h u$]$_\sigma$[gəh_2]$_\sigma$[trés]$_\sigma$			*!	
d.	☞	[$d^h uk$]$_\sigma$[trés]$_\sigma$				*

In (28) we see that candidate (a) fails, since it violates the undominated constraint *CHCC. Candidate (b) is not the most optimal, since the *h_2 is not syllabified in the output and all word-medial consonants must be syllabified. Lastly, candidate (d) triumphs over (c), since it was better for PIE speakers to delete a consonant than it was for them to insert a vowel.

(29) PIE *$ph_2 trés$ → [pəh_2]$_\sigma$[trés]$_\sigma$

/ph_2 + trés/			INTACT	*CHCC	*$_\sigma$[PPR	DEP-V	MAX-C
a.		[ph_2trés]$_\sigma$		*!			
b.		[ptrés]$_\sigma$			*!		*
c.		[trés]$_\sigma$	*!				**
d.		[pə]$_\sigma$[trés]$_\sigma$				*	*!
e.	☞	[pəh_2]$_\sigma$[trés]$_\sigma$				*	

In the derivation of *$ph_2 trés$, the most optimal form shows one instance of vowel epenthesis with no consonant deletion. The difference between the derivation of *$ph_2 trés$ in (29) and that of *$d^h ugh_2 trés$ in (28) is that consonant deletion does not improve the output in (29) in any way. If *h_2 were deleted, then the resulting [ptrés]$_\sigma$ would violate the undominated constraint *$_\sigma$[PPR (the sequence which prompts schwa secundum). If both *p and *h_2 were deleted, then the output form violates the undominated constraint INTACT, which requires that the underlying root be continued in the output form through at least one segment.[53] This results in the choice of the most optimal candidate, *[pəh$_2$]$_\sigma$[trés]$_\sigma$, with vowel epenthesis (vocalization).

2.3.3 Why does *#CHC- simplify to *#CC-?

It's worth digressing and investigating why, in response to the ban of an onset *#CHC- in PIE, *#CHC- should simplify to *#CC- (with loss of *H) and never to *#CH- (with loss of final *C) or *#HC (with loss of initial *C). One could simply suppose that laryngeals were marked consonants in PIE and more prone to deletion than other consonants, but it is conceivable that the reduction of *#CHC- to *#CC- may be attributed to a general markedness (but not necessarily complete avoidance) of the PIE clusters *#CH- and *#HC- themselves.

Roots of the shape *#CH- are very poorly attested in PIE. The LIV lists seven: five of the shape *P + *H; two of the shape *s + *H. The evidence for the onset *#sH- (at least, *#sh_2-) in PIE is fairly strong; though the two verbal roots given by the LIV, *$sh_2a\underset{\frown}{i}$- 'bind' (544) and *$sh_2a\underset{\frown}{u}$- 'pour, rain' (555) do not conclusively point back to PIE *sh_2-, the root equation of Hitt. *(i)šḫamai-* 'song' and Skt. *sā́man-* 'song' (< *sh_2om-) assures that the onset *sh_2- was a possible one in PIE.[54]

Roots of the shape *P + *H are much less securely reconstructable.[55] The first given in the LIV, *$k^{(w)}h_2ad$- 'crush', is only attested in IIr., continued by Ved. *khā́dati* 'chews', YAv. *vī-xaδa* 'smash apart!', Khot. *khad-* 'wound' and Bal. *khāδ-* 'eat' (LIV

[53]Note, too, that there is no onset of the shape *#HPR- reconstructable for PIE (aside from *$h_3b^h ruh_x$- 'brow'), which presumably is why **h_2*trés* does not surface as the winning candidate.

[54]Kloekhorst 2008:395.

[55]For a more extensive study, see Elbourne 2000.

359-60). Since this root is restricted to the IIr. branch, aspiration does not necessarily have to be derived from an initial cluster *k + *h_2, and may only be reconstructed with any certainty as PIIr. *k^h*ad*-. The next root, *$k^{(w)}h_2$*aĝ*- ‘consume’, is found solely in Iranian (Khot. *khāś*- ‘drink’, Parth. *x’z*- ‘devour’) and Arm. *xacanem* ‘I bite’ (LIV 360). However, as Klingenschmitt 1982:210 discusses, it is conceivable that the Arm. form is an Iranian loanword (cf. Parth. *x’z*-), and therefore we would only need to reconstruct a PIr. root *k^h*aȷ́*-, though as Martirosyan (2010:324) points out, “one [would need to] assume a very old borrowing with consonant shift **j* > *c*, cf. the well-known case of *partēz* ‘garden’.”

The root *kh_2*ai̯d*- ‘hit’, is only attested in Italic (*caedō* ‘I strike’) and perhaps in Alb. *qeth* ‘cuts; shears (hair)’.[56] If the LIV is correct in separating Skt. *khidáti* ‘rips’ from these roots, neither the Latin nor Albanian form requires the sequence *kh_2- and may just as easily go back to an original **kaid*- (with *a*-vocalism). Next, the LIV cites the root *th_2*au̯s*- ‘be quiet’; cf. Hitt. *tuḫussiyezzi*, CLuv. *taḫušii̯a*- ‘keep silent/quiet’ (< *th_2*us-i̯é*-). Melchert (AHP 108-9) and Kloekhorst (hesitantly in 2008:894-5) assume an inherited *th_2- onset, with later anaptyxis specific to each Anatolian language: PIE *th_2*u*- > Hitt. **tuḫu*-, CLuv. **taḫu*- . Oettinger (1979:326) reconstructs *tuh_2*s*- as the protoform, with expected metathesis (see below) of the laryngeal and high vowel dating back to PIE (PIE *th_2*u*- > PIE *tuh_2- > Hitt. **tuḫ*-), comparing Skt. *tūṣṇī́m* ‘quietly’.[57] Lastly, probably the most widely attested root that points back to an initial *PH- onset is *(z)$g^{wh}h_2$*al*- ‘make a mistake’, which is continued by Ved. (AB) *skhalate* ‘makes a mistake’, Arm. *sxalem* ‘I make a mistake’, Gk. *sp*^{h}*állō* ‘trip up, overthrow’, *sp*^{h}*állomai* ‘be mistaken’ and, without initial *s*-mobile, Lat. *fallō* (LIV 543). Perhaps,

[56] See LIV 360 for references. Orel 1998:359 reconstructs **kaitsa*, which he connects with Skt. *keśa*- ‘hair’, Lith. *káisti* ‘scrape, shave’.

[57] GAv. *tušnā*, OPruss. *tusnan* ‘still’; MIr. *tó*, *tuae* ‘silent’, Wel. *taw* ‘silence’; see Schumacher 2000:179.

however, this root's reconstructable voiceless aspirate may be attributed to Siebs' Law (*$s + D^h \rightarrow$ *sT^h) and not to the cluster *PH.[58]

The strength of each instance of reconstructed initial *#PH- onset varies significantly case-by-case, with roots like *$k^{(w)}h_2ad$- (which occurs solely in Indo-Iranian), which in no way need to be derived from PIE vs. the well-attested root *(s)$g^{wh}h_2al$-, which most certainly does. Of course, the main reason a laryngeal is reconstructed in most of these roots is the presence of aspiration in Indo-Iranian, Greek and Armenian (x < *k^h), though, as Jasanoff 2008:156[4] points out, when reconstructing any of these roots the onset stop + laryngeal is "simply the LIV notational substitute for "classical" (and likewise unsatisfactory) [voiceless aspirate] (AMB)". I myself am not presenting this evidence in order to argue on behalf of the reconstruction of voiceless aspirates.[59] I am merely arguing that the scarcity of *#PH- initial roots indicates that onsets of this shape were disfavored in PIE.[60]

This markedness of the onset *#CH-, as Brent Vine reminds me, may also be reflected in a securely reconstructable phonological process in PIE: the metathesis of the sequence *#CHU- to *#CUH-.[61] This sequence (*#CHU-) is regularly generated in the zero-grade to "long diphthong roots".[62] For example, the zero-grade of the well-attested root 'drink', *peh_3i̯- is realized as *pih_3-*tó*-, with metathesis from original

[58]Lat. *fallō* may go back to *$g^{wh}al$- (cf. *-fendō* < *Cf.* *$g^{wh}en$-; Weiss 2009:79). For another possible etymon with voiceless aspirate derived from Siebs' Law, cf. Ved. *sphuráti* 'jerk, kick', Av. *sparaiti* 'push, tread', OE. *spurnan* 'spurn' vs. Ved. *bhuráti* 'jerk, quiver, move rapidly' (Southern 1999:43). See Siebs 1904, Collinge 1985:155-8, Southern 1999:49ff.

[59]Note that advocates of this view would also have to explain the paucity of initial voiceless aspirates! I find it more plausible that PIE speakers would avoid a complex onset *P + *H over a singleton voiceless aspirated stop.

[60]Word-medially, it is likely that *$PH was present in PIE in the second person singular perfect suffix *-th_2a and mediopassive suffix *-th_2*or/i̯*.

[61]Mayrhofer 1986:174-5, first seen by Winter 1965:192.

[62]Roots of the shape *Ceh_x*U̯*-. Cf. Mayrhofer 1986:174-5.

ph_3i-tó-. This metathesis is assured by the consistent long vowel found throughout the IE languages: Skt. *pītá-* 'having (been) drunk', OCS *pitъ* 'drank' (< *pih_3-tó-*) and Att. *$p\tilde{\imath}t^hi$* 'drink!' (< *pih_3-d^hi*). On the other hand, when *#PPU- sequences are created within the grammar, no such metathesis takes place: *d^hg^{wh}(e)i̯-* 'perish' + *-tó* → *$d^h(z^h)g^{wh}$itó* (Gk. *ép^ht^hito* '(s)he perished') NOT **d^hik^wtó* (Gk. ***ét^hipto*).[63] If, as seems reasonable, metathesis in *#CHU- sequences was driven by an avoidance of *#CH- sequence, then this would demand a constraint ranking *#CH ≫ LINEARITY (= 'Don't metathesize'):

(30) *ph_3itós → *pih_3tós*

/ph_3itós/	*#CH	LINEARITY
a. ph_3itós	*!	
b. ☞ pih_3tós		*

Now, if *#CH and *#CC were equally marked within the PIE grammar, then our constraint ranking would be *#CH, *#CC ≫ LINEARITY, thereby requiring metathesis to occur in the sequence *#CCU-:

(31) *d^hg^{wh}itó → **d^hik^wtó*

/d^hg^{wh}itó/	*#CC	LINEARITY
a. ☹ $d^h(z^h)g^{wh}$itó	*!	
b. ☞ d^hik^wtó		*

Naturally, metathesis did not occur. We may therefore conclude that the constraint ranking in PIE was *#CH- ≫ LINEARITY ≫ *#CC-, and thus the sequence *#CH- was more marked than *#CC-.

[63] *g^{wh} → *k^w by assimilation of laryngeal features with the following *-t-*.

On the other hand, word-initial sequences of the shape *#HC- are much more robustly attested. Well-established roots of the shape *#HT- include _*$h_1$ger-_ ‘wake up’ (Gk. _egeírō_, Ved. _jāgā́ra_; LIV 245-6), _*h_2k̂ou̯s-_ ‘hear’ (Gk. _akoúō_, Goth. _hausjan_),[64] _*$h_2$seu̯s-_ ‘become dry’ (Gk. _(h)aũos_ ‘dry’, Ved. _śuṣyati_ ‘will dry’; LIV 285) and _*$h_2$teu̯g-_ ‘terrify’ (Hitt. _ḫatukzi_ ‘terrifies’, Gk. _atúzomenos_ ‘terrified’).[65] Even more common are roots of the shape *#HR-: cf. _*$h_1$leud^h-_ ‘climb; grow’ (Gk. _ḗlut^{h}on_ ‘I came’, Goth. _liudan_ ‘grow’; LIV 248-9), _*$h_1$nek̂-_ ‘take’ (Gk. _enegkeĩn_ ‘bring’, OCS _nošǫ_ ‘I carry’; LIV 250-1), _*$h_2$leg-_ ‘attend to, worry’ (Gk. _alégō_ ‘worry about’, Lat. _-legō_; LIV 276-7), _*$h_3$reĝ-_ ‘reach out’ (Gk. _orégō_, Lat. _regō_; LIV 304-5).

If the sequence *#HC- was so common in PIE, then how could one claim that *#HC- was more marked than *#CC-? At this point, I find no evidence in favor of such an assumption. The total number of roots of the shape *#PP in PIE is vanishingly small – perhaps almost as small as *#PH- roots. Aside from the root/stem _*pter-_ ‘wing’, _*ptah$_2$k-_ ‘duck, crouch’[66] and _*pk̂u-_, the ø-grade of _*pek̂u-_ ‘livestock’,[67] the most common type of root with an onset *#TT- contains so-called ‘thorn’ clusters, such as _*t(s)k̂ei̯-_ ‘inhabit’ (Ved. _kṣéti_ ‘dwells’; LIV 643-4); _*t(s)ken-_ ‘strike, wound’ (Gk. _kteínō_ ‘kill’; LIV 645-6) and _*d^h(z^h)g^{wh}ei̯-_ ‘be destroyed’ (Gk. _ép^{h}t^{h}itai_ ‘has

[64] Likely from _*h_2k̂h$_2$ou̯s-_ ‘be sharp-eared’, a denominative verb composed of the root _*$h_2$ak̂-_ ‘sharp’ + _$h_2$áus-_ ‘ear’ (Ringe 2008:28). Even if this clever etymology were true, it does not prove that an onset _*h_2k̂h$_2$-_ was permissible in PIE – for all we know _*h_2k̂h$_2$ou̯s-_ > _*h_2k̂ou̯s-_ upon compound formation. See IEW 587-8.

[65] LIV 286. The LIV also cites _*$h_3$pus-_ ‘copulate; marry’, which Watkins 1982 has proposed as the source of Gk. _opuíō_ ‘wed; have sex with’ and Hitt. _ḫapuša-_ ‘reed, shaft, penis’. Kloekhorst 2008:299 argues that due to problems of vocalism and semantics, Hitt. _ḫāpūša-_ ‘shaft (of an arrow, reed), shinbone’, cannot be connected with the Gk form.

[66] Gk. _ptḗssō_ ‘I duck’, Arm. _t‘ak‘eaw_ ‘concealed (oneself)’, Lat. _taceō_ ‘am silent’, OHG _dagēn_ ‘be silent’ (LIV 495)

[67] > YAv. _fšu-_, Ved. _kṣu-_ (Hoffmann & Forssman 2004:105) and perhaps Gk. _ku-klṓps_ ‘Cyclops’ (< ‘cattle thief’).

perished'; LIV 150-1).[68]

In conclusion, with no evidence of a markedness scale *$HP- ≫ *$PP- visible in our reconstruction of PIE, perhaps it would be better to assume simply that the laryngeals *themselves* were marked in PIE (relative to other consonants), a proposal perhaps corroborated by a sizeable number of laryngeal deletion rules reconstructable for PIE.[69] This markedness may be illustrated using two faithfulness constraints, MAX-C and MAX-H.[70]

(32) PIE CONSONANTAL FAITHFULNESS CONSTRAINTS.

a. MAX-C: Every non-laryngeal consonant in the input has a correspondent in the output. Assign a $*$ for each instance.

b. MAX-H: Every laryngeal consonant in the input has a correspondent in the output. Assign a $*$ for each instance.

---

[68]See Schindler 1977b.

[69]See Mayrhofer 1986 for examples. Note that this does not imply that EVERY instance of laryngeal deletion should be motivated by the constraint ranking MAX-C ≫ MAX-H. For instance, we will later see that loss of *H in CHCC > CCC in word-medial position is to be attributed to a violation of PIE syllable markedness constraints.

[70]Kie Zuraw has pointed out to me that it is typologically common for languages to preserve the first consonant of a word-initial cluster, if simplified. However, the reduction of the sequence *#P_1P_2R̥ → *#P_2R̥ as seen in PIE **$d^h\hat{g}^h$m̥-* → **$\hat{g}^h$m̥-* (Gk. *k^hamaí* 'on the ground') perhaps speaks against such an explanation in this case.

(33) PIE *$ph_2t\acute{\bar{e}}r$ → *[ptḗr]$_\sigma$

/ph$_2$trés/		INTACT	*CHC	DEP-V	MAX-C	MAX-H
a.	[ph$_2$tḗr]$_\sigma$		*!			
b.	[h$_2$tḗr]$_\sigma$				*!	
c. ☞	[ptḗr]$_\sigma$					*
d.	[tḗr]$_\sigma$	*!			*	*
e.	[pəh$_2$]$_\sigma$[tḗr]$_\sigma$			*!		

2.4 The DECOMPOSITION THEOREM.

Returning now to the more general problem of reconstructing PIE syllabification, we've seen that the rule given in (26) above explicitly delineates which environments vowel epenthesis and consonant deletion should occur in, if a particular consonant cluster has been determined to be a 'bad' one. Of course, in order to understand when a cluster should be fixed, we must have some sort of metric to know what was a 'good' and 'bad' cluster in Proto-Indo-European.

One such metric that we may utilize in reconstructing PIE syllabification is a method of phonotactic analysis called the DECOMPOSITION THEOREM (DT), given in (34) below.

(34) DECOMPOSITION THEOREM (DT).[71]

"All medial clusters should be decomposable into a sequence composed of an occurring word-final cluster and an occurring word-initial cluster."

In other words, understanding word-edge phonotactics should allow us to predict what

[71] Hammond 1999:68-69.

were possible word-medial codas and onsets in PIE, and from there, word-medial syllabification. Earlier instantiations of the DT may be found in Vennemann (1972; 1985), who splits up the DT into two syllable-preference laws, the LAW OF INITIALS and the LAW OF FINALS. For other similar proposals, see Pulgram 1970:97, Kahn 1980:57-8 and Steriade 1999:223ff., among others.

To illustrate the claims of the DT, let's test its predictive power in an analysis of PIE. A word such as **tréyes* 'three (nom.pl.)' is securely reconstructable through a comparative analysis of forms attested in many IE languages (Skt. *tráyaḥ*, Gk. *treĩs*, Eng. *three*, etc.). This word has a word-final coda **-s* and a word-initial onset **tr-*, which predicts the existence of a word medial cluster **-s.tr-*, with the syllable boundary located between **-s-* and **-t-*. This sequence has been confirmed in the reconstruction of a word such as PIE $*b^hes.trah_2$, a derivative of the PIE root $*b^hes$- 'blow',[72] which is continued by Sanskrit *bhas.trā-* 'tube, bottle'.

An excellent example of the DT being able to predict medial syllabification in an attested language may be found in Modern Icelandic, as discussed by Vennemann (1972:11). Here we see that the process of syllabification can play an important role in determining legal sequences in a language, and if the sequence proves to be illegal, either deletion or epenthesis will occur. In Modern Icelandic there is a rule in which a *t* is deleted in between an *s* and another consonant, except when the consonant is *r*. For example, /t/ is lost in /sɪstkɪn/ → [sískjɪn] 'siblings', but not in /vestra/ → [vɛ́stra] 'in the west'. Vennemann notes that only *-str-* is allowed to remain because it is syllabified as *s.tr*, as *#tr-* is a legal onset in Icelandic. On the other hand, *stk* cannot be syllabified as *st.k* or as *s.tk*, since neither is a legal syllable margin in Modern Icelandic. The consonant /t/ cannot be syllabified; hence /t/ deletion occurs.[73]

[72]IEW 146. Cf. Skt. *bábhasti* 'blows', Gk. *psū̃ō* 'I blow'.

[73]In word-final *-st* clusters (e.g. *breytast* 'gets dressed'), the /t/ is 'extrasyllabic'. See 2.5.5 and Ch. 3 for definition and discussion.

2.5 Exceptions to the DECOMPOSITION THEOREM.

As it is currently formulated in (34), the DT cannot be universally true for medial consonant clusters in all of the world's languages, as there are situations that arise where restrictions on word-medial sequences are stricter or looser than those at word's edge. In order for the DT to be a viable tool in our study of PIE medial consonant sequences, we must address and explain all of these exceptions and then alter our model accordingly. It is crucial that the DT work for all of the world's languages; if it does not, how can we be certain that PIE was not an exceptional language in this regard? The remainder of this chapter will be organized as follows. First, we will examine situations where the DT (as given in (34) above) is violated, discussing when these exceptions are likely relevant for PIE. We will then turn to an earlier study by Steriade (1999), which deals with the interaction of word-edge phonotactics and syllable structure. This will lead to a revision of the DT such that we may state it as a typological universal, thereby allowing us to securely reconstruct PIE syllabification.

2.5.1 Sequential Constraints.

While it is true that every well-formed word consists of well-formed syllables, it is not true that any combination of well-formed syllables may create a well-formed word, since certain combinations of syllables are ruled out by sequential markedness constraints. These are constraints that apply in a linear fashion, such that if a sequence XY is blocked in one position it is blocked everywhere else, regardless of its location in the syllable. We have already seen one such constraint reconstructable for PIE, the OCP. Given the existence of words such as *b^hered* 'he carried (impfct.)' and **didoh$_3$mi* 'I give (pres.)', the DT would predict a sequence such as **edde* to occur, but we know this sequence never surfaced in PIE (see section 1.3.1). Another example of a sequential phonological constraint reconstructable for PIE is *K^{w}u̯/*u̯K^w, the so-called *bou̯kólos* rule, which was first proposed by Saussure 1889:161-2 (MSL, 6, 1889, 161-2) and has been established by Weiss (1995). From an analysis of the PIE words *suh$_x$nḗu̯[74] 'son (loc.sg.)' and *k^{w}ís[75] 'who? (nom.sg.)', the DT would predict the sequence *-u̯k^w- in the word *g^{w}ou̯k^{w}ólos to be a perfectly acceptable consonant cluster, but a linear constraint delabializes *-k^w- after a preceding *-u̯-, resulting in *g^{w}ou̯kólos 'cowherd' (> Gk. *boukólos*, OIr. *búachaill*). Thus we find that the presence of these two sequential constraints, the OCP and *K^{w}u̯/*u̯K^w, results in overgeneration within a DT-based analysis, since medial sequences which are otherwise predicted are in fact blocked.

As Pierrehumbert (1994:168) points out, standard syllable-based phonological theory – just as our formulation of the DT in (34) above – predicts that "in the absence of additional provisos, any concatenation of a well-formed [word-final (AMB)] coda and a well-formed [word-initial (AMB)] onset is predicted to be possible medially in a word." However, this is not the case for modern English, as Pierrehumbert demonstrates in her study of its inventory of medial clusters consisting of three or more consonants. English has a particularly rich array of consonant clusters at word's edge, with 147 different consonantal sequences in word-final position and 129 in word-initial position. Should it be possible for any combination of an existing word-final coda and word-initial onset to create a possible word-medial consonant cluster, we would expect 18,963 possible medial sequences in English! In reality, however, only 675 distinct clusters of three or more consonants are found word-medially in English, and only fifty clusters are found morpheme-internally (ibid.).

Looking at this problem from a stochastic point of view,[76] Pierrehumbert dis-

[74] Ved. *sūnáu*.

[75] Hitt. *kuiš*, Lat. *quis*, etc.

[76] By which I mean "that [which] follows some random probability distribution or pattern, so that its behaviour may be analysed statistically but not predicted precisely" (Oxford English Dictionary on-

covers that the combination of a low-frequency word-final coda and a low-frequency word-initial onset results in a non-occurring or low-frequency word-medial consonant cluster. Conversely, the combination of a high-frequency word-final coda and a high-frequency word-initial onset produces 200 likely medial sequences, from which almost all of the actually attested fifty morpheme-internal sequences may be taken. To rule out the 150 likely, but non-occurring word-medial sequences, Pierrehumbert invokes various sequential markedness constraints in her analysis of English, such as the OCP, which rules out geminates (**vekkro*) and more loosely clusters such as **-lCl-* (**velclo*), a constraint that requires agreement in labiality within nasal-stop sequences (**vemkro*) and the ban of velar obstruents before labials (**vekpro*). In short, one may gauge the likelihood of the occurrence of monomorphemic word-medial clusters in a language by discovering what are the most common edge sequences attested in a language and by determining relevant sequential constraints that may rule out concatenations of those sequences. Of course, since there exists no corpus of PIE texts it is impossible to do a statistical phonotactic analysis of PIE words as Pierrehumbert has done for English. However, her findings confirm our discussion of PIE above: word-medial sequences otherwise predicted by the existence of edge clusters may be ruled out by sequential markedness constraints.[77]

2.5.2 Consonant Licensing.

Typologically it is common for languages to require a certain consonant or category of consonants to be followed by another consonant in order to be syllabified. For example, it is obvious that the DT is not able to account for the medial sequence in

line).

[77]Moreover, since disyllabic morphemes were very rare in PIE (cf. **atta* 'daddy' and complex suffixes such as **-is-th$_2$o-*, **-e-ro-*, etc.), one wonders if we should focus on the 675 attested heteromorphemic consonant clusters in English instead for a better comparison with PIE.

Gk. *léktron* ‘bed’. Like all stops in Greek, the coda *-k#* is not legal word-finally[78] and the onset *#ktr-* is not legal word-initially. Following Steriade (1982:223ff.), the coda consonant /k/ is allowed to stand in *léktron* since it is licensed by the following segment; in other words, /k/ can only exist in a coda if it is followed by and linked with another consonant.[79] In a similar fashion, in Italian and Japanese, obstruent codas are licensed only when they form the first part of a geminate sequence (cf. Ital. *struz.zo* ‘ostrich’, Jap. *gakkō* ‘school’) and are strictly banned word-finally (Itô 1988:17ff.).

In Greek consonant licensing is accompanied by complete laryngeal feature assimilation (voicing and aspiration); in Italian, it is accompanied by laryngeal feature and place assimilation. Examples include Gk. /klep/ ‘steal’ → [kleb] in *klebdēn* ‘stealthily’, Gk. /strep^{h}/ ‘turn’ → [strep] in *strépsomai* ‘I will turn’ and Ital. /leg(g)/ ‘read’ → [let] in the past participle *letto* ‘read’. We have seen in section 1.2.1 that PIE coda obstruents were also required to assimilate completely to the following obstruent with respect to laryngeal features (but not place); cf. */skab^{h}/ ‘shave’ (cf. Eng. *shave*) + *-s- (aor.) > *skaps- > Gk. *éskapsa* ‘dug up’ and *$h_1$leng^{wh}-t- > *$h_1$lenk^{w}-t- > Germ. *le(n)h^{w}t(a) ‘light’ (> Eng. *light*). However, PIE obstruents do not have to be linked to a following consonant in order to be licensed in the coda, which may be inferred from the existence of obstruent-final words such as *úd ‘up; on high’, *ég(z) ‘out(side)’, *úb(z) ‘up’, *u̯ih$_1$roms ‘men (acc.pl.)’,[80] *b^{h}eront ‘they carried (impfct.)’,[81] *u̯ḗkst

[78]This phonotactic generalization of course does not apply to the preposition *ék* ‘out’ or *ouk* ‘not’, since both are proclitics and thus must always be followed by another word.

[79]Likewise the /k/ of *ónuks* ‘fingernail’ is segmentally linked to the following /s/ (Steriade 1982:223ff.).

[80]In early PIE *Vms > *V̄m by Szemerényi’s Law: *d^{h}ég^{h}ōm < *d^{h}ég^{h}oms (cf. ph$_2$tḗr < ph$_2$térs (Szemerényi 1970:109)9; in later PIE, the sequence *-Vms was restored (see section 3.3.3).

[81]As Craig Melchert had (tentatively) suggested to me many years ago, it is possible that word-final *-nt > *-n already in PIE, as no IE language provides direct evidence for this ending (Ved. *-n*, Av. *-n*, Gk. *-n*, Goth. -na). This would only hold true, however, if Faliscan *-od* (/-ond/) does not continue the original 3pl. secondary ending *-nt (Weiss 2009:387).

'carried' and **b*^{h}*eronts* 'carrying'.[82]

2.5.3 Lexical & Morphological Gaps.

Edge consonant sequences can sometimes be more restricted than medial ones. For instance, even though the sequence *-mpt-* occurs word-medially in Lat. *ēmptus* 'the act of purchasing', a DT analysis of Latin predicts that this word should not exist, since *mp#* and *#pt* are illegal edge codas and onsets, respectively.[83] This is the opposite of the situation discussed above in section 2.5.1, where a linear constraint may override the predictions of the DT, thus resulting in overgeneration, since **g*^{w}*ou̯k*^{w}*ólos* was predicted to occur but does not. Here in the case of *ēmptus*, the DT results in undergeneration, since no word-edge cluster is attested that would predict the syllabification of a sequence *-mpt-* in Latin.

We may view this undergeneration by the DT in one of two ways. First, in the case of *ēmptus*, the epenthesis of a stop [p] is required for the transition from /m/ to /t/ in the output: *em-* 'buy' + *-tus* > *ēm[p]tus*.[84] For this reason, perhaps we should view the underlying syllable structure of the medial sequence as *m$t*, which speakers would have constructed based on the already existing edge sequences in the language (cf. *hominem* 'man (acc.sg.)', *tū* 'you (nom.sg.)'). Since there is no word-final ***-mt$* and no word-initial ***$mt-* in Latin, the syllabification could only have been *-m$t-*. Should this hypothesis be true, the DT would predict medial syllabification based on underlying structures, not surface ones. However, the [p] in question must have been syllabified somewhere within *ēmptus*; otherwise, speakers of Latin would not have

[82] An additional example may be found in Keydana 2004:186-9, who argues **-nts* not to have been licit coda in IE, which was sometimes realized as **-ns*, other times **-nt*.

[83] Hermann 1923:360.

[84] See Weiss 2009:173.

been able to pronounce it.[85] Thus, should we follow a formulation of the DT that applies at the phonemic level, we would lack the means of predicting not only which syllable [p] should be parsed with, but also its pronunciation as well! Consequently, it seems prudent not to formulate the DT to apply at the phonemic level, since our goal is to predict the syllabification of all consonants in the analysis of a language, which of course also includes consonants that are epenthetic.

A more serious exception lies in Lat. *sculptor*. As with *ēmptus* above, there exists no word-final *-lp* coda, though this particular [p] may in no way be claimed as epenthetic, since the root in question is *sculp-*. For this reason, undergeneration produced by the DT may be better attributed to a gap within the morphology or lexicon. In Latin, as is the case in almost all ancient IE languages and in PIE itself, the vast majority of words must end in some kind of overt morphology, and there exist no morphological endings of the shapes *-mp* or *-lp*. In fact, the inventory of word-final suffixes in IE by and large consists of a single consonant (**-s*, **-m*, **-H*, etc.), and if there is a consonant cluster in either a single suffix or through a concatenation of morphemes, the last consonant is invariably a coronal obstruent (**-nt* '3rd pl.', **-ms* 'acc.pl.', **-s-t* 'aor. 3rd sg.', etc.) or laryngeal (**meĝ*h_2 'great', **-me(s)d*$^h h_2$ '1st pl. M/P').[86] That said, though word-final codas of the shapes *-mp* and *-lp* do not exist in Latin, a coda consisting of a resonant + coronal stop does exist in word-final position in the sequences *-nt* (*sunt* 'they are'), *-rt* (*fert* 'he carries') and *-lt* (*volt* 'he wants'). The undergeneration produced by the DT may therefore be attributed to a gap in the Latin lexicon and grammar. Similarly, it would be hard to deny the existence of words such as **i̯eu̯ktrom* 'instrument of binding, yoking' (Skt. *yoktra-* 'cord') and **denk̂tro-* 'instrument of biting' (Skt. *dáṁṣṭra-* 'tusk') in PIE, even though there are no recon-

[85] See above for discussion of the undominated constraint PARSE.

[86] Hitt. *-wašta*, Ved. *-mahi*, Gk. *-me(s)t*^{h}*a*.

structable word-final clusters of the shapes *-$\underset{\frown}{u}k$/-$\underset{\frown}{u}g$ or *-$n\hat{k}$/-$n\hat{g}$.[87] These facts suggest that the DT should not be so strictly formulated as it is in (34), and that perhaps a more successful formulation of the DT should not require specific sequences to occur word-finally in order for them to occur word-medially. This possibility will be discussed in more detail in section 2.6 below.

2.5.4 Morpheme Structure Constraints.

Furthermore, there are often language-specific restrictions upon the possible shapes of words within a particular morphological category, which are called morpheme structure constraints (MSCs).[88] Adherents of (monostratal) OT explicitly reject the existence of MSCs because they run counter to the assumptions of Richness of the Base (Kager 1999:19-20), and have also been rejected in the past due to their frequent redundancy with respect to existing phonological rules. Nevertheless, a brief mention of MSCs has been included here, due to their relevance for the reconstruction of PIE. As first seen by Meillet (1934:173ff.), there were three root shapes that are strikingly absent: roots of the shape *TeD^h-, **DeD*-, and *D^heT-, where **T* = any voiceless stop, **D* = any voiced stop and *D^h = any voiced aspirated stop.[89] The DT would predict that the syllable **ded* would be legal anywhere in PIE (cf. *b^hered 'he carried', *$didoh_3mi$ 'I give'), though no root of this shape is reconstructable. Just as we saw above in the case of sequential constraints, the existence of MSCs leads to instances of overgeneration by the DT, predicting phonological sequences that do not occur in the

[87] One conceivable situation in which such word-final sequences could have arisen is in the suffixless 2nd sg. root imperative, as seen in Lat. *ī* 'go! (2nd sg.)' (< *$h_1e\underset{\frown}{i}$). However, the word equation Gk. *ít*^{h}*i* = Skt. *ihí* makes it likely that the original PIE form was *$h_1id^h\acute{\imath}$, which for our purposes means that the 2nd sg. root imperative did not contribute to the inventory of consonant sequences at word's edge. Cf. Fortson 2010:105.

[88] Cf. Halle 1959:56ff.

[89] See Mayrhofer 1986:95[19], Szemerényi 1990:90ff. and Fortson 2010:78.

language.

2.5.5 Extrasyllabicity.

Lastly, it is typologically very common for a language's word-edge clusters to allow more complex sequences than the onsets and codas in medial position, a phenomenon called extrasyllabicity or edge effects. A cursory glance at an English word such as *texts* [teksts] proves this to be the case – there is no word in English with a word-medial coda of the shape -ksts$]_\sigma$. Similar facts will become apparent in our discussion of PIE extrasyllabicity in Chapter 3, when we revisit the laryngeal loss rule CH.CC > C.CC.

2.5.6 Review of Exceptions to the DECOMPOSITION THEOREM.

So, to review, it appears that there are two main problems with the strict formulation of the DECOMPOSITION THEOREM as given in (34). First, it overgenerates, or predicts non-existent sequences, in at least three cases: 1) instances of sequential constraints such as the OCP, 2) the occurrence of Morpheme Structure Constraints and 3) the existence of extrasyllabic consonants at word's edge. All three are reconstructable for PIE, the third of which will be discussed in great detail in the following chapter. Second, this strict formulation of the DT undergenerates, or fails to predict existing sequences, in at least two instances: non-existent clusters at word's edge due to 1) consonant licensing and 2) gaps within the lexicon and morphology. The first appears not to be relevant for PIE, since the same basic consonant clusters that occur word-medially may occur word-finally in PIE. However, there are accidental consonant cluster gaps in the lexicon that give us a deficient view in predicting medial consonant sequences. In short, in order to be able to predict a well-formed word in PIE, we must be fully aware that it is simply not the case that there is a one-to-one correspondence between word-edge clusters and word-medial clusters. In order to adopt the DT as a successful metric of syllabification, we must be able to explain all five of these exceptions and incorporate them into our model.

2.6 Establishing the DECOMPOSITION THEOREM as a Typological Universal.

This section will elucidate the typological universality of a broader, less restricted version of the DECOMPOSITION THEOREM, which will be able to accommodate four of the five exceptions given in section 2.5. Universality, of course, is crucial in our study, as it allows us to reconstruct PIE syllabification in a reliable and credible fashion. To do so, we must establish a psychological connection between word-edge phonotactics and the process of medial syllabification. Consequently, it will become clear that the DT is not some ungrounded analytical tool devised for the study of medial syllabification; rather, it is a heuristic guideline innate to all humans for the purpose of dividing medial sequences into two distinct syllables, as Steriade (1999) proposes.

In the introduction of this seminal work, Steriade discusses the correlation between phonotactics, stress and intuitions of syllable division, which has been viewed by many to be strong evidence in favor of the assumption of the syllable as a unit of hierarchical structure. To give an example, Steriade (ibid.) compares Spanish and (Cairene) Arabic, where three distinct facts appear to be motivated by a unitary phonological phenomenon.[90] In Spanish, words can begin with clusters consisting of stop + liquid (PR), such as *tres* 'three', but in Arabic they cannot. Similarly, post-consonantal PR is allowed in Spanish (*semblanza* 'sketch'), but in Arabic, it is not. In Spanish, clusters of the shape PR do not count as heavy for stress assignment (*fúnebre* 'sad', not ***funébre*; contrast *solémne* 'solemn', not ***sólemne*), while those in Arabic do

[90]Data taken from Harris (1983) and Broselow (1976), respectively.

(*tanábla* 'extremely lazy (pl.)'). Lastly, the native intuitions of syllabic division for word-medial sequences of the shape VPRV are V.PRV by speakers of Spanish (*o.tros* 'others') but are VP.RV by speakers of Arabic (*zak.ru* 'they studied').[91]

These differences in Spanish and Arabic with respect to phonotactics, stress assignment and speaker intuition have in the past been interpreted to be the result of a single phonological fact: Spanish allows complex onsets of the shape PR, but Arabic does not.[92] This has been seen by many to justify the assumption of the syllable as a phonological unit. However, for Steriade (ibid.) this assumption is unnecessary: "Arabic lacks all word-initial clusters whereas Spanish allows TR [= PR (AMB)] clusters word-initially. Word finally Arabic permits a broad range of C's, whereas native Spanish words end in sonorants or [s] only, not in stops. The syllable intuitions can be deduced entirely from the word edge differences: Spanish favors V.TRV over VT.RV parses because (a) word final stops are missing in the native lexicon and (b) TR initials are possible. Arabic on the other hand rejects V.TRV in favor of VT.RV because (a) TR initials (and all CC initials) are impossible and (b) VT finals are not ruled out."

Steriade's explanation, of course, looks very much like the DECOMPOSITION THEOREM. There is a key difference, however. Steriade does not in fact believe that syllable boundaries really exist word-medially; rather, she believes that speakers, when asked to perform the task of syllable division, infer syllable boundaries word-medially based on possible sequences at word's edge. Steriade calls this inference the WORD-BASED SYLLABLES (WBS) hypothesis, which she explains as follows: "Speakers rely on inference when they attempt to locate syllable boundaries in a multi-vowelled string, and one guideline in this process is that the segmental composition of word and syllable edges must be similar." Though for our purposes I will not reject the

[91] Ibid.

[92] Ibid.

phonological assumption of syllables altogether as Steriade has done, I will follow her insight into how speakers may choose to parse syllables word-medially. I contend that medial syllabification (and all phonological processes that derive therefrom) relies heavily upon a speaker's knowledge of possible sequences at word's edge in a language, but is not completely dependent upon it. Unlike the earlier formulation of the DT (34), which requires that medial consonant clusters be decomposable into an occurring word-final coda and word-initial onset, I propose that speakers infer the medial syllabification based on what is possible at word's edge.

Much of the evidence that Steriade (1999) presents in favor of her WBS hypothesis comes from instances of variability of syllable divisions,[93] which she claims to arise when both phonotactic and syllabic preferences are in conflict with one another in a single parse. The word *lemon*, for example, exhibits variation in syllabification among speakers of English when surveyed. On the one hand, the parsing of a VCV sequence as V.CV is preferred, which would produce [lɛ.mn̥], while on the other hand open syllables with a lax vowel are strictly prohibited, which would produce [lɛm.n̥]. This conflict is made clear through the variation of syllabification among speakers surveyed (Derwing 1992): 51% of those polled syllabified lemon as [lɛm.n̥], 37% as [lɛ.mn̥] and 12% preferred the parse [lɛm.mn̥] with ambisyllabic *m*. Contrast this variation with the syllabification of the word *demon* among those same speakers surveyed; there was an 82% consensus for the parse [di.mn̥]. The difference here, of course, is that the initial syllable has a tense vowel, which is permissible in word-final position (e.g. Eng. *see*), whereas there is no word like [lɛ][94] in English.[95]

---

[93]Cf. Treiman & Danis 1988.

[94]An exception lies in the very recent and very popular internet interjection 'meh' [mɛ] "[an] expression of apathy or indifference" (American Heritage Dictionary).

[95]Similarly, in an analysis of syllabification and syllabically-driven phonological processes in Italian, McCrary (2004) assumes the perceptually driven phonotactic constraint LEX-C / IN V_ OR _V/L ("In the native lexicon, a consonant may only occur if it is after a vowel or followed by a vowel or liquid.")

Further and more striking evidence comes from Arrernte,[96] where all words begin with a vowel and end in at least one consonant. Steriade's WBS hypothesis correctly predicts the Arrernte syllable to be invariably of the shape VC(C_0), whereas standard syllable-based phonological theory cannot (cf. Prince & Smolensky 1993). The unusual syllable structure of Arrernte (VC(C_0)) is confirmed through phonological processes such as reduplication tests, in which we see the reduplicant is clearly of the shape VC_0: *unt-em* 'is running' + *ep*-RED → *untepunt-em* 'keeps running', *atw-em* 'is hitting' + *ep*-RED → *atwepatw-em* 'keeps hitting'. Similar facts become apparent through language games such as 'Rabbit Talk', which moves the leftmost syllable to the end of the word (*itirem* 'thinking' → *iremit*, not ***tiremi*).[97]

Translating her findings into an OT framework, Steriade (1999:226) tentatively proposes two word-to-syllable identity constraints.

(35) Word-to-syllable identity conditions.

"W-S(I): For any I, a syllable-initial segment, there is a word such that its initial segment is identical to I.

W-S(F): For any F, a syllable-final segment, there is a word such that its final segment is identical to F."

---

to explain instances of speaker variation of syllabification of Italian words such as *pasta* ($[\text{pas}]_\sigma[\text{ta}]_\sigma$). Without the assumption of the LEX-C constraint, the WBS hypothesis would predict that the abundance of native words with initial onsets of the shape *st-*, as in *stare* 'stand' and the absence of native words ending in *-s* should always lead to the parse *pa.sta*. Those speakers with the ranking LEX-C ≫ W-S(I) chose the parse *pas.ta*, whereas *pa.sta* was chosen if a word-to-syllable identity was deemed to be more important (W-S(I) ≫ LEX-C). Another example may be seen in Polish, where a minority of Polish speakers surveyed by Dubiel (1994) parsed the words *karta* and *pokorny* as [ka.rta] and [po.ko.rny], respectively, despite the resulting SSP violation (see section 11). English speakers, on the other hand, unanimously reject such parses (Steriade 1999).

[96] A language spoken in central Australia of the Pama-Nyungan family. See Breen & Pensalfini (1999).

[97] As briefly mentioned in chapter 1, English has a similar language game called Pig Latin, which moves a word-initial onset to the end of the word (in part or in whole), followed by the diphthong *-ay* (*Pig Latin* → *Igpay Atinlay*, *dissertation* → *issertationday*, etc.).

Should we follow Steriade's hypothesis, medial syllabification will crucially depend on the ranking of the word-to-syllable identity conditions W-S(I) and W-S(F) vis-à-vis the other constraints within a language's phonology.

2.7 Fine-tuning the DECOMPOSITION THEOREM.

Keeping the previous section in mind, let us now revisit the DECOMPOSITION THEOREM, which was first formulated in (34) above. In section 2.5 we saw a number of exceptions to this formulation, leading us to the conclusion that the presence of particular consonants and consonant sequences at word's edge cannot be the sole factor in determining what may or may not be syllabified in word medial position. Rather, as Pierrehumbert (1997) convincingly shows, it is only one of many.

Our revised, and more successful formulation of the DT must therefore be much less stringent than (34) and will follow Steriade's crucial insight that speakers base medial syllabification on, and are not necessarily constrained by, the existence of similar, or identical, codas and onsets at word's edge. The DT should not require medial clusters to directly mirror edge sequences, such that a heterosyllabic cluster AB$CD requires AB to be an occurring word-final coda and CD to be an occurring word-initial onset. Rather, a heterosyllabic cluster ABCD is likeliest to be parsed as AB$CD if AB and CD (or sequences similar to AB and CD) occur at word's edge as a legal coda and onset, respectively, and if BCD and ABCD do not occur as legal onsets.

(36) DECOMPOSITION THEOREM (REVISED).

Medial consonant clusters are decomposable into a sequence consisting of a coda plus onset, whose syllable division is produced by the interaction of a speaker's knowledge of consonant sequences at word's edge and syllable markedness constraints.

If speakers do construct syllable boundaries through inference of edge sequences as formulated in (36), it follows that we as non-speakers can also construct medial syllable boundaries in a similar fashion. It is in this way that I justify the use of the DT as a tool in the reconstruction of Proto-Indo-European syllabification.

Let us now put the newly formulated DT to the test, seeing how it measures up with the exceptions given in 2.5 above. First, sequential constraints and morpheme structure constraints may apply without fail, since the DT does not directly control what may or may not occur in word-medial position. In the case of 'cowherd' in 2.5.1, the candidate *g^{w}ou̯k^{w}ólos* never even surfaces due to a violation of the undominated sequential markedness constraint **K^{w}u̯/*u̯K^{w}*. However, the DT does determine how a candidate that does not violate **K^{w}u̯/*u̯K^{w}* will be syllabified. We know that *g^{w}ou̯kólos* could not have been syllabified as **g^{w}o.u̯kó.los* because of a W-S(I) violation,[98] since **u̯k-* does not occur as a word-initial onset in PIE, and it could not have been syllabified as **g^{w}ou̯k.ólos* because of a violation of the markedness constraint ONSET. This leaves us with the most optimal candidate **g^{w}ou̯.kólos*, which satisfies both word-to-syllable identity constraints and the constraint ONSET.[99] Note that I do not assume a violation of W-S(F) in candidate (c) *[g^{w}ou̯k]$_{\sigma}$[ó]$_{\sigma}$[los]$_{\sigma}$, given the similarity[100] of the coda *-*u̯k* with the reconstructable word-final coda *-*u̯d* (PIE **ĝ^{h}eu̯d* 'poured (3rd sg. impfct.)'; cf. Skt. *ajuhot* 'sacrificed (3rd sg. impfct.)').[101]

[98]As well as an SSP violation; see sections 1.4 and 3.3.1.

[99]The constraint rankings assumed below are preliminary and are given on a case-by-case basis. In (37), I have ranked ONSET ≫ W-S(I), based on (39) below.

[100]How does one gauge 'similarity'? At this time, I have no good answer, which I acknowledge poses a major problem for my hypothesis. Though it is beyond the scope of this dissertation to give an accurate metric on how one defines 'similarity', I refer the reader to Bailey & Hahn 2001 for a survey of numerous approaches.

[101]LIV 179.

(37) PIE *g^{w}*ou̯kólos* → *[g^{w}ou̯]$_{\sigma}$[kó]$_{\sigma}$[los]$_{\sigma}$.

*/g^{w}ou̯k^{w}ólos/	*K^{w}u̯/*u̯K^{w}	ONSET	W-S(I)	W-S(F)
a. *[g^{w}ou̯]$_{\sigma}$[k^{w}ó]$_{\sigma}$[los]$_{\sigma}$	∗!			
b. *[g^{w}o]$_{\sigma}$[u̯kó]$_{\sigma}$[los]$_{\sigma}$			∗!	
c. *[g^{w}ou̯k]$_{\sigma}$[ó]$_{\sigma}$[los]$_{\sigma}$		∗!		
d. ☞ *[g^{w}ou̯]$_{\sigma}$[kó]$_{\sigma}$[los]$_{\sigma}$				

Turning to the problem of undergeneration, we now understand that medial syllabification may be based on similar, though not identical consonant sequences at word's edge. For this reason, instances of lexical and morphological gaps are quite straightforward. Words such as Lat. *sculptor*, *ēmptus* are likeliest to be parsed as [sculp]$_{\sigma}$[tor]$_{\sigma}$, [ēmp]$_{\sigma}$[tus]$_{\sigma}$ because RP sequences occur word-finally (see 2.5.3) but PP sequences of the type *$pt-* are non-existent in the native Latin lexicon. In the case of Greek *léktron*, the winning candidate is [lék]$_{\sigma}$[tron]$_{\sigma}$, and not [lékt]$_{\sigma}$[ron]$_{\sigma}$ or [lé]$_{\sigma}$[ktron]$_{\sigma}$. All three candidates violate one or more word-to-syllable identity constraints. However, the most optimal form, [lék]$_{\sigma}$[tron], is chosen because it violates the fewest number of markedness, faithfulness and word-to-syllable identity constraints.[102]

(38) Greek *léktron* → [lék]$_{\sigma}$[tron]$_{\sigma}$.

/léktron/	LICENSE	W-S(I)	W-S(F)
a. [lé]$_{\sigma}$[ktron]$_{\sigma}$	∗!	∗	
b. [lékt]$_{\sigma}$[ron]$_{\sigma}$	∗!		∗
c. ☞ [lék]$_{\sigma}$[tron]$_{\sigma}$			∗

[102]For this tableau I have posited an ad hoc, Greek-specific constraint LICENSE: 'Consonants must be properly licensed.'

It has also been noted that word-initial **r-* was blocked in Greek, Hittite and likely PIE, despite the fact that syllable-initial **r-* is almost certainly assured. Through our knowledge of Greek meter, we know that the first syllable of Greek *ára* was light and therefore was syllabified as $[\text{á}]_\sigma[\text{ra}]_\sigma$. Similarly, we may assume that PIE **h_2n̥rés* 'man (gen.sg.)' was syllabified as *$[\text{h}_2\text{n̥}]_\sigma[\text{rés}]_\sigma$, despite the *W-S(I) violation.[103] Both syllabify as such in order to satisfy the more highly ranking constraint, ONSET.

(39) Greek *ára*, PIE **h_2n̥rés*

/ára, *h_2n̥rés/		ONSET	W-S(I)
a.	$[\text{ár}]_\sigma[\text{a}]_\sigma$	*!	
b. ☞	$[\text{á}]_\sigma[\text{ra}]_\sigma$		*
c.	$[\text{h}_2\text{n̥r}]_\sigma[\text{és}]_\sigma$	*!	
d. ☞	$[\text{h}_2\text{n̥}]_\sigma[\text{rés}]_\sigma$		*

There is one remaining exception discussed in 2.5 that cannot be handled by the present formulation of the DT: extrasyllabicity. It is for this reason that the majority of scholars assume extrasyllabicity (a.k.a. edge effects, syllable appendices, etc.) as an additional theoretical mechanism in the study of syllabification. For example, Pierrehumbert (1997) was required to assume final extrasyllabic consonants in her analysis of medial consonant clusters in English, positing that any coronal obstruent in word-final position that does not follow a vowel, offglide or nasal is extrasyllabic. For instance, the /t/ in *weft* is extrasyllabic, whereas the /d/ in *mad* is not (Pierrehumbert 1997:172). This is confirmed by a lack of medial codas of the shape $\text{-ft}]_\sigma$ (***weft.kry*) and by the existence of medial codas of the shape $\text{-d}]_\sigma$ (*vod.ka*) in English. In the next chapter, we will see that it is also necessary to reconstruct extrasyllabicity in both word-initial

[103] One wonders, though, if the syllabification of such forms fluctuated among speakers of Ancient Greek and PIE, if questioned, as was seen in the example of Eng. *lemon* above.

and word-final position in Proto-Indo-European. This assumption will allow us to explain certain phonological rules of vowel epenthesis and consonant deletion in the PIE grammar.

CHAPTER 3

Livin' on the Edge: PIE Extrasyllabicity and the MAXST.

3.1 Overview.

In the previous chapter, we established a universal means of predicting medial syllabification called the DECOMPOSITION THEOREM, but we were unable to account for the existence of extrasyllabic consonants in our model. This chapter will first begin with a discussion of the process of extrasyllabicity and will establish it as a feature within the PIE phonology, providing typological parallels when relevant. We will then proceed to examine particular phonological rules that occur in word-medial position, which I will claim are motivated by violations of syllable structure. These rules will allow us to formulate a PIE MAXIMUM SYLLABLE TEMPLATE, which I postulate is the driving force behind at least five phonological processes: the *métron* rule, CH.CC > C.CC, the loss of **t* in PIIr. **$h_x$ok̂th$_x$tí-* 'eighty', schwa secundum and vowel epenthesis in the cluster CHC(C).

3.2 Past Uses of the DECOMPOSITION THEOREM in Indo-European Studies.

The DECOMPOSITION THEOREM has been used before in various studies of Indo-European phonology, though never by that particular name. Both Juret (1913) and Wolff (1921) implied DT-like explanations for various phonological changes within the prehistories of Latin and Germanic, respectively. In his 1923 treatment of IE syllabification, Hermann argued against both scholars, demonstrating that the use of the DT in their analyses leads to false conclusions. Against Juret, Hermann (1923:214-216) pointed out that if *Ses.tius* reflected *Seks.tius* one would expect the same simplification of **ks* in *rēx* and *coniux*. Against Wolff, Hermann (1923:271-272) pointed out that if the cluster *VχsCV* was simplified to *VsCV* in German because *s* closed the syllable, the same treatment would also be expected in *sechs*. In both instances, Hermann comes to the conclusion that the reason for simplification is that the clusters themselves were bad, regardless of syllable structure. This leads him to conclude that in an analysis of PIE syllabification (1923:360): "Scrutinizing the beginnings and ends of words doesn't work [in determining the syllabification of] multipartite consonant groups in the middle of the word."[1]

Where Hermann is incorrect, however, is his assumption that word-edge onsets and codas have a status equal to those in word-medial position. As we saw earlier in section 2.5.5, it is possible for a language to permit consonant clusters in a word-initial onset or a word-final coda, but not allow those same clusters in the middle of a word. Word edges are known to license special syllable structures, and certain consonants or consonant sequences may not be incorporated into the onset or the coda of the syllable in question. When these segments occur in medial position they cannot be syllabified,

[1]"Das Schielen nach dem Wortanfang und Wortende nützt also bei den mehrteiligen Konsonantengruppen des Wortinnern nichts."

and so they are either deleted through STRAY ERASURE or supported through STRAY EPENTHESIS (Itô 1988).

These segments are called extrasyllabic.[2] Extrasyllabic segments must be: **consonants** – vowels cannot be extrasyllabic – and **at word's edge** – segments must be peripheral (Hdouch 2008:69).[3] As discussed by Bagemihl (1991:625): "…segments must belong to syllables, syllables to metrical feet, and metrical feet to phonological words or phrases. The only way that a segment can avoid being syllabified is through extraprosodicity [for our purposes, extrasyllabicity (AMB)], which is only available at the edges of well-defined domains (usually words). If a segment is not licensed through either syllabification or extraprosodicity, it is subject to one of two operations: the language may undergo a process of STRAY EPENTHESIS, in which case a degenerate syllable is assigned to the stray consonant (this syllable will ultimately receive a vocalic nucleus through default rules). Otherwise, unsyllabified segments are subject to the process of STRAY ERASURE..., which deletes all unlicensed material from the representation."[4]

Thus we may be able to explain certain instances of medial consonant deletion in PIE as the result of STRAY ERASURE, when the segment is not licensed through syllabification or extrasyllabicity. For example, it is simplest to assume that *-$\hat{g}h_2$# in

[2]See Vaux and Wolfe 2009 for the most recent defense of extrasyllabicity.

[3]Lunden (2006) suggests that word-final extrasyllabicity is ultimately due to word-, phrase- and utterance-final lengthening, whose purpose is to make the boundaries of prosodic categories easier to perceive. This phenomenon is found in all of the world's languages, including sign language, as well as in music. Lunden demonstrates that if a shorter and a longer duration are increased by the same amount, listeners will perceive the increase to the shorter duration as the greater one. Thus, in non-final position a coda consonant will create length that distinguishes itself from a CV syllable, but in final position the increase in duration is not sufficient to metrically distinguish it from the syllable shape CV. In short, Lunden demonstrates that the reason certain languages allow super-heavy syllables only in absolute word-final position is that they are perceived as heavy syllables, providing a perceptual reason for the existence of extrasyllabic consonants in word-final position.

[4]For a perceptual account of edge effects and processes of STRAY ERASURE and STRAY EPENTHESIS, see Côté 2000:267ff.

$*méĝh_2$ ‘great’ was pronounced as a bipartite cluster in PIE, since Hittite *mēg* shows laryngeal deletion[5] whereas inner IE shows vowel epenthesis (Inner IE *$*mégh_2ə$* > Skt. *máhi*, Gk. *méga*).[6] In the sequence *$*\text{-}ĝh_2\#$*, *$*h_2$* was extrasyllabic at a word’s edge and allowed to remain. However, in medial position extrasyllabicity is not allowed and therefore *$*h_2$* is not syllabifiable in the sequence *$*\text{-}gh_2$*. This results in deletion via Stray Erasure: $*[d^h ug]_\sigma h_2[trés]_\sigma$ > $*[d^h uk]_\sigma[trés]_\sigma$.

3.3 Extrasyllabicity in Proto-Indo-European.

Examples like these abound in the phonological literature. In Negev Bedouin Arabic (McCarthy 1994:213-4), there is a prohibition against syllable-final gutturals preceded by *a*: CaGCVC → CaGaCVC (where C = any consonant and G = any guttural). One finds *ašrab* ‘I drink’ alongside *aħalam* (/aħlam/) ‘I dream’ and *tašrab* ‘you drink’ beside *taħalam* (/taħlam/) ‘you dream’. Epenthesis does not occur after word-final or stem-final gutturals: *rawwaħ/rawwaħna* ‘he/we went home’. Word-finally, the guttural is tolerated via extrasyllabicity, but word-medially the guttural must be licensed to a syllable, and therefore vowel epenthesis occurs. Within French there are words such as *arbre* ‘tree’ /aʁbʁ/ (Hdouch 2008:72), which may be realized as [aʁbʁ] (extrasyllabic C pronounced), [aʁb] (deletion) or [aʁbʁə] (epenthesis). This precisely parallels the Hittite outcome of PIE *$*meĝh_2$* (extrasyllabic C pronounced) as *mēg* (deletion) and the inner-IE outcome as *$*megh_2ə$* (epenthesis).

[5]Melchert 1994:87; Kloekhorst 2008:572-3. Cf. PIE *$*éĝh_2$* ↣ Hitt. *ūg* ‘I’.

[6]See section 1.2.2 for explicit discussion of this methodology.

3.3.1 Extrasyllabic Consonants in Coda Position.

Not all laryngeals were extrasyllabic in word-final position. Since we find compensatory lengthening (CL) in the sequences *-VH# (*-ah_2 > Skt. -*ā*, Gk. -*ā*, etc.)[7] and *-RH# (pre-PIE **wédorh*$_2$ 'waters' > PIE **wédōr*),[8] *H in these two sequences could not have been extrasyllabic. Consonants that are licensed by a syllable carry moras, while extrasyllabic segments do not.[9] Since the deletion of an extrasyllabic segment would not have produced CL, *H in the sequences *VH# and *RH# was not extrasyllabic. Thus, *H was only extrasyllabic in the position *VPH#, where there was a violation of the Sonority Sequencing Principle.

(40) Sonority Sequencing Principle (SSP).
Between any member of a syllable and the syllable peak, only sounds of higher sonority rank are permitted.[10]

We may infer that the SSP violation in the coda *PH$ is the reason for the extrasyllabicity of *H, and therefore any segment that violates the SSP in this position should be extrasyllabic. Typologically this makes good sense, as many languages that prohibit SSP violations word-internally allow them at word-edge (Steriade 1982:92, Hdouch 2008:72). French *arbre* /aʁbʁ/ is allowed but there exists no word **/aʁbʁ.po/. In Rus-

[7]CL here, of course, cannot be reconstructed for PIE, since there is no indication the sequence *-VH$ became *V̄ in the proto-language. Nevertheless, the fact that laryngeal deletion in the sequence *-VH$ invariably results in CL in the daughter languages makes it highly likely that laryngeals were syllabified (i.e., not extrasyllabic) in this position in PIE.

[8]Hitt. *úidār*, Gk. *húdōr*. See Szemerényi 1970:155 & 159 and Nussbaum 1986:129ff.

[9]In Attic Greek, when an unsyllabifiable (i.e. extrasyllabic) coda consonant is lost via STRAY ERASURE, it never results in compensatory lengthening: /galakt/ 'milk' → [gala] (not **[galā]); /RED-komid-k-a/ 'I have eaten' → [kekomika] (not **[kekomīka]). See Steriade 1982:227. For a view of compensatory lengthening that is not mora-based see Kavitskaya 2002.

[10]See section 1.4. Note once again that I consider both sonority rises (**ste*-, *-*ets*) and sonority plateaus (**pte*-, *-*ept*) to be violations of the SSP.

sian, we find words such as *mgla* 'mist' and *rubl'* 'ruble', but words such as **glub.mgla* and **rubl'.to* are strictly prohibited.

(41) Rule of Coda Extrasyllabicity in PIE.

Any SSP violation from a consonant in the coda renders that consonant extrasyllabic.

3.3.2 Revisions to CHCC > CCC.

Our rule of coda extrasyllabicity allows us to account for the exceptions presented in (24), which are given again in (43) below. If *H were extrasyllabic only in the word-final sequence *PH# and the loss of a laryngeal occurs in the sequence CH.CC because *H could not be syllabified, then we should expect a more specific environment for laryngeal deletion; namely, *PH.CC > *P.CC.[11]

(42) PH.CC > P.CC

*H → ø / P __ $CC

A post-consonantal laryngeal is lost at a syllable boundary before two consonants if its sonority value is of greater or equal value than that of the preceding consonant.

This revision perfectly handles the previously unexplained exceptions found in section 2.2.2 above:

(43) Counterexamples to CHCC > CCC: Word-Internal Position

a. **k̂erh$_2$srom* > Lat. *cerebrum* 'brain'

b. **temh$_x$srah$_2$* > Skt. *tamisrā*, Lat. *tenebrae* 'darkness'

[11] This analysis holds should we assume that the laryngeals were of equal or higher sonority as stops (as discussed in section 1.4), since both scenarios result in an SSP violation in the coda.

c. *$\hat{g}enh_1trih_2$ > Lat. *genitrix*, Ved. *jánitrī-* 'bearer, mother'

d. *$\hat{g}enh_1d^hlo$- > Gk. *génet^{h}lon* 'relative'

e. *h_2arh_3trom > Gk. *arotron*, OIr. *arathar*, Arm. *arawr* 'plow'

f. *(h_x)i̯enh$_2$trih$_2$- > Lat. *ianitrīcēs* 'brothers' wives'

g. *$térh_2trom$ 'auger' > Gk. *tératron*, OIr. *tarathar*

It would be difficult to explain away by analogy a number of the counterexamples in (43), such as *$\hat{k}erh_2.srom$, *$temh_1.sreh_2$ and *(h_x)i̯enh$_2$.trih$_2$; this fact, combined with the threefold attestation of the highly archaic *$h_2arh_3.trom$, makes it very likely that the sequence RH.CC did not undergo a regular rule of laryngeal loss in PIE.

3.3.3 *RF$.

If the laryngeal loss rule only applied to sequences of the shape *PH.CC, how does one explain the loss of laryngeal in *janman-* and *génnā* (< *$\hat{g}énh_1mnV$-) discussed above? Here, if laryngeal deletion had occurred, one might expect to find compensatory lengthening (CL), just as in word-final position (*u̯édorh$_2$ > *u̯édōr), since the laryngeal would not have been extrasyllabic and therefore would have carried a mora. The fact that we do not (*$\hat{g}énh_1mnV$- > *janman-*, not ***jānman-*) provides us with the clue for understanding this phonological paradox.

No CL occurs due to the (P)IE tendency to avoid superheavy syllables (syllables consisting of more than two morae) in medial position, resulting in the loss of a mora. As we will discuss in greater detail in Chapter 4, this tendency also underlies Schwebeablaut, Osthoff's Law, the replacement of certain *e*-grade oblique stems with ø-grade forms, and Sievers' Law. Perhaps a conflict between the avoidance of mora loss and the avoidance of a superheavy syllable was the motivation for variation in the proto-language. Either speakers deleted a laryngeal and lost a mora (*$\hat{g}enh_1.mn$- > *$\hat{g}en.mn$-

$\rightarrowtail$ *janman-*) or dealt with the disfavored cluster and retained the moraic structure (**ĝénh$_1$.trih$_2$- > jánitrī-*). This variation may be best represented with an OT analysis using the following four constraints:

(44) CONSTRAINTS FOR *RH$ ANALYSIS:

a. *RH$: The output may not have the sequence sonorant + laryngeal immediately preceding a syllable boundary. Assign one $*$ for each violation.

b. MAX-H: Every laryngeal consonant in the input has a correspondent in the output. Assign one $*$ for every loss.

c. *SUPERHEAVY: No medial syllable may consist of three or more morae. Assign one $*$ for each violation.

d. MAX-μ: Every mora in the input has a correspondent in the output. Assign one $*$ for every loss.

A vacillation of two constraints in the PIE grammar gives us the source of the conflicting set of data attested in the IE languages, a vacillation, which, as we will see, is likely to be attributed to a reranking within late PIE. The grammar in which laryngeal deletion occurred had the constraint ranking *RH$ $\gg$ MAX-H, which means that it was more important for PIE speakers to avoid the coda sequence *RH$ than it was for them to retain a laryngeal. The reason **[ĝēn]$_\sigma$[trih$_2$]$_\sigma$ is not chosen results from the underlying tendency within the grammar to avoid superheavy syllables, represented by the constraint ranking *SUPERHEAVY $\gg$ MAX-μ.

(45) VARIANT ONE: /ĝenh$_1$trih$_2$/ → [ĝen]$_\sigma$[trih$_2$]$_\sigma$

/ĝenh$_1$trih$_2$/	*RH$	MAX-H	*SUPERHEAVY	MAX-μ
a. [ĝenh$_1$]$_\sigma$[trih$_2$]$_\sigma$	*!		*	
b. ☞ [ĝen]$_\sigma$[trih$_2$]$_\sigma$		*		*
c. [ĝēn]$_\sigma$[trih$_2$]$_\sigma$		*	*!	

In the grammar that does not delete a laryngeal the constraint ranking was MAX-H ≫ *RH$, which indicates that it was more important for speakers not to delete a laryngeal than it was for them to avoid the sequence *RH in the coda.

(46) VARIANT TWO: /ĝenh$_1$trih$_2$/ → [ĝenh$_1$]$_\sigma$[trih$_2$]$_\sigma$

/ĝenh$_1$trih$_2$/	MAX-H	*RH$	*SUPERHEAVY	MAX-μ
a. ☞ [ĝenh$_1$]$_\sigma$[trih$_2$]$_\sigma$		*	*	
b. [ĝen]$_\sigma$[trih$_2$]$_\sigma$	*!			*
c. [ĝēn]$_\sigma$[trih$_2$]$_\sigma$	*!		*	

Note that the constraint *SUPERHEAVY applies only to medial syllables, which will allow for compensatory lengthening to occur in final syllables: **u̯édorh$_2$* > **u̯édōr*. Otherwise, if *SUPERHEAVY had applied to all syllables, we would expect **u̯édorh$_2$* > ***u̯édor*, a form which did not occur.[12]

[12]Why were superheavy syllables in PIE marked in word-medial position but not word-finally? Given our rule of coda extrasyllabicity (41) above, it would be impossible to assume that all word-final consonants were extrasyllabic (such as */r/ in **p(h$_2$)tēr* 'father'). But can we assume that */r/ was extrametrical? Cf. Watson 2002:92: "In contrast to extrametrical consonants, which link directly with the syllable node of the peripheral syllable, the extrasyllabic consonant falls into . . . a degenerate syllable". Perhaps the solution lies in Lunden 2006, whose work was discussed in fn. 3 above. I leave this problem open for future research.

(47) /u̯édorh$_2$/ → [u̯é]$_\sigma$[dōr]$_\sigma$

/u̯édorh$_2$/	*RH$	MAX-H	*SUPERHEAVY	MAX-μ
a. [u̯é]$_\sigma$[dorh$_2$]$_\sigma$	*!		*	
b. ☞ [u̯é]$_\sigma$[dōr]$_\sigma$		*		
c. [u̯é]$_\sigma$[dor]$_\sigma$		*		*!

The PIE loss of a laryngeal in the word-final sequence *-VRH# with subsequent compensatory lengthening was first recognized by Szemerényi (1979:155,159).[13] As we would expect from the above analysis of medial *RH$ sequences, there also seems to have been variation in word-final syllables of this shape. One could cite the Skt. form *nā́māni* 'names (nom./acc.)' as evidence of variation, though the lengthened vowel of the suffix (*-ān-*) undoubtedly derives from **-onh$_2$* by Szemerényi's Law, with the **-h$_2$* (or its later outcome) being analogically reintroduced from other neuter paradigms such as vowel-final and obstruent final stems, where the **-h$_2$* had not been lost. More compelling evidence of the non-application of Szemerényi's Law in PIE may be cited in the archaic forms Skt. *jáni-*, YAv. *jaini-*, Arm. *kin*, TA *śaṃ*, TB *śana* from **g^{w}énh$_2$* 'woman'[14] and Hitt. *kit-kar* 'at the head', which derives either from *kēd* + **k̂érh$_2$* 'head (loc.sg.)' or *kēd* + **k̂arh̯* (< **k̂r̥h$_2$*) 'id.'.[15]

Given the existence of variation within sequences of the shape *-VRH$ in both medial and final position, it is conceivable that both rules were driven by the same phonological phenomenon. This view is only made possible through the assumption

[13] Discussed in further detail by Nussbaum 1986:129ff. and Jasanoff 1989:137.

[14] See NIL 177-85 for discussion and references for each form. Contrast OIr. *bé* < **g^{w}en* < **g^{w}ēn* < **g^{w}énh$_2$*, which according to Jasanoff 1989:140, provides evidence for an earlier, laryngeal-less form produced by Szemerényi's Law.

[15] Nussbaum 1986:96-99. Whether from PIE **k̂érh$_2$* or much later **k̂arh̯*, Szemerényi's Law ceased to exist at some point within the prehistory of Hittite.

that there was a tendency within PIE to avoid superheavy syllables in word-medial position. Just as there were root variants *$\hat{g}$*en-* (Skt. *janman-*), *$\hat{g}enh_1$- (Skt. *jánitrī-*) in medial position in PIE so were there variants *$g^{u̯}ēn$ (OIr. *bé*), *$g^{u̯}enh_2$ (Skt. *jáni-*) in final position. The forms without a laryngeal are the more opaque and therefore the more archaic; this entails that the constraint ranking *RH$ ≫ MAX-H was older than the ranking MAX-H ≫ *RH$ in PIE. Like the nt. nom/acc. pl. endings in *-V̄R, the laryngeal-less *CERH roots should be viewed as archaims and were likely to have been lexicalized at a later stage of PIE.

Szemerényi's Law is also found in the environment *VRs#, with loss of **s* and subsequent compensatory lengthening (→ *V̄R#).[16] Should one assume that the laryngeals were all fricatives in manner (see section 1.2 for discussion), then both processes in question may be collapsed as targeting the word-final sequence *-VRF#.

(48) SZEMERÉNYI'S LAW.
PIE *VRF → V̄R / _#
In a word-final sequence of vowel + sonorant + fricative (*/s, h_1, h_2, h_3/), the fricative is lost with subsequent compensatory lengthening on the preceding vowel.

Just as in sequences of the shape *-VRH#, Szemerényi's Law ceased to be productive in those of the shape *-VRs# – forms such as *$u̯ih_1roms$[17] 'men' (not **$u̯ih_1rōm$), **sals*[18] 'salt' (not ***sāl*) and **séms*[19] 'one (masc.nom.sg.)' (not ***sḗs*) are frequently reconstructable. Those instances where Szemerényi's Law did apply, such as PIE

[16]Szemerényi 1970:109.

[17]Goth. *wairans*.

[18]Gk. *háls*, Lat. *sāl*.

[19]Gk. *heís*, not ***hḗs*.

$p(h_2)t\acute{\bar{e}}r$ (< *$p(h_2)t\acute{e}rs$*), *$d^h\acute{e}\hat{g}^h\bar{o}m$* (< *$d^h\acute{e}\hat{g}^h oms$*) and the third plural perfect ending *-ēr* (< *-ers*)[20], were all morphologized and the long vowel produced was no longer synchronically viewed as the outcome of a process of **s*-loss.

If Szemerényi's Law did apply within sequences of the shape *-VRH# and *-VRs#, and if codas of the shape *RH]$_\sigma$ were also avoided word-medially, then we would also expect there to have been instances of deletion in sequences of the shape *RsCC, as there were in those of the shape *RHCC. Unfortunately I have not yet found any such examples of **s* deletion in PIE. However, it is curious to note that an examination of the forms listed in the IEW, LIV and NIL reveals a number of roots of the shape **-Rs*, such as *d^hers-* 'take courage',[21] **ĝeu̯s-* 'taste'[22] and **(s)kers-* 'cut',[23] though only a fraction attest derivatives with the sequence RsCC. However, in almost each instance the attested RsCC sequence is secondary. Lith. *žiezdrà* 'gravel; grain', *žiẽz(g)dros* 'gravel' to the root **ĝei̯s-* 'gravel' (IEW 356); PSlav. **męzdra* continued by RCS *męzdrica* 'egg shell' and Russ. *mjazdra* 'the flesh-side of the hide' to **mēms-* 'meat' (IEW 725); and OPruss. *tiēnstwei* 'excite' from **tens-* 'thin out' (IEW 1069), are all formations which cannot be reconstructed back to PIE. In fact, I have found only one reconstructable sequence of the shape *RsCC: *(h_1)oi̯s-tro-/-trah_2*, continued by Gk. *oístros* 'rage', Lith. *aistrà* 'vehement passion' and Lith. *aistrùs* 'passionate', formed to the root *$h_1$eis-*, as seen in Lat. *īra* 'anger' (Plautus *eira*).[24]

It is unclear whether the paucity of reconstructable sequences of the shape *RsCC is of any significance, though it would be explained nicely by a general avoidance of the shape *RF$. If I am correct that Szemerényi's Law applied also in word-medial

[20] Seen in Lat. *-ēr-e* and likely Hitt. *-er, -ir* (Jasanoff 2003:32).

[21] IEW 259; LIV 147; NIL 120-2. Ved. *dhr̥ṣṇóti* 'is courageous', Goth. *ga-dars* 'dares'.

[22] IEW 399-400; LIV 166-7. Ved. *juṣāṇá-* 'taking pleasure', Goth. *kausjan* 'taste; meet'.

[23] IEW 945; LIV 355-6. Hitt. *karašzi* 'cuts, fails', TB *śarsa*, TA *śärs-* 'knew'.

[24] IEW 299ff.; Olsen 1988:16.

position, then we may tentatively reformulate the rule as follows:

(49) SZEMERÉNYI'S LAW (REVISED).

$$\text{PIE *VRF} \rightarrow \begin{cases} \bar{\text{V}}\text{R} / \_\# \\ \text{VR} / \_\$\text{C}_0 \end{cases}$$

A fricative is deleted in a coda sequence of the shape vowel + sonorant + fricative (*/s, h_1, h_2, h_3/), with compensatory lengthening on the preceding vowel if lost word-finally.

This rule was no longer productive in late PIE, which explains the variation found in the attested IE languages.

3.3.4 Extrasyllabicity Test #1: Monosyllabic Lengthening.

Returning now to the more general problem of reconstructing word-final extrasyllabicity in PIE, it is helpful that there exists an independent, non-circular method to test for coda extrasyllabicity in languages where monosyllabic lengthening is a synchronic phonological process (Itô 1988:123). This phonological process is employed by languages to make light monosyllables heavy, in order to satisfy a minimal word requirement that demands that a lexical item consist of at least two morae.[25]

(50) MINIMAL WORD REQUIREMENT.

Any word bearing stress must consist of at least two morae.

For example, in Old Irish, vowels found in monosyllables are lengthened if the word bears stress. Thurneysen (1946:32) lists *sé* besides *sessed* 'sixth', *mé* 'I' besides the emphasizing particle *messe*, *gé* 'pray (3 sg. subj.)' besides *gessam* (1 pl. subj.), *tó* 'yes' and *trú* 'doomed person (nom.sg.)' beside *troch* (gen.sg.). However, in unstressed

[25] See McCarthy & Prince 1986:6-7 for discussion, with references.

monosyllables (enclitics and proclitics) the vowel is never lengthened, as we see in the unstressed clitic *de* ‘from him, it’ (not ***dé*). Whether the lengthened vowel in the monosyllable is historically or synchronically derived is irrelevant; *synchronically* there are no stressed monosyllables in Old Irish that consist of a single mora.

As for extrasyllabic coda consonants, if a word of the shape #CV# lengthens to #C$\bar{V}$# and #CVX# also lengthens to #C$\bar{V}$X# by a process of monosyllabic lengthening, we may assume that X is extrasyllabic. For example, in Ponapean, a Micronesian language spoken primarily on the island of Pohnpei and the Caroline Islands, one finds vowel lengthening in certain monosyllabic nouns ending in a consonant (Rehg & Sohl 1981:117; McCarthy apud Itô 1988:123ff.).[26]

(51) Ponapean Monosyllabic Noun Lengthening.

a. /pik/ → [piik] ‘sand’ (vs. inflected *pik-en* ‘sand of’)

b. /pet/ → [peet] ‘bed’ (< Eng. *bed*)

but

c. /keep/ → [keep] ‘yam’, not **[keeep]

d. /kent/ → [kent] ‘urine’, not **[keent]

In Ponapean, a monosyllabic noun of the shape CVC behaves as if it is light, or consisting of only one mora, while a noun of the shape CVVC and CVCC behaves as if it is heavy. McCarthy argues that by assuming the final consonant of nouns to be extrasyllabic, the forms presented in (51) may be syllabically parsed as $[\text{pi}]_\sigma$k, $[\text{pe}]_\sigma$t, $[\text{kee}]_\sigma$p, $[\text{ken}]_\sigma$t, etc. This is why the process of monosyllabic lengthening, or the lengthening of a vowel to satisfy the minimum word requirement, occurs only in monosyllabic nouns of the shape CVC.

[26]Vowel-final roots are inherently long in Ponapean (e.g. /ntā/ ‘blood’), which ironically have shortened within the history of the language: /ntā/ ‘blood’ > [nta], through a more general process of apocope (Rehg & Sohl 1981:89).

As early as Hirt (1921-37:II/227), it has been suggested that a process of monosyllabic lengthening existed in PIE, in order to account for long and short variants of some very common Indo-European words. Sihler (1995:38) proposes that this lengthening occurred only when PIE monosyllables were stressed (cf. Old Irish above). If a minimal word requirement was present in the proto-language, a restriction of vowel lengthening to stressed monosyllables is to be expected, since lexical items always require accentuation, while grammatical items such as clitics do not.[27] Kapović (2006:151[442]) argues that Sihler's explanation of monosyllabic lengthening is impossible, since there is lengthening of the 1st pl. pronominal clitic **nōs* 'us' found in Latin *nōs* and Slavic *na-*, which should never have occurred, since, having been a clitic, it was always unaccented. Kapović's argument, however, is not fatal to Sihler's suggestion, since as we will see, both instances of lengthening in Latin and Slavic may be viewed as secondary within the individual prehistories of each language.

Kapović (2006) presents an excellent overview of the problem of monosyllabic lengthening and convincingly argues for the reconstruction of this phonological process for PIE. The most solid examples are presented in (52) below and are primarily taken from Kapović (2006:147ff.). Examples (vi) - (ix) of set (a) and (iv) of set (b) have been graciously suggested to me by Brent Vine.

(52) Possible examples of lengthened monosyllables in PIE.

a. In open syllables.

i. ****me*** 'me (acc.)' (Gk. *emé*, *me*, OIr. *mé*, Goth. *mi-k*) vs. ****mē*** (Skt. *mā*, Av. *mā*, Lat. *mē-d*, Gaulish *mi*, Welsh *mi*)[28]

ii. ****nu*** 'now' (Skt. *nú*, Gk. *nu*, Lat. *nu-nc*, OIr. *nu*, *no*, Latv. *nu*) vs. ****nū***

[27] Cf. Blevins & Wedel 2009:158.

[28] IEW 702.

(Skt. *nū́*, Av. nū, Gk. *nũn*, OCS *ny-ně*, Lith. *nū-naĩ*)[29]

iii. ***tu*** 'you' (Gk. *sú*, Latv. *tu*, OE *þu*, OIr. *tu-ssu*, Hitt. *tu-k*) vs. ***tū*** (Lat. *tū*, Hom.Gk. *tū́nē*, Av. *tū*, OCS *ty*, OE *þū*)[30]

iv. ***ne*** 'not' (Skt. *ná*, OCS *ne*, Lat. *ne-que*, Goth. *ni-h*) vs. ***nē*** (Lat. *nē*, OIr. *ní*, OCS *ně-*)[31]

v. ***t(u̯)e*** 'you (acc.)' (Gk. *sé*, *té*) vs. ***t(u̯)ē*** (Skt. *tvā́*, Av. *θβā*, Lat. *tē-d*)[32]

vi. ***b^{h}e*** 'emphasizing particle' (East Lith. *bè*), ***b^{h}o*** (Goth. *ba*, OCS *bo*) vs. ***b^{h}ē*** (perhaps Av. *bā*), ***b^{h}ō*** (Pol. *ba*) [33]

vii. ***de*** 'directional particle' (Lat. *quan-de*, *in-de*), ***do*** (OCS *do*) vs. ***dē*** (OIr. *dí*, Gk. *dḗ*, Lat. *dē*), ***dō*** (Lat. *dōnec*, OIr. *do, du*, PGmc. **tō*)[34]

viii. ***ku*** 'interrogative particle' (Skt. *kútaḥ*, Av. *kuθa*) vs. ***kū*** (Skt. *kū́*, Av. *kū*)[35]

ix. ***u̯e*** 'disjunctive particle' (Lat. *-ve*, Skt. *i-vá*) vs. ***u̯ē*** (Skt. *vā*, Av. *vā*, Gk. *ḕ-(w)e*, *ḕ*)[36]

b. In syllables with one coda consonant.

i. ***úd*** 'on high' (Skt. *úd*) vs. ***ūd*** (Gmc. **ūt* > Goth. *ut* [ūt], Eng. *out*, Germ. *aus*)[37]

---

[29]IEW 770.

[30]IEW 1097.

[31]IEW 756-8.

[32]IEW 1097.

[33]IEW 113.

[34]IEW 181ff.; de Vaan 2008:162.

[35]IEW 647ff.

[36]IEW 75.

[37]IEW 1103-4.

ii. ***u̯os*** ‘y’all’ (Skt. *vas*, YAv. vō, GAv. *və̄*) vs. ***u̯ōs*** (Lat. *vōs*, OCS *va*)[38]

iii. ***nos*** ‘we’ (Skt. *nas*, Hitt. *-naš*, YAv. *nō*, GAv. *nə̄*, Alb. *na*, Gk. *nós-p*^{h}*in* ‘apart (from)’) vs. ***nōs*** (Lat. *nōs*, OCS *na*)[39]

iv. ***ub*** (Lat. *sub*) vs. ***ūb*** (OHG *ūf*, OCS *vy-sokъ*)[40]

Other purported examples of monosyllabic lengthening in PIE include **mūs* ‘mouse’, which has traditionally been connected to **mus-* ‘steal’ (‘mouse’ < ‘the thief’) and **i̯ūs* ‘y’all’ (Av. *yūš*, Skt. *yūyám*, Lith. *jū́s* and OCS *vy*), from an earlier **i̯us*. However, the root ‘steal’ clearly is to be reconstructed as **mush*$_{x}$-,[41] in order to account for seṭ-forms such as *moṣīs* ‘steal (2nd sg. injunctive)’. Therefore, the root ‘steal’ does not give us any particular reason to assume monosyllabic lengthening in **mus* ‘mouse’ versus a straightforward development of **u* plus laryngeal (**muh*$_{x}$*s*).[42] In addition, the pronoun **i̯ūs* has no attested short variant, and as Kapović himself admits (2006:148[431]), may just as easily derive from PIE **i̯uh*$_{x}$*s*.

If monosyllabic lengthening did exist as a synchronic phonological process within PIE, our rule of coda extrasyllabicity should show precisely where it occurred. We have seen that in PIE, extrasyllabic consonants were only present in codas when they violated the SSP. Thus, in order for a consonant to be extrasyllabic, another consonant would necessarily precede it, and therefore any word-final syllable with extrasyllabic consonant(s) would ALWAYS be heavy. So, if our rule of coda extrasyllabicity (41) is correctly formulated, monosyllabic lengthening should only be found in words of

---

[38]IEW 514.

[39]IEW 758.

[40]IEW 1106-7; Derksen 2008:535.

[41]Cf. LIV 445.

[42]If ‘steal’ and ‘mouse’ are to be connected, a possible explanation may be found in Olsen 1999:130-1, with references to previous literature.

the shape CV, which would explain the long variants of *me* 'me', *nu* 'now', *tu* 'thou', *ne* 'not', *t(u̯)e* 'thee', *b^he/*b^ho* 'emphasizing particle', *de/*do* 'to', *ku* 'interrogative stem' and *u̯e* 'or'.

In (52b) above, there are four possible examples of monosyllabic lengthening in words ending in a single final consonant. Note, however, that all of the long variants occur in Western Indo-European (sub-)branches: Italic, Slavic and Germanic. In Latin, monosyllables of the shape /CVs/ always have a long vowel (cf. *ās* 'as, a type of coin', *pēs* 'foot', *vīs* 'strength', *ōs* 'mouth', *mūs* 'mouse', etc.),[43] and therefore the forms *nōs* and *vōs* provide no evidence for the original vowel length in PIE, as any monosyllable of the shape /CVs/ would automatically have been realized as [CV̄s] in Latin. In Slavic, coda consonants were completely eliminated by the Law of Open Syllables (Carlton 1990:100), and so it is conceivable (though ad hoc) that in a prehistoric Slavic grammar, either the loss of *-*s* (via an intermediate stage *-*h*) had resulted in compensatory lengthening in this particular form: **nos* > **noh* > **nō* > *na*, or **s*, at some point along path to zero, had become extrasyllabic, at which time monosyllabic lengthening had taken place: *$[\text{nos}]_\sigma$ > *$[\text{no}]_\sigma$s > *$[\text{nō}]_\sigma$s > $[\text{na}]_\sigma$.[44] Neither are particularly compelling scenarios since, to my knowledge, there are no parallel examples. Perhaps, however, an ad hoc explanation may be maintained since we are dealing with a high-frequency word, a type which tends to undergo unusual phonological developments.[45] Though I currently do not have an explanation for the vowel lengthening found in the Germanic forms **ūt* 'out' and **ūp* 'up', the reader should note that monosyllabic lengthening did not occur in PGmc. *h^wát* 'what' (< PIE *k^wód*; not PGmc. **h^wṓt*) or

[43]Brennan 2006. For a detailed discussion of word minimality in Latin, see Mester 1994:19ff.

[44]A similar derivation would also be assumed for PIE **ub*.

[45]In other words, high-frequency words are more likely to undergo lexical diffusion. Cf. Labov 1994:483.

**át* ‘at’ (< PIE **ád*; not PGmc. ***ṓt*), two words of similar shape (Ringe 2006:98).[46]

3.3.5 Extrasyllabic Consonants in Onset Position.

As in PIE word-final codas, multiple violations of the SSP were allowed in PIE word-initial onsets. Bipartite onsets include **stah*$_2$- ‘stand’, **u̯reh*$_1$- ‘find’, **h*$_2$*ḱou̯s*- ‘hear’, and tripartite onsets include **psten*- ‘breast’, **strew*- ‘strew’, **h*$_2$*ster*- ‘star’, **h*$_1$*sti*- ‘existence’. The key difference between medial onsets and codas, however, lies in the fact that the SSP may be violated in bipartite medial onsets: **d*^{h}*ug.h*$_2$*ter*-, **h*$_1$*et.sḱé/ó*- ‘eat (iterative)’, **h*$_2$*uk.sto*-[47] ‘grown’, **pn̥k.sti*- ‘fist’,[48] **s(u̯)ek.sto*-[49] ‘sixth’. This suggests that an SSP violation was allowed in a PIE bipartite onset. Note that **d*^{h}*ugh*$_2$*ter*-, **h*$_1$*etsḱé/ó*-, **h*$_2$*uksto*-, etc. could not have been syllabified as **d*^{h}*ugh*$_2$*.ter*-, **h*$_1$*ets.ḱe/o*-, **h*$_2$*uks.to*-, because there would have been an SSP violation in the coda; see rule (41) above.

3.3.6 Extrasyllabicity Test #2: Reduplication.

In section 3.3.4 we saw that monosyllabic lengthening provides us with an independent method of determining coda extrasyllabicity in languages that have a Minimal Word Requirement. Depending on one’s theoretical view of reduplication, there may also exist an independent test of onset extrasyllabicity: the analysis of a language’s redupli-

[46] Perhaps the presence of **u* in these protoforms – being inherently shorter and less sonorous than **a* – is a factor here?

[47] Continued by Dardic, Pashai (dialectal) *ūṣ/ux* ‘long’ (Turner 1966:74, Nr. 1627). Sanskrit **uṣṭa*-, the expected reflex of **h*$_2$*uksto*-, has been replaced by *ukṣita*- (NIL 354-5).

[48] OHG *fūst*, OE *fȳst*, OBulg. *pęstь* (IEW 839).

[49] Lat. *sextus* (← **sektos* < **sekstos*), Gk. *héktos*, Ved. *ṣaṣṭhá*-, Goth. *saihsta*, Lith. *šẽštas*, TB *ṣkaste* (Weiss 2009:293).

cation patterns.[50] According to this particular theoretical view, if a certain consonant occurs in absolute root-initial position and participates normally in the reduplication process, we may say that that consonant is syllabifiable in that language. Conversely, if a particular root-initial consonant is not copied into the reduplicant then we may say that that consonant is extrasyllabic.

For example, Sanskrit roots with no extrasyllabic segments in the onset reduplicate the initial consonant (*prā-* 'fill' → *pi-pra-*) while those roots with extrasyllabic consonants in the onset reduplicate the second (*sthā-* 'stand' → *tí-ṣṭhati* 'stands' not **sí-ṣṭhati*).[51] As Steriade (ibid.) shows, only those roots with an SSP violation in the onset behave differently from the normal pattern of reduplication, since those violating consonants are extrasyllabic. In (Attic) Greek, reduplication is blocked for onsets with extrasyllabic consonants: the perfect of *gnō* 'know' is *é-gnōka*, not **gé-gnōka*, since *#gn-* contains an extrasyllabic *g* that blocks reduplication.[52]

Thus analyzing how reduplication worked in PIE may be able to tell us in a non-circular fashion whether there were extrasyllabic consonants in word-initial position and exactly which consonants could function as extrasyllabic. Fortunately for our study, reduplication is a common morphological process in Proto-Indo-European. It is found in many verbal categories (reduplicated presents, aorists, desideratives and perfects) and very rarely in nominal formations, in such forms as *k^we-k^wlós* 'wheel' (Skt *cakrá-*, Gk. *kúklos*, OE *hwēol*, etc.).

[50] See Steriade 1982:312ff., Keydana 2006, Keydana, forthcoming a. An alternate view is given by Carlson 1997 (see also Morelli 1999, Cho & King 2003 and Vaux & Wolfe 2009) who points out that the choice of consonant in reduplicative templates is rather due to reduplicant-specific markedness constraints and therefore nothing may inferred about the underlying syllabic structure.

[51] Steriade 1982:312ff., Kobayashi 2004:43, Keydana 2006:95-7.

[52] Steriade (1982:221) argues that in Greek adjacent tautosyllabic consonants must be at least four intervals apart on the sonority scale (see top of p. 221). The reason *g* is extrasyllabic in the onset *#gn-* is because there is only a sonority distance of three intervals in this cluster.

For the most part reduplication is a well-understood morphological process in PIE.[53] There are two reconstructable types of reduplication, partial and full, with the latter utilized only in intensive/iterative formations, such as Skt. *dár-darti* ‘pry open’ to *dar-* ‘split’ and Gk. *por-p*^{h}*ū́rō* ‘swell, surge (of the sea)’ to an earlier **p*^{h}*ur-*.[54] All other morphological categories reduplicated partially. Some reduplicated with *e*-vocalism (perfect, aorist, certain presents, and nominal formations), while others showed *i*-vocalism (certain presents and desiderative).

Roots with a single consonant in the onset simply reduplicated that particular consonant, followed by a vowel: **g*^{w}*em-* ‘come’ → **g*^{w}*e-g*^{w}*óm-e* ‘came’ (> Skt. *jagā́ma*); **b*^{h}*er-* ‘bear’ → **b*^{h}*i-b*^{h}*érti* ‘carries’ (> Skt. *bibhárti*). Roots with two consonants in the onset reduplicated with the first: **drem-* ‘run’ → **de-drom-* (> Hom. Gk. *ana-dédrome* ‘towered’); **mnah*$_2$*-* ‘remember’ → **me-mnóh*$_2$*-/me-mn̥h*$_2$*-* (> Gk. *mémnēmai* ‘I remember’); **smei̯-* ‘smile’ → **se-sm(o)i̯-* (> Ved. *siṣmiyāṇa-* ‘smiling’), **h*$_2$*neḱ-* ‘reach’ → **h*$_2$*a-h*$_2$*noḱ-* (> Skt. *ānáṃśa*, OIr. *-ánaic*), etc.[55] Based on these facts, the reduplication template for PIE may be reconstructed as *C$_1$V-C$_1$(C)V- for morphological categories with partial reduplication. Complications arise, however, in the reconstruction of reduplicants for roots whose onsets consisted of /s/ + stop and /H/ + stop, namely, those roots where the first consonant of a bipartite onset was a fricative and violated the SSP.[56]

Let us first address the reduplication template for roots of the shape *sP-. Here, the uncertainty for reconstruction lies in the many different types of reduplicants attested in the IE languages. As we saw, Sanskrit shows *PVsP-* (cf. *tí-ṣṭ*^{h}*āti* ‘stands’ to *st*^{h}*ā-*)

[53] For excellent and thorough discussions of the process of reduplication in PIE, see Brugmann 1897-1916:20-41, Niepokuj 1997 and Keydana 2006.

[54] Cf. Hitt *paraḫzi* ‘chases’ < **b*^{h}*erh*$_2$*-* ‘move quickly’.

[55] Kümmel 2000:287.

[56] See Keydana 2006.

while its most closely related branch, Iranian, attests to *sVsP-* (Av. *hi-štaiti*, *vi-šastarᵊ*, OPers. *a-hi-štatā*, all to *stā-* 'stand'), as does Celtic (cf. OIr. *se-scaind* 'jumped' < **se-skond-e*). In non-productive reduplicated *i*-presents in Greek, we find *sVsP-* (cf. *hí-stēmi*),[57] while productive perfect formations block reduplication altogether (*e-stratóōnto* 'was on the battlefield', with simple prefixed *e-*). In Latin, the highly archaic and non-productive reduplicated present *si-stō* (= Umbr. **se-stu**) attests to *sVsP-*, while later productive reduplicants to roots of the shape *sPV-* exhibit a curious process of reduplicative infixation: *stā-* 'stand' → *ste-t-*; *spond-* 'libate' → *spo-po-nd-*, etc.[58] In Gothic the reduplicative template is always *sPVsP-* (*stai-stald* to *stald-*, etc.),[59] while in Tocharian, we find three reduplicant shapes attested: in A, *sV-* (*spārtw*$^{(\bar{a})}$- 'turn, behave, be' → *sā-spärtwu*), and in B, *PV-* (*spārtt*$^{(\bar{a})}$- 'turn, behave, be' → *pa-sparttau*) and *sPV-* (*stäm*$^{(\bar{a})}$- 'stand' → *śce-ścamoṣ*; *staukk*$^{(\bar{a})}$- 'swell' → *sta-staukkauwa*).[60]

Many scholars[61] have viewed this disparity among reduplicants in the IE languages as proof that the reduplication template for roots of the shape **sP-* was **sPVsP-*, exactly as is attested in Gothic. The basic idea is that those languages whose reduplicants have a single consonant (**s* or **P*) have simplified the original, more complex reduplicant beginning in **sP-* through a process of dissimilation: PIE **sti-stah*$_2$*-ti* 'stands' > **si-stah*$_2$*-ti* or **ti-stah*$_2$*-ti*.[62] However, as early as Brugmann,[63] it has been recognized that since Lat. *sistō* and Gk. *hístēmi* do not follow their respective synchronic pattern

[57] As well as the archaic *hé-stēka* 'I stood'.

[58] Fleischhacker 2002:5. See Keydana, forthcoming a for an alternative analysis.

[59] I follow Keydana 2006 and forthcoming a, who assumes a subsegmental status of /s/ in /sP/ clusters in Gothic (suffricates), which easily accounts for the entire copy of the initial cluster of the root.

[60] Forms taken from Malzahn 2009:964, 963, 959-60 and 990, respectively.

[61] For instance, Szemerényi 1970:249 and LIV 590.

[62] Keydana 2006:81 convincingly dispels this notion, demonstrating that dissimilation cannot produce the attested reduplicants as long as reduplication is still an active morphological process.

[63] 1897-1916:40-1.

of reduplication, they must be archaisms derived from an older reduplicative template **sVsP-*. It is simplest to assume that both forms go back to a common archaism, and that the reduplication pattern found in Avestan and Old Irish continues the original state of affairs.

Those scholars who reconstruct a PIE **sti-stah*$_2$*-ti* 'stands' typically see confirmation in their reconstruction of the reduplicant to **h*$_1$*ger-* 'awaken', the only root of the shape *HP- with an attested reduplicated form (Kümmel 2000:191-4). There are three attested reflexes of the perfect to **h*$_1$*ger-* 'awaken': Skt. *jāgā́ra* 'wakes, is awake', Av. *jaɣāra* 'is awake' and Gk. *egrḗgore* 'is awake' (LIV 245-6). The Avestan form must be secondary, having eliminated the lengthened grade in the reduplicant by analogy to other perfects. The Greek and Sanskrit forms, however, appear to derive from an original **h*$_1$*ge-h*$_1$*gor-e* 'woke up', with the [r] in the Greek reduplicant secondary by analogy to the aorist *égreto* (LIV 245).

However, as Keydana (2006:104-5) demonstrates, the Sanskrit form is more easily explained as deriving from **gēgor-* < **ge-Hgor-*, following the reduplication template **PVHP-*, analogous to the **PVsP-* reduplication template discussed above. Just as *PVsP-* replaced **sVsP-* as the reduplication template to roots of the shape *sP-* in the prehistory of Sanskrit, so may we posit that **PVHP-* replaced an earlier template **HVHP-* to roots of the shape **HP-*. Similarly, Greek *egrēgore* may be viewed as a later formation, remade from original **ēgore* (< **h*$_1$*e-h*$_1$*góre*) and re-marked as a reduplicated form through the process of Attic reduplication (cf. Rix 1992:204-5), as has occurred, for example, in the perfect of **h*$_1$*nek̂-* 'take': **h*$_1$*e-h*$_1$*nok̂-* > **ēnok̂-*[64] → Att. *en-ēnok*^{h}*a* 'has carried' (LIV 250). Likewise, **h*$_1$*e-h*$_1$*góre* > **ēgore* → **eg-ēgore*, with the *-r-* secondarily inserted into the reduplicant.

Note also that if the reduplicant to **h*$_1$*ger-* 'awaken' had consisted of two conso-

---

[64] cf. Skt. *ānáṃśa*, OIr. *-ánaic* < **h*$_2$*a-h*$_2$*nok̂-*

nants in the onset, it would be the sole example of such a reduplication type reconstructed for PIE, since, as we saw, roots of the shape *_sP_- reduplicated with a simple *_sV_-. This makes it increasingly likely that a single reduplication template existed in PIE: *C_1V-C_1(C)V-. In short, PIE reduplication proceeded in the same way for all bipartite onsets: *_pleh_$_1$-'fill' → *_pi-pleh_$_1$- (Gk. _pímplānō_ 'fill'), *_stah_$_2$- 'stand' → *_si-stah_$_2$- (Lat. _sistō_, Gk. _hístēmi_, etc.), *_u̯reh_$_1$_ĝ_- → *_u̯e-u̯roh_$_1$_ĝ_- (Gk. _érrōge_), *_h_$_2$_nek̂_- 'reach' → *_h_$_2$_a-h_$_2$_nok̂_- (> Skt. _ānáṃśa_, OIr. _-ánaic_), and *_h_$_1$_ger_- 'awaken' → *_h_$_1$_e-h_$_1$_gor_-.[65]

3.3.7 The Rule of Onset Extrasyllabicity.

Reduplication tests indicate that because the reduplication pattern in PIE was the same for all roots with bipartite onsets, an SSP violation was allowed in the PIE onset, unlike in Sanskrit. This is corroborated by allowance of an SSP violation in word-medial onsets: *_d_^{h}_ug.h_$_2$_ter_-, *_h_$_1$_et.sk̂e/o_- 'eat (iterative)', *_s(u̯)ek.sto_- 'sixth'.[66] Curiously, none of the verbal roots with tripartite onsets have reduplicated forms; we cannot reconstruct *_pipsterti_ 'sneezes', *_sestrou̯e_ 'strewed' or the like.[67] If not simply by chance, it is likely these roots possessed an extrasyllabic segment that blocked reduplication,

---

[65]Cf. Keydana 2006:66-7 & forthcoming a.

[66]SSP violations in word-medial onsets are not allowed in many of those languages where the synchronic reduplication template for roots of the shape #sP- has diverged from that of the protolanguage due to onset extrasyllabicity: Lat. _steti_, Gk. _estratóōnto_, Skt. _tíṣṭhati_. Thus, _sP_- was licensed as an extrasyllabic onset at word's edge (_stāre_, _statós_, _sthitá-_) but blocked word-medially, reflected by the following cluster simplifications: _Sestius_ from _Sekstius_ (the _k_ in _sextus_ is analogical; Weiss 2009:375), Gk. _-istos_ and Skt. _vittá-_ from _u̯itstós_ 'known' (for Sanskrit, cf. Kobayashi's (2004:38) PRINCIPLE OF COHESIVE CLOSURE: "In Indo-Aryan, the closure of two plosives in the same consonant cluster should not be interrupted by a continuant consonant").

[67]Though reduplicated forms of *_kseu̯_- 'sharpen' are found in both Av. _kuxšnuuąna-_ (to _xšnu-_ 'agitate') and Skt. _cukṣṇāva_ 'whet (3rd sg. perf.)', it is unclear whether this provides a direct example of reduplication in a tripartite onset for PIE, as *_kseu̯_- is attested as a verbal root only in Indo-Iranian. If it is necessary to reconstruct a *_ke-ksnou̯_- 'sharpen (perfect)' back to PIE, then perhaps at least three consonants were allowed in the PIE onset, though the first (of three) could not violate the SSP. Thus, *_ksneu̯_- → *_ke-ksnou̯_- but *_pster_- 'sneeze' could not form a perfect *_pe-pstor_-.

just as in Greek above. This, coupled with the fact that there is no reconstructable medial onset consisting of more than two consonants, allows us to postulate a rule of onset extrasyllabicity for PIE.

(53) Rule of Onset Extrasyllabicity in PIE.
The maximal onset in PIE consisted of two consonants. Any consonant preceding this sequence is to be considered extrasyllabic.

The following underlined consonants were extrasyllabic: *<u>p</u>sten-, *<u>s</u>treu̯-, *<u>h</u>₂ster-, *<u>k</u>ʷspent-. Note that while *pst- was a legal word-initial onset in PIE, *ptr- was not (see 2.3.1 above). It appears that – with the sole exception *$\hat{g}^h d^h$i̯és 'yesterday' (Skt. *hyáḥ*, Gk. *$k^h t^h$és*) – in all legal tripartite onsets, either the first or second consonant was an *s.[68] At this time I am unsure of this fact's significance: was the presence of *s a requirement for extrasyllabicity or merely a coincidence?[69]

3.4 Review of PIE Extrasyllabicity.

Our rules of coda and onset extrasyllabicity ((41) and (53) above) correctly predict 1) deletion in the coda because of SSP violation (*$[d^h ug]_\sigma h_2[trés]_\sigma$ > *$[d^h uk]_\sigma[trés]_\sigma$ 'daughter (gen.sg.)'), 2) retention in the onset despite SSP violation (*$[h_1 ed]_\sigma[s\hat{k}é/ó]_\sigma$ > *$[h_1 et]_\sigma[s\hat{k}é/ó]_\sigma$ 'eat (iterative)') and 3) deletion in the coda & retention in the onset (*$[Vd^h]_\sigma h_1[s\hat{k}é/ó]_\sigma$ 'put (iterative)' > *$[Vt]_\sigma[s\hat{k}é/ó]_\sigma$).[70] The last example,

[68] Then again, maybe *$\hat{g}^h d^h$i̯és 'yesterday' was NOT an exception, if it was realized as *[$\hat{g}^h zd^h$i̯és], with *s-epenthesis (as per Schindler 1972b; for a different view, see Melchert 2003:153). If this were the case then the syllable structure of 'yesterday' would be *$\hat{g}^h z[d^h$i̯és$]_\sigma$ with two extrasyllabic consonants. For extensive discussion of this form with literature see Vine 2008.

[69] Once again, note that I do not reconstruct *pkten- 'comb' as a legal consonant cluster in PIE, since there is no direct or indirect evidence in favor of such a reconstruction. It is likely that Gk. *kteís, ktenós* 'comb' continues PIE *kten-, with loss of *p according to rule (55) below.

[70] See section 2.2.1 for examples, taken from Hackstein 2002.

*-$Vd^hh_1s\hat{k}$é/ó- 'put (iterative)' > *$Vts\hat{k}$é/ó-, reinforces the fact that only two consonants were allowed in the onset. Although word-initially the onset *h_xst- was legal in *h_2ster- 'star' and *h_1sti- 'existence', word-medially it was not. Thus we have deletion of *h_1 in *-$Vd^hh_1s\hat{k}$é/ó- for two reasons: 1) *h_1 violated the SSP in the coda *-d^hh_1 and 2) a maximum of two consonants was allowed in a PIE onset.

3.5 The Maximum Syllable Template.

Through an examination of reconstructable consonant clusters in PIE, we find that the maximal medial consonant cluster in PIE consisted of four consonants. While *i̯éu̯ktro- 'cord' (NIL 399) was a well-formed PIE word, those such as *i̯eu̯k.stro-, *i̯eu̯rk.stro- were impossible, indicating that the maximum syllable template in PIE was CCVCC.

(54) Maximum Syllable Template (MAXST) in PIE: CCVCC.
The maximum PIE syllable consists of two consonants in the onset and two consonants in the coda. The onset may violate the SSP; the coda may not.

Any violation of the maximum PIE syllable template should result in STRAY ERASURE or STRAY EPENTHESIS, following a modified version of the laryngeal deletion rule given in (26) as proposed in chapter 2 above.

(55) Conditions for PIE Cluster Repair.
If a PIE syllable violates the MAXST and the violating consonants cannot be realized as extrasyllabic, delete a consonant if the result would produce a legal consonant sequence; otherwise, insert a schwa.

Ultimately (55) should allow us to collapse any syllable-driven phonological process of epenthesis and deletion in PIE. In addition to the process of STRAY ERASURE in the sequence *PH.CC and STRAY EPENTHESIS in the sequence *#CHC(C) (cf. *ph*$_2$*trés* and *d*^{h}*h*$_1$*só-* above), it is very likely that the epenthesis of schwa secundum also results from a MAXST violation. This would render schwa primum and schwa secundum the exact same phonological process: fixing a cluster that violates the maximum syllable template with an epenthetic vowel, a process we may simply call "Indo-European schwa".[71]

Moreover, we should expect STRAY ERASURE to eliminate any consonant that violates the SSP in a medial coda. For an excellent example, see Rau's 2009 historical analysis of Ved. *aśītí-* 'eighty', which he straightforwardly reconstructs as **h*$_x$*ok̂(t)h*$_x$*-tí-*, an abstract *-ti-* formation to the PIE word for 'eight', **h*$_x$*ok̂toh*$_x$.[72] According to Rau, **h*$_x$*ok̂th*$_x$*-tí-* reduced to **h*$_x$*ok̂h*$_x$*tí-* in PIE or at a very early stage of Proto-Indo-Iranian via "dissimilation... in order to break up the difficult cluster **k̂th*$_x$*t-* which ensued."[73] As Rau correctly recognizes, this deletion must have been very early because PIE **k̂* in **h*$_x$*ok̂(t)h*$_x$*-tí-* becomes Skt. *ś*, not ***ṣ*, which is the expected development of PIE **k̂* in the cluster **-k̂t-* (cf. Skt. *aṣṭāu* 'eight' < **h*$_x$*ok̂toh*$_x$). Of course, what made this cluster "difficult" was not the close proximity of two dental stops (cf. PIE **wit.sto-* 'known', PIE **mat.h*$_2$*tó-* 'torn off' > Skt. *mathitá-* 'shaken', etc.), but rather the violation of the SSP in the coda of the first syllable, rendering the violating conso-

[71]It is beyond the scope of this dissertation to perform a thorough analysis of Indo-European schwa primum and secundum; nevertheless, this tentative theory may be put forth. Note that in order to view these two rules of epenthesis as a unitary phonological process, we must envision two 'rounds' of schwa primum (laryngeal 'vocalization') to have occurred within the IE languages. The first occurred in PIE and was inherited by all IE languages: **ph*$_2$*trés* → **pəh*$_2$*trés* > Arm. *hawr*, OIr. *athar*; **dh*$_1$*só-* → **dəh*$_1$*só-* > Gk. *t*^{h}*eós*, PAnat. **daso-*. The second occurred independently in particular IE languages. This is essentially why some IE languages continue laryngeals as vowels in certain environments (PIE **dhugh*$_2$*ter-* > PGk. **dhugh*$_2$*əter-* > Gk. *t*^{h}*ugátēr*; PIE *h*$_2$*stér-* > PGk. *h*$_2$*əstér-* > Gk. *astér-*) while others do not (PIE **dhugh*$_2$*ter-* > Lith. *duktẽ*; PIE *h*$_2$*str̥-* > Skt. *str̥bhis*).

[72]See Rau 2009 for discussion, with references.

[73]Under the present analysis, Gk. *ogdo(w)os* 'eighth' could not derive directly from PIE **h*$_x$*ok̂th*$_x$*u̯os*. See Beekes 2010:1044 for further discussion.

nant unsyllabifiable: *[h$_x$ok̂]$_\sigma$t[h$_x$tí-]$_\sigma$ > *[h$_x$ok̂]$_\sigma$[h$_x$tí-]$_\sigma$.[74] Thus, we see that there was nothing inherently "bad" about the Indo-European laryngeals, phonemically or phonetically, that would result in their loss in the position PH.CC. Rather, there was something inherently "bad" about the syllable structure – a violation of the SSP in the coda. I would contend that the reason there are more examples of PH.CC > P.CC than, say, PP.CC > P.CC is that there is a greater number of roots of the shape *-PH than of the shape *-PP.[75]

---

[74]Another possible example of word-medial consonant deletion driven by a MAXST violation may be found in PIE **tek̂slah*$_2$ 'weapon' (IEW 1058), which is continued by ON *þexla* 'mattock', OHG *dehsa(la)* 'hatchet', Lat. *tēlum* 'spear, missile' (for an alternative etymology of Latin *tēlum*, see de Vaan 2008:609), RCS *tesla* 'axe'; OIr. *tál* 'axe', Ogam *TALA-GNI*. Joseph 1982:43, followed by Olsen 1988:16, has suggested that **tek̂slah*$_2$ is a **-tlo-* derivative to **tetk̂-* 'fashion', with the original meaning 'instrument of fashioning, tool, axe'. One might expect **tetk̂tlo-* > **tetlo-*, with **k* deletion driven by a MAXST violation (cf. '80') and **t* deletion driven by an OCP violation, just as in the *métron* rule (see below). Since this is not the case, I *very* tentatively suggest the following derivation, which would produce the reconstructable form **tek̂slah*$_2$:

**tetk̂tlah*$_2$	→	**tetsk̂tlah*$_2$	**s*-Epenthesis
**tetsk̂tlah*$_2$	→	**tek̂sttlah*$_2$	Metathesis
**tek̂sttlah*$_2$	→	**tek̂slah*$_2$	The *métron* rule (C deletion driven by OCP violation).
**tek̂slah*$_2$	→	*[tek̂][slah$_2$]	Syllabification

Note that while the MAXST violation does not lead to expected deletion of the SSP violating **k̂* (which in turn would have led to an output **tettlo-* thereby violating the OCP), perhaps its violation prohibits the input form **tetk̂tlo-* from being realized as such in the output.

[75]The LIV lists only two verbal roots of the shape *-PP, both ending in thorn clusters: ***h*$_2$*ad*^{h}*g*^{h}*-*** 'push' (LIV 255; Gk. *ák*^{h}*t*^{h}*omai* 'am oppressed', Hitt. *ḫatki* 'closes') and ***tetk̂-*** 'fashion, create' (IEW 1058-9; LIV 638-9. Gk. *téktōn* 'craftsman', Ved. *tákṣaṇ-* 'carpenter', Av. *tašan-* 'creator' < PIE **tétk̂-on-*). It lists seven of the shape tectal + **s*: ***dek̂s-*** 'be brave, to suit' (IEW 189; LIV 112. Ved. *dakṣáyanti* 'make suitable', Gk. *deksiterós*, Lat. *dexter* 'on the right', etc.), ***dek*$^{(w)}$*s-*** 'indicate' (IEW 189; LIV 112. Only attested in IIr.; Av. *daxšta-* 'feature'), ***h*$_2$*leks-*** 'protect' (IEW 32; LIV 278. Gk. *aléksō* 'ward off', Ved. *rákṣati*), ***h*$_2$*u̯eks-*** 'grow' (IEW 84-5. LIV 288-9. Eng. *wax*, Gk. *aúksomai* 'I grow', etc.), ***i̯ek*$^{(w)}$*s-*** 'appear' (IEW 502; LIV 312. Only attested in IIr.; cf. Ved. **prá yakṣanta* 'they distinguish themselves', Yagh. *yaxš-* 'appear'), ***mi̯ek̂s-*** 'set oneself' (LIV 445. Only attested in IIr.; Ved. *mimyákṣa* 'belongs to') and ***(h*$_1$*)rek̂s-*** 'harm' (LIV 505; cf. IEW 864. Gk. *erék*^{h}*t*^{h}*ō* < **erekst*^{h}*ō*, Ved. *rákṣas-* 'injury', Av. *rašah-* 'damage'). The nominal root ***(h*$_2$*)aks-*** 'axis, axle' (IEW 5; NIL 259-62), seen in Lat. *axis*, Ved. *ákṣa-* and Gmc. **ahslō* 'axle' (OE *eaxel*, OSax *ahsla*, etc.), also belongs here. Crucially, I have found no evidence that any of the derivatives formed to these roots violated the MAXST.

3.6 The *métron* rule.

If the above hypothesis successfully accounts for the data attested in IE, then the *métron* rule may be explained in a fashion more typologically natural than previously (Mayrhofer 1986:111; Hill 2003:23ff.): namely, as the result of a violation of the SSP in a medial coda. Saussure's and Schindler's syllabification of **médtrom* as **médt.rom* makes little sense typologically[76] and runs counter to what we presently know about PIE syllabification. Not only does PIE allow the sequence **tr* as an onset word-initially (**tréyes*) and word-medially (**d^{h}uk.trés*), it prefers onset maximization of complex consonant sequences to a simple consonant in the coda, realized by two high-ranking constraints: *ONSET 'Syllables must have onsets.' and ALIGNNUC 'Align the right edge of a syllable nucleus with the right edge of a syllable.'

As we have seen in section 1.3.1, all geminates were strictly banned across morpheme boundaries in PIE, and an illegal sequence *VTTV was fixed by *s*-insertion. This *s*-insertion (VTTV > VTsTV) was possible since one violation of the SSP was allowed in the onset: **u̯id-tó-* > **u̯it.stó-*. However, if an **s* were epenthesized into a sequence of the shape VTTRV then there would have been an SSP violation in the coda, which was strictly banned: **méd.trom* ↛ ***méts.trom*. Resyllabification of ***méts.trom* to ***mét.strom* would not have been allowed because of the MAXST, CCVCC: **méd.trom* ↛ ***mét.strom*.[77] However, the OCP constraint must always be obeyed in PIE across morpheme boundaries, which results in the deletion of the final root consonant: **méd.trom* → **mé.trom* (**mét.rom*?).[78] Thus we see that the strict ban

---

[76]See Jasanoff 2002:291 for a similar criticism of Mayrhofer's treatment of the Skt. imperatives *bod^{h}i* 'heed' and *yód^{h}i* 'fight'.

[77]A glaring exception is PIE **u̯oi̯tsth$_2$a* 'you know' if we are to view this word as having a quintipartite consonant cluster (and not **u̯oi̯tst^{h}a* or the like). I am indebted to Jessica DeLisi for pointing this out to me.

[78]If the /d/ of the root was lost, we would expect **mé.trom*; if the /t/ of the suffix was lost, we would expect **mét.rom* (see section 5.2 for discussion).

of an SSP violation in PIE codas is the driving force behind the *métron* rule. Consonant deletion of the root dental occurs to avoid the violation of two constraints: the OCP and the MAXST.[79]

(56) PIE **méd-trom* → [mé]$_\sigma$[trom]$_\sigma$

/méd-trom/		OCP	MAXST	MAX-T	DEP-[s]
a.	[mét]$_\sigma$[trom]$_\sigma$	*!			
b.	[méts]$_\sigma$[trom]$_\sigma$		*!		*
c.	[mét]$_\sigma$[strom]$_\sigma$		*!		*
d. ☞	[mé]$_\sigma$[trom]$_\sigma$			*	
e.	[més]$_\sigma$[trom]$_\sigma$			*	*!

Candidate (e), [més]$_\sigma$[trom]$_\sigma$, is not the most optimal because it violates two constraints (MAX-T & DEP-[s]), whereas the winning candidate (d) only violates one.

Note that we must explicitly view *TT > *TsT as a rule of *s*-insertion and not as a rule of affrication of the first dental. This is because affrication would lead to a different syllabification of **u̯id-tó-* (**u̯it͡s.tó-*), producing the same coda as would be found if the metron rule did not occur (**mét͡s.tro-*). A similar process of /s/ epenthesis may be demonstrated to occur in the Limburg dialect of Dutch, where /s/ is inserted and subsequently palatalized to [ʃ] before the diminutive suffix *-kə* if the root ends in a velar consonant: *bok* 'book' → *bækʃkə* 'little book'; *ɛk* 'corner' → *ɛkʃkə* 'little corner'.[80]

[79] See section 1.3.1 for the motivation behind the constraint ranking, OCP ≫ MAX-T ≫ DEP-[s].

[80] Hinskens 1996:137-8. For perceptual motivation of *s*-insertion after a stop, see Côté 2000:49. The insertion of /s/ is not a process restricted to the Double Dental Rule. For example, **s* appears in the enigmatic *s-mobile*: **pek̂-* ~ **spek̂-* 'see' (See Southern 1999 and Keydana, forthcoming a). There were alternations within PIE of the shape *-men* ~ *-smen* (cf. Attic *parádeigma* < **dei̯k̂mn̥* vs. Epidaurian *pardeik^{h}matōn* < **dei̯k̂smn̥* [Stüber 1998:52]); see Brugmann 1897-1916 II/1:242-3 for further

3.7 Exceptions to the MAXST?

As formulated, the MAXST explicitly banned PIE medial coda sequences where there was an SSP violation. Therefore if any such sequence is attested or reconstructable in an IE language, it cannot derive directly from PIE. For example, Lat. *extrā* 'outside' clearly violates the PIE MAXST, but we know the medial cluster /kstr/ in this word to be secondary, having arisen through a prehistorical process of syncope[81] from the sequence **eksterād*, with the /k/ being reintroduced into the word by analogy with the simplex preposition, *ex*.[82] Also a derivative of the PIE adposition **eḱs/*eǵʰz* (IEW 292-3), Gk. *ekʰtʰros* 'outsider' at first glance appears to show a MAXST violation (< **eḱstros*). However, this form is not likely to derive from PIE, and according to Beekes (2010:488-9), it may simply be a *-ro-* derivative of *ékʰtʰos* 'outside', a form that may be reconstructed as **éḱstos*, with a perfectly licit syllable structure in PIE under the rules of PIE syllabification set forth above: *$[\acute{e}\hat{k}]_\sigma[\text{stos}]_\sigma$.

Other apparent MAXST violations need not be so. For instance, Lat. *lūstrum* '(ceremony of) purification' has long been reconstructed as **l(e/o)u̯k-s-tro-* 'illumination' (vel sim.),[83] a **-s-tro-* derivative[84] to **leu̯k-* 'light'; cf. Lat. *lūx*. However, de Vaan (2008:354-5) objects to this etymology, arguing that "there is no good evidence for 'enlightening' in the meaning of *lūstrum*". Instead he proposes two alternate ety-

discussion. And lastly, if Merlingen (1957:51) is correct, thorn clusters also illustrate a process of **s*-epenthesis: **h₂artḱos → *h₂artsḱos* 'bear' (Schindler 1977b:32-3); **dʰǵʰōm → *dʰz(ʰ)ǵʰōm* 'earth' (Melchert 2003).

[81] See Nishimura 2008 for an excellent discussion of syncope in the Italic languages.

[82] Meiser 1998:152. Likewise, the English word *extra* /ekstrə/ is an explicit violation of the PIE MAXST. Though the maximum medial syllable in English monomorphemic words is also of the shape CCVCC and SSP violations in medial codas are rare (cf. Hammond 1999:84), I find it highly unlikely for it in any way to be a (modified) continuation of the original PIE MAXST.

[83] Walde & Hofmann 1982:839; IEW 688.

[84] As for the origins of the suffix *-stro-* in Latin, Leumann (1977:313) deems it "nicht erklärt", though it certainly does not derive from PIE.

mologies for this word: 1) that it may be connected with _lavō_ 'wash',[85] though there are no clear passages in Latin with the sense of 'washing' and 2) following Serbat (1975:312), _lūstrum_ is originally from *luh$_x$- 'set free'[86] + -_stro_-, as found in Gk. _lúō_ 'loosen' and Lat. _luō_ 'atone for; liberate', which de Vaan ultimately chooses as his preferred etymology. Note, however, that the assumption of the PIE MAXST does not necessarily rule out the reconstruction of _lūstrum_ as *l(e/o)u̯k-s-tro-; rather, it requires that at the time *l(e/o)u̯k-s-tro- was formed, the PIE MAXST no longer synchronically functioned as the maximum syllable template at that point within the prehistory of Latin.[87]

There are a number of instances of a sequence PiCC attested within Sanskrit, which at first glance suggest an inherited sequence *-PH]$_\sigma$[CC-. For example, to the root *k̂neth$_2$- 'pierce' (LIV 337) one finds Ved. _śnathiṣṭam_ 'pierce (2nd du. aor.)' (virtual *k̂neth$_2$st-); to k̂reth$_2$- 'slacken' (IEW 620; LIV 338), _śranthiṣyati_ 'will slacken', _śra(n)thitvā_ 'having slackened' (virtual *k̂reth$_2$-CC-); to *ku̯ath$_2$- 'froth, foam' (IEW 627-8; LIV 374), _kvathiṣyati_ 'will boil' (virtual *ku̯ath$_2$si̯e/o-); to k$^{(w)}$Reph$_x$- 'wail' (IEW 569; LIV 370), _akrapiṣṭa_ 'wailed' (virtual *ek$^{(w)}$Reph$_x$sto); and lastly to *g^{h}reb^{h}h$_2$- 'grab', we find _agrahīṣṭa_ 'grabbed (3 sg. aor M/P)' (virtual *eg^{h}reb^{h}h$_2$sto). However none of these examples argues decisively in favor of a reconstructed PIE sequence *PHCC, since they all may be instances where *h$_x$ (or just plain _i_) has been restored by analogy with other, phonologically regular _seṭ_ forms in the paradigm. Thus beside _śnathiṣyati_ we find _śnathitá_-; beside _agrabhiṣṭa_ the forms _gr̥bhītá_- '(having been) seized' and _ágrabhīt_ 'seized (3 sg. aor.)'; and certain forms, such as the form

[85] PIE *leu̯h$_3$-. IEW 692; LIV 418.

[86] IEW 681-2; LIV 417.

[87] Olsen (1988:17) provides another potential violation of the MAXST in her reconstruction of Umbr. **perkslum, pesklu** 'prayer' as *perk̂s-tlo-. However, Untermann (2000:540) derives these forms from *perk̂-sk̂-elo-, formed in the same fashion as **tiçel** '(sakrale) Deklaration, nuncupatio' < *dik̂-elo-, which would make this exception not probative.

akrapiṣṭa, might even be secondarily *seṭ* (LIV 370).

In short, I have been unable to find any reconstructable exceptions to the MAXST in PIE, and any instance of what looks like *POCC may easily be viewed as secondary.[88]

3.8 Conclusions.

In this chapter, we have seen that Schmidt's CHCC > CCC rule is more precisely formulated as PH.CC > P.CC and is driven by a strict ban of an SSP violation in medial codas, just as the *métron* rule and the loss of /t/ in PIE *h_xoḱth$_x$tí-* > Skt. *aśītí-* 'eighty'. This ban is part of the MAXST constraint reconstructed for PIE, which was deduced through phonotactic analysis of edge clusters (the DECOMPOSITION THEOREM) and the assumption of extrasyllabic consonants at word's edge.

Though it may seem strange at first to posit extrasyllabic consonants for PIE, common sense reminds us that while the consonant clusters in the words *u̯ḗḱst* 'he carried' and *$h_2$ster-* 'star' are reconstructable for PIE, there is no such word as *u̯eḱst.$h_2$ster*. In fact, this word looks decisively *un*-Indo-European. With the assump-

[88]The following roots of the shape *-PH (all taken from the LIV) have been examined in this study: ***b^{h}edh$_2$-** 'dig' (IEW 113-4; LIV 66), ***b^{h}leu̯dh$_2$-** 'dissolve' (IEW 159; LIV 90), ***g(^{w})renth$_2$-** 'knot, tie' (IEW 386; LIV 191), ***ghrebh$_2$-** 'grab' (IEW 455; LIV 201), ***h$_2$aḱh$_3$-** 'eat (up)' (IEW 18; LIV 261), ***h$_2$u̯edh$_x$-** 'speak' (IEW 76-7, LIV 286), ***h$_3$u̯ath$_2$-** 'wound' (IEW 1108; LIV 307), ***i̯eu̯ĝh$_x$-** 'become restless' (IEW 512, LIV 315), ***ḱneth$_2$-** 'pierce' (LIV 337), ***ḱreth$_2$-** 'slacken' (IEW 620; LIV 338), ***keu̯bh$_2$-** (IEW 590; LIV 357-8), ***k(w)Reph$_x$-** 'wail' (IEW 569; LIV 370), ***Kreph$_x$-** 'crack' (IEW 569; LIV 370), ***ku̯ath$_2$-** 'froth, foam' (IEW 627-8; LIV 374), ***k^{w}erpH-** 'turn' (IEW 631; LIV 392-3), ***lembh$_x$-** 'droop, sag' (IEW 656-7; LIV 411), ***meikh$_2$-** 'flash, shine' (IEW 712-3; LIV 429), ***mei̯th$_2$-** 'change (out), remove' (IEW 715; LIV 430), ***menth$_2$-** 'stir around' (IEW 732; LIV 438-9), ***meth$_2$-** 'rip off' (IEW 732, LIV 442-3), ***peth$_1$-** 'fall' (IEW 825-6; LIV 477-8), ***peth$_2$-** 'spread out' (IEW 824-5; LIV 478-9), ***peth$_2$-** 'fly' (IEW 825-6; LIV 479), ***pleth$_2$-** 'spread out' (IEW 833; LIV 486-7), ***preu̯th$_2$-** 'snort, foam' (IEW 810; LIV 494), ***rei̯k^{w}h$_2$-** 'scratch' (IEW 858; LIV 504), ***reu̯dh$_x$-** 'cry, wail' (IEW 867, LIV 508), ***sekh$_x$-** 'separate; cut; distinguish' (IEW 895-6; LIV 524), ***skeb^{h}h$_x$-** 'support' (IEW 916; LIV 549), ***(s)kedh$_2$-** 'split apart, spread out' (IEW 918-9; LIV 550), ***stemb^{h}h$_x$-** 'support, fight against' (IEW 1012-3; LIV 595), ***TerKh$_2$-** 'let, allow' (IEW 258: LIV 635), ***u̯ed^{h}h$_1$-** 'push' (IEW 1115: LIV 660), ***u̯eth$_2$-** (LIV 694), ***u̯i̯eth$_2$-** 'waver, roll' (IEW 1178; LIV 696).

tion of syllabically driven phonological rules, we may now see the PIE MAXST as the driving force in at least four major phonological processes in PIE: PH.CC > P.CC, the *métron* rule, schwa primum and schwa secundum. As we will see in the next chapter, the MAXST may also be utilized in the explanation of another phonological process in PIE: Sievers' Law.

As an aside, though this analysis assumes a phonological framework that employs extrasyllabic segments, it is conceivable that another framework would describe the phonological phenomena equally as well, or perhaps even better. For instance, Keydana (forthcoming a) proposes that PIE fricatives (*s, *h_x) were able to form semi-syllables, which were licensed in the onset of phonological words and restricted to one C-slot. Of course, if Keydana's framework (or anyone else's) proves to be more explanatory, this dissertation's analysis should be modified accordingly. The key point to our discussion is that PIE medial consonant clusters may be accurately predicted by the assumption of a special status of certain consonants at word's edge.

CHAPTER 4

Motivating Sievers' Law.

4.1 Introduction and Overview.

In the previous two chapters we have established a maximum syllable template (MAXST) for Proto-Indo-European, through the phonotactic analysis of PIE edge consonant clusters and the assumption of extrasyllabic consonants at word's edge. The MAXST makes explicit predictions about the possible shapes of PIE words. In this chapter, we will apply the MAXST to a much pored-over problem within Indo-European Linguistics and theoretical phonology: Sievers' Law (SL). We will first revisit the evidence and original formulation of SL, demonstrating it to have most likely been a postlexical phonological process in PIE, continued in varying degrees by a number of the IE daughter languages. We will then look in detail at Schindler's 1977 analysis of SL, which has been the most widely held view of SL since its publication. From what we now know of the PIE MAXST, it will become apparent that his analysis must be reconsidered, which will lead us to motivate SL through the avoidance of superheavy syllables at the postlexical level. Lastly, we will apply the findings from the following analysis of SL to three other instances of 'exceptional' syllabification in PIE, refining our conception of PIE syllabification in further detail.[1] The reader should bear in mind that the goal of this chapter is by no means to do a complete reanalysis of all the data that is relevant to the problem of Sievers' Law; rather, the goal is to apply

[1]The following chapter is an extended version of Byrd, forthcoming b.

the findings of the previous two chapters to our current knowledge of the problem at hand. For this reason, discussion of (the voluminous) past scholarship will be kept to a bare minimum.

4.2 Overview of Sievers' Law.

Sievers' Law has been one of the most perennially discussed phonological problems in Indo-European linguistics and the focus of numerous treatments in the generative phonological literature, particularly with regard to Germanic and Sanskrit.[2] For a comprehensive review of the Indo-European scholarship I refer the reader to Edgerton 1934, Seebold 1972:25ff., Horowitz 1974:11ff., Collinge 1985:159ff., Mayrhofer 1986:164-7, Meier-Brügger 2003:90-1, Sihler 2006 and Fortson 2010:71-2. While this chapter will not exhaustively cover the more than 100 years of scholarship devoted to the topic in question, it is necessary to provide the reader with a basic overview of the problem.

The original formulation of the law appeared in Sievers 1878, in a discussion of the Goth. *ja*-stem masculine nouns. In this particular class, some genitive singulars are of the shape *-jis*, while others are of the shape *-eis* (*-īs*). Sievers argued that this distinction resulted directly from the length of the preceding syllable. The genitive singular of *hairdeis* 'shepherd' ended in *-eis* (< **-ijis*) because the underlying suffix *-jis* was preceded by two consonants (*rd*), whereas the genitive singular of *harjis* 'army' remained in *-jis* because the suffix was only preceded by one (*r*). Similar phonological alternations may be found in the verbal system as well: classic examples include Goth. *waurkeiþ* 'works' vs. *satjiþ* 'sets', which are both underlyingly suffixed by *-j-*.

[2]See Kiparsky 1998 and Calabrese 1994, 1999 for discussion and references.

Sievers recognized similar alternations of glides and high vowels in Sanskrit,[3] which were restricted to the Rig Veda (Sihler 2006:5). To explain contrasting pairs such as *ajuryá-* 'unaging' / *asūria-* 'sunless' and *ávya-* 'woolen; made from sheep' / *mártia-* 'mortal' Sievers proposed the following rule: "If, in Indic, unaccented (without svarita) *i* or *u* occurs before a vowel ... then ... this segment is realized as a consonant after a light syllable and as a vowel after a heavy syllable."[4] For Sievers, a heavy syllable consisted of a short vowel plus two consonants (VCC) or a long vowel plus one consonant (V̄C); a light syllable consists of a short vowel plus one consonant (VC). Sievers conceived of this phonological process as the alternation of a glide and corresponding high vowel depending the shape of the preceding syllable.

(57) SIEVERS' LAW (SIEVERS 1878).
[+sonorant, -syllabic, +high] → [+syllabic] / VXC __ V (X = V, C)
If a glide is preceded by a heavy syllable, it is realized as [+syllabic].

Since Sievers' influential and now canonical observations, scholars have steadily proposed a number of parallels from within Indo-European that bolster his observation that this alternation of high vowels and glides may go back to PIE times. Furthermore, many scholars have extended SL to apply to the entire class of PIE resonants (**m*, **n*, **r*, **l*, **i̯*, **u̯*), such as Osthoff & Brugmann 1879:14-6[1], Edgerton 1934 and Schindler 1977a. This reformulation is given in (2) below.

(58) SIEVERS' LAW (RESYLLABIFICATION).
[+sonorant, -syllabic] → [+syllabic] / VXC __ V (X = V, C)

[3] Sievers 1878:129.

[4] Translation by Collinge (1985:159): "Im Indischen unbetontes (nicht svaritiertes) *i* oder *u* vor einem vocal ist consonant nach kurzer, vocal nach langer silbe ohne rücksicht auf die sonstige accentlage des wortes."

If a non-syllabic sonorant is preceded by a heavy syllable, it is realized as [+syllabic].

While Sievers had originally envisioned this phonological process as being an alternation of high vowel and glide depending on the weight of the preceding syllable, the following analysis will assume SL to be a rule of vowel epenthesis, such that _*terti̯o_ → _*tertii̯o_ and not _*tertio-_,[5] since, as has long been recognized, there appears to have been a strong preference in PIE for syllables to contain at least one consonant in the onset.[6]

(59) SIEVERS' LAW (EPENTHESIS).

ø → R̥$_1$ / VXC __ R$_1$V (X = V, C)

If a prevocalic non-syllabic sonorant follows a heavy syllable, epenthesize a corresponding syllabic resonant before the resonant in question.

In the following analysis, I will make use of three hypothetical 'roots': _*tert-_, _*tēt-_ and _*tets-_,[7] where _t_ = any consonant, _e_ = any short vowel, _ē_ = any long vowel, _r_ = any sonorant and _s_ = any consonant of equal or higher sonority than the preceding _t_.[8] In the derivations, _*-i̯o-_ will represent any glide- or sonorant-initial suffix that may potentially participate in SL, such that _*-i̯o-_ may stand for _*-i̯ah_$_2$_-_, _*-u̯o-_, _*-ri-_, etc. Though the present analysis assumes SL to have operated upon all resonants for simplicity of presentation, the core arguments of this paper by no means rest on this assumption. Should one prefer to restrict SL to the PIE glides (_*i̯_, _*u̯_) or even to

---

[5]Cf. Seebold 1972:29, Kobayashi 2001:93 and Weiss 2009:39.

[6]Cf. Mayrhofer 1986:123-4[108].

[7]Should one prefer attested roots, one may replace _*tert-_ with _*derk̂-_ 'see' (Gk. _dérkomai_, etc.), _*tēt-_ with _*h_$_1$_ēd-_ (Hitt. _ēdmi_), lengthened grade of _*h_$_1$_ed-_ 'eat' and _*tets-_ with _*h_$_2$_u̯eks-_ 'grow' (Av. _uxšiieitī_, Gk. _aúksomai_).

[8]Such that an SSP violation occurs if syllabified in the same syllable, through sonority rise (_mats.ya-_) or through sonority plateau ($dik^ht^h.ios$). See sections 1.4 and 3.3.1 for further discussion.

just **i̯*, (s)he may easily do so through the assumption of additional markedness (or faithfulness) constraints in this paper's analysis, as given in footnote 48, below.

4.2.1 Evidence in the daughter languages.

Languages that provide evidence for SL may be separated into two types: those where some semblance of the law is still productive and those where it is moribund and has been lexicalized. The former is only true in the oldest attested Germanic and Indic. In the Rig Veda we find that SL is most regularly attested in formations with the suffix *-ya-*. After a heavy syllable, there are 1552 instances of *-i(y)a-* but only 91 instances of *-ya-*, a 17:1 ratio; after a light syllable there are 462 instances of instances of *-i(y)a-* and 1747 instances of of *-ya-*, a 5:19 ratio.[9] Other suffixes behave much less consistently. For instance, the dative/ablative case ending *-bhyas* occurs as expected *-bhiyas* only 38.5% of the time after a heavy syllable (Seebold 1972:35), and a resistance to Sievers' Law in the Rig Veda is also found in the dual ending *-bhyām* and the gerund *-tvā*. Moreover, while there is some indication that the 3rd pl. ending *-iré* (< **-r̥ré*) originated as a Sievers variant of **-ré* in prehistoric Sanskrit,[10] nowhere does **-tr̥ra-* occur as a variant of the instrumental suffix *-tra-* (< **-tro-*, **-tlo-*).[11] It thus appears that Sievers variants are more regularly attested for suffixes of the shape *-RV-* than among those of the shape *-CRV-* in Sanskrit. As we will see, this fact is significant.

In Germanic, most notably in Gothic, I follow Kiparsky (1998) in assuming that

---

[9] Seebold 1972: 31, citing Edgren 1885:78. Of course, not all instances of *-iya-* (especially after a light syllable) are to be attributed to SL; see Schindler 1977:58.

[10] According to Praust, the original distribution was *-iré* (< **-r̥ré*) after heavy syllable (*īdhiré*, *ūciré*, *vavāśire*) and *-ré* after light (*riricré*, *nunudré*, *ānajre*, *jagr̥bhré*). Praust also points to other possible examples of SL affecting non-glides, such as Ved. *índra-* (see Sihler 2006:98-100 for opposite view) and the pair Ved. *cyautná-* 'work, deed', Av. *šiiaoθna-*, which at times must be scanned as three syllables. See Praust 2000:429-30 for further discussion and references.

[11] Sihler 2006:7.

SL was driven by a preference for moraic trochees, or feet of the shape (LL) or (H).[12] This most easily explains the application of SL in both monosyllabic roots such as _hairdeis_ 'shepherd' (< *_hairdijis_ < *_haird_ + _jis_) and in disyllabic roots such as _ragineis_ 'counselor' (< *_raginijis_ < *_raginjis_). Kiparsky (ibid.) argues that Gothic _(hair)(dei)<s>_ 'shepherd' and _(ra.gi)(nei)<s>_ 'counselor' underwent vowel epenthesis (with subsequent glide deletion) in order to avoid a suboptimal footing *_herd(jis)_[13] and *_ra(gin)(jis)_. According to Kiparsky, the form _(har)(jis)_ does not epenthesize -_i_-, since this would have led to **_ha(rei)<s>_, resulting in the prosodic structure L(H).

Traces of SL are found elsewhere in Germanic, such as in Old English.[14] As in Gothic, heavy root syllables (e.g. *_wītjas_) require high vowel epenthesis before a suffix consisting of a jod + vowel (*_wītijas_, footed as _(wī)(ti.ja)<s>_), with subsequent vowel contraction (_wītes_ 'punishment (gen.sg.)'). Polysyllabic roots also show evidence for SL, such as _æðeles_ 'nobility (gen.sg.)', which derives from *_æðaljas_, via *_æðalijas_ (footed as _(a.ða)(li.ja)<s>_). As elsewhere in Germanic and Vedic, SL does not occur within light monosyllabic stems such as *_kunjas_ > *_kunnjas_ > _kunnes_ 'kin (gen.sg.)', proved by the gemination of _n_ in the root _kunn_ 'kin'.[15] Although SL only applies to _j_-initial suffixes in attested Germanic, it is possible that it operated on other resonants in Proto-Germanic, such as in the case of the first person plural preterit ending *-_mé_ > *-_m̥mé_ > *-_um_ (e.g., PIE *_widmé_ 'we know' > Goth. _witum_).[16]

Elsewhere in Indo-European SL has disappeared as a productive phonological process, vanishing without a trace in some, while leaving behind subtle clues of its exis-

[12] See Kiparsky 1998 for references.

[13] Since [herd]$_\sigma$ is a superheavy syllable (syllable of three or more morae), it has one too many morae to form a moraic trochee (Kiparsksy 1998) and so cannot be footed.

[14] See Kiparsky 1998 and Adamczyk 2001.

[15] Dahl 1938:74ff.

[16] Ringe 2006:116ff.

tence in others. In the former category, one may include Albanian, Armenian, Celtic[17] and Slavic. In the latter lies the rest of Indo-European in varying degrees, both in quantity and in certainty. In Iranian, according to Schindler 1977a:58, the only certain example of SL is Av. _huuō.guua-_ (_hauguu̯a-_), a vr̥ddhi formation to an unattested *_hu-gú-_, cognate with Ved. _su-gú-_ 'having fine cows'. In Tocharian, the suffix *_-iyë_ possibly originates as a Sievers variant of *_-yë_; cf. PToch. *_ñəkciyë_ 'divine', a derivative of *_ñəktë_ 'god'.[18] In Greek, it is conceivable that the distribution of _-an-_ and _-n-_ presents originated as Sievers' variants (cf. _lambánō_ vs. _kámnō_), though far from certain; more likely is the contrast between *_h_$_2$_álg(i)i̯os_ > Gk. _álgios_ 'more painful' and *_pedi̯ós_ > _pezós_ 'on foot' (Skt. _pádya-_, Lat. _peius_).[19] In Italic, SL _might_ explain certain distributions of roots in the third and fourth verbal conjugations, such as the difference between _capiō/capere_ and _sāgiō/sāgīre_, if both derive from simple *_-i̯e/o-_ presents (*_kap-i̯e-_ but *_sāg-(i)i̯e-_).[20] In Baltic, SL may have produced two separate suffixes from the feminine *_-i̯ah_$_2$ declension, _-ia_ (< *_-i̯a_) and _-ė_ (< _-ii̯ā_), though the distribution has become opaque (cf. _eilė_ 'row' besides _eilia_).[21] And lastly, in Anatolian, SL is possibly found at a morpheme boundary in forms such as _ardumeni_ 'we cut (with a saw)' (< *_arduweni_), which may be contrasted with _tarweni_ 'we speak', which lacks an epenthetic vowel.[22] Note that in all instances but Av. _huuō.guua-_, the target

[17]See Schrijver 1995.

[18]Ringe 1991. It has been proposed by Jasanoff (apud Nussbaum 1986:8) that the pair TA _śaru_, TB _śerwe_ 'hunter' (< PT *_k̂eruwe)_ provides additional evidence for SL in Tocharian, originating from *_k̂ēruwo-_ < *_k̂ērwo-_, a vr̥ddhi-derivative of PIE *_k̂eru̯os_ 'stag' (Lat. _cervus_). See, however, Pinault 2008 for an alternate view. I am indebted to Melanie Malzaln for the reference.

[19]See Peters 1980:127ff., Ruijgh 1987 (1996) and Rix 1992.

[20]Meiser 1998:90. According to Weiss 2009:40: "The Latin evidence for the operation of Sievers' Law is very scant due to the interference of later sound changes, especially anaptyxis between most consonants and a following _i̯_."

[21]Sommer 1914; Horowitz 1974:19.

[22]Melchert 1994:57-8.

sonorant in question is suffix-initial: *-RV- → *-R̥RV-.

4.2.2 Einzelsprachlich or Inherited?

Should we follow Kiparsky in understanding the process of SL in Germanic as being driven by a preference for moraic trochees, then we must accept that the function of SL was different in the other IE languages and in PIE itself, as there is no indication that SL was motivated by such a preference in these languages. For this reason, a number of scholars[23] have argued that the many different reflexes of SL in the attested IE languages prevent us from projecting SL back to the proto-language. Rather, they argue, we must view the several instantiations of "Sievers' Law" as separate phonological processes occurring individually in the daughter languages.

However, as Ringe 2006:120 correctly points out: "The reapplication of Sievers' Law is hard to understand if it was an ordered rule, fossilized within the phonology of the language but no longer operative on the postlexical phonetic level; but it makes sense if Sievers' Law was operating as a surface filter, applying to any derived input that met its structural description in much the same way as modern German obstruent devoicing." Thus, if SL was truly inherited from PIE, we should view it as a postlexical rule carried across the generations, rather than as a historical event. Postlexical rules are phonological processes that apply at the phrasal, or syntactic, level. They typically apply across the board with no regard to morphological boundaries, tend to be exceptionless and frequently produce allophonic variation.[24] Ringe's implicit (and my explicit) claim is that Sievers' Law was a postlexical phonological process in PIE, which was lost, lexicalized or continued in different guises in the IE daughter languages.

---

[23] Beginning with Kluge 1891:502 and followed most recently by Sihler 2006:188-91.

[24] See Kiparsky 1982.

To illustrate my claim, let us take the example of final devoicing as a postlexical rule in some hypothetical language, which we will call Proto-XYZ. This process occurs across-the-board, meaning that no phrase-final obstruent can be realized as voiced in an utterance. Throughout the hypothetical years, Proto-XYZ evolves into three daughter languages: X, Y and Z. Each language evolves independently as does the inherited post-lexical rule of final devoicing. In language X, final devoicing disappears altogether. Perhaps this is due to sociolinguistic factors, such that X is influenced areally by a speech community that lacks this particular phonological process. Perhaps the loss of this rule is driven by linguistic factors, such that the final sequence *-DV, where D = any voiced obstruent, undergoes apocope, reintroducing the phonemic status of voiced obstruents in absolute word-final position.[25] In language Y, only a handful of words show alternation of a stem ending in a final voiced obstruent (imagine the opposite of Eng. *wife* : *wives*). Here, too, the rule is lost entirely, with only a handful of traces in the lexicon. Lastly, in language Z, final devoicing persists, but its purpose is altered such that it is unrecognizable. To be on the exotic (and quasi-ridiculous) side, let us say that in language Z all syllable final voiceless stops create a high tone (*-VT$ > -V́$),[26] and this high tone alternates with allomorphs containing a voiced stop (-V́ ~ -VDV). Thus, we may say that final devoicing has continued as a (morpho)phonological process in language Z, but with a function entirely different from that of the proto-language.

It is in this fashion that Sievers' Law, which originally was a postlexical rule, evolved in the attested Indo-European languages. In some subgroups, such as Albanian, Armenian and Slavic, SL was lost without a trace. In others, such as Greek, Iranian and Italic, SL only persists through a handful of archaic forms. And lastly, in

[25] Such as Yiddish. See Albright 2008.

[26] For a thorough discussion of 'tonogenesis', see Matisoff 1973.

other subgroups, such as Indic and Germanic, SL persists to a greater or lesser extent as a synchronic process, which has been altered to suit the needs of the speakers of those languages. Our current task is to devise a postlexical rule that can conceivably evolve into each of the IE systems attested and to arrive at an independently motivated reason for the existence of Sievers' Law at the PIE postlexical level.

4.2.3 Schindler 1977.

To the best of my knowledge, Schindler (1977a), in his influential review of Seebold 1972, did not conceive of SL as a postlexical process in PIE. His paper arguably constitutes to date the most successful attempt in motivating the original conditions of SL in PIE; it has undoubtedly been the most widely held view in the literature since its publication.[27] In his review, Schindler points out two curious instances of SL not applying in Vedic. First, Schindler (1977a:61) argued that words of the shape **tetti̯o-* are not realized as ***tettii̯o-*, where *tt* = any sequence of two obstruents (regardless of sonority level). The best example is Ved. *matsya-* 'fish', which is never scanned as ***matsiya-* and must go at least as far back as Proto-Indo-Iranian, as is evidenced by its Avestan cognate *masya-* 'id.' Ruijgh (1996:354) also points to a similar treatment of **tetti̯o-* in Ionic Gk., where *diksós* 'double' derives from **dik*^{h}*t*^{h}*i̯ós* (cf. *dik*^{h}*t*^{h}*á* 'in two'). If SL had occurred, one would expect ***dik*^{h}*t*^{h}*ii̯ós*. Second, Schindler (1977:60-1) pointed out that the absolutive **-tu̯V-* (Sanskrit *-tvā́*, *-tvī́*, *-tvā́ya*), never shows Sievers variants, with forms such as *gūdhvī́*, *yuktvā́*, *bhūtvā́*, *jagdhvā́ya*, etc., never being scanned as ***gūdhuvī́*, ***yuktuvā́*, ***bhūtuvā́*, ***jagdhuvā́ya*, etc.

To prevent forms like *matsya-* and the absolutives from participating in Sievers' Law, Schindler proposes the following syllabifications of our three hypothetical forms:

[27]Cf. Peters 1980:129ff., Mayrhofer 1986:164-8, Meiser 1998:89ff., Praust 2000, Meier-Brügger 2003:90-1, Weiss 2009:39-40 and Fortson 2010:72.

ter.ti̯o-*, **tē.ti̯o-* and **tett.i̯o-* (= my **tets.i̯o-*).[28] These syllabifications, he argues, were not chosen in an entirely ad hoc fashion, but rather confirmed Saussure's syllabifications of the double dental clusters (méd-t.ro-*),[29] a problem we addressed in section 3.6. Thus, Schindler explains SL to be the result of an avoidance of a complex onset C + R (or C + U̯) in a word-final syllable in PIE (so Meiser 1998:89-90).

Structurally Schindler's hypothesis has proved very attractive, since it allows us to collapse SL together with Lindeman's Law[30] (LL), thereby viewing long Lindeman variants such as *dii̯ḗu̯s* 'sky (god)' (← *di̯ḗu̯s*) as the result of SL in certain sandhi configurations.[31]

(60) LINDEMAN'S LAW (SCHINDLER 1977).

$[\text{+son, -syll}] \rightarrow [\text{+syll}] \,/\, \#[\text{+obst}]_0 \,\_\_\, [\text{+syll}][\text{-syll}]_0\#$

If a non-syllabic sonorant follows a single obstruent and precedes a vowel in a monosyllable, it is realized as [+syllabic].

Classic examples include **di̯ḗu̯s* 'sky god' (Gk. *Zéus*), which alternates with **dii̯ḗu̯s* (Skt. *diyā́us*), and **du̯oh*$_x$ 'two' (OE *twā*), which has a variant **duu̯oh*$_x$ (Lat. *duō*).

Schindler envisioned SL and LL to be identical processes, most notably in two regards. First, he claimed, both processes occurred in the final syllable of a word. While this goes without saying for LL (which applies only in monosyllables), Schindler noted that SL is frequently not attested in words such as *vāiśvānará-* 'pertaining to all men',

[28] Schindler 1977a:60[4]. Note that in this analysis, $[\text{tets}]_\sigma$ represents any sequence in which there is a SSP violation. Should fricatives (**s* and laryngeals) be more sonorous than stops as discussed in section 1.4, then this would imply that the sequences $*[\text{test}]_\sigma$ and $*[\text{teh}_x\text{t}]_\sigma$ were legal syllables in PIE. See footnote 65 below for how this may be relevant to our understanding of Sievers' Law.

[29] Mayrhofer 1986:111.

[30] Lindeman 1965.

[31] Cf. Meier-Brügger 2003:142: "On the one hand, we have *##... V̄#*dii̯ḗu̯s* and *##... V̄R#*dii̯ḗu̯s*, on the other *##... K#*di̯ḗu̯s...*"

which never scans in five syllables (i.e. **vāiśuvānará-). He explains this lack of application by restricting SL to the final syllable of a word. Second, once Schindler had identified the syllabifications *ter.ti̯o-, *tē.ti̯o- and *tett.i̯o- (= my *tets.i̯o-) he was able to define the targeted sequence as a syllable onset consisting of a consonant + non-syllabic sonorant ($TR-). His reformulation is given in (61).

(61) SIEVERS'/LINDEMAN'S LAW (SCHINDLER 1977:64).
$[+son, -syll] \rightarrow [+syll] / \$[-syll]_{1}\_\_ [+syll][-syll]_{0}\#$
In the final syllable of a word, if a non-syllabic sonorant is preceded by a consonant and followed by a vowel, it is realized as [+syllabic].

There are two problems with Schindler's ingenious analysis, however. First, as Collinge (1985:165) points out, we have no reason to believe that a sequence *C + R was disfavored in PIE onsets, as can be shown by a sizeable number of roots and words in PIE: *ti̯eg^w- 'withdraw', *tu̯erk- 'cut', *tréi̯es 'three', *pleh$_1$- 'fill', etc. Second, the syllabification of *tetsi̯o as *tets.i̯o is problematic, since there is no direct evidence for this type of syllabification attested in the IE languages. This renders Schindler's analysis completely circular. Of course, more seriously, if the independently-motivated MAXST holds true for PIE syllabification, Schindler's syllabification of *tetti̯o- (= my *tetsi̯o-) as *tett.i̯o- (= my *tets.i̯o-) was in fact impossible for a speaker of PIE, since an SSP violation was prohibited in a medial coda. Following the MAXST, we now contend that this sequence must have been syllabified as *tet.si̯o- and so we need to devise an alternate solution.

4.3 Motivating Sievers' Law: the Avoidance of Superheavy Syllables.

The MAXST, which dictates the largest possible syllable in PIE, predicts the following syllabifications of our hypothetical root shapes _*tert_, _*tēt_ and _*tets_ to have been possible:

(62) */tert/ → *[tert]$_{\sigma}$ (superheavy σ)
/tēt/ → *[tēt]$_{\sigma}$ (superheavy σ)
but */tets/ → *[tet]$_{\sigma}$s, (heavy σ, with an unsyllabified consonant)

The first two roots _*tert_ and _*tēt_ are both entirely syllabifiable and form a superheavy syllable, or a syllable consisting of more than two morae. The last root _*tets_ must be syllabified as [tet]$_{\sigma}$s, since the second obstruent in coda position violates the SSP.

Let us assume that the syllabifications of these roots were realized as such at the derivational stage where SL occurred, driven by the desire for PIE speakers to keep morphemes syllabically distinct from one another. If so, the suffixation of _*-i̯o-_, _*-u̯o-_, etc. would have resulted in the following syllabifications:

(63) _*tert-_ + _*-i̯o-_ → *[tert]$_{\sigma}$[i̯o]$_{\sigma}$
*tēt- + _*-i̯o-_ → *[tēt]$_{\sigma}$[i̯o]$_{\sigma}$
*tets- + _*-i̯o-_ → *[tet]$_{\sigma}$[si̯o]$_{\sigma}$

Should we assume this typologically common tendency in the phonological derivation, it now becomes clear what the motivation was for SL in PIE: the avoidance of a superheavy syllable. While *[tet]$_{\sigma}$[si̯o]$_{\sigma}$ does not contain a superheavy syllable at the postlexical level, *[tert]$_{\sigma}$[i̯o]$_{\sigma}$ and *[tēt]$_{\sigma}$[i̯o]$_{\sigma}$ do, which is what prompts the resyllabification and insertion of a syllabic sonorant (SL).

Unlike Schindler's explanation of SL above, the avoidance of superheavy syllables is extremely well founded typologically[32] and is seen elsewhere in PIE and the attested IE languages.[33] In PIE, we saw earlier that compensatory lengthening is blocked in medial position if it would result in a superheavy syllable: **ĝenh$_1$mn-* > PIE **ĝen.mnV-*, not ***ĝēn.mnV-* (Skt. *janmane*, Gk. *genná*).[34] Certain instances of Schwebeablaut, or the metathesis of a root sonorant from coda to onset position (**derk̂-* 'see' > **drek̂-*), suggest a dispreference for superheavy syllables as well.[35] Lastly, the avoidance of (super)heavy syllables may also have played a part in the analogical replacement of weak full-grade forms with the zero-grade in roots of the shape *TeR(T) (i.e., **ter(t)-* → **tr̥(t)-*): e.g., *$[\text{kom}]_\sigma[\text{h}_1\text{eit}]_\sigma$s 'fellow traveller' → *$[\text{kom}]_\sigma[\text{h}_1\text{it}]_\sigma$s > Lat. *comes, -itis* 'companion'.[36]

Later in IE there are additional, language-specific processes of superheavy syllable avoidance. The most widespread is Osthoff's Law, whereby a long vowel is shortened if it precedes a resonant + consonant: */V̄/ → [V] / __ RC.[37] For example, the well attested word for 'wind', PIE **h$_2$u̯eh$_1$n̥to-*, is realized as Post-PIE **u̯ēnto-* with loss of laryngeal, which is subsequently shortened to **u̯ento-* (Latin *ventus*, OIr. *fét* 'whistle', Goth. *winds*, etc.). Kobayashi has also argued for Brugmann's Law to be the result of the blocking of the sound change **o* > **ā* in IIr. closed syllables, once again to avoid a superheavy syllable.[38]

---

[32]Cf. Zec 1995:100ff. and see Sherer 1994:11 for references.

[33]Seebold 1972:132, Hoenigswald 1988:202-3, Kobayashi 2001:92ff. & 2004:26

[34]See section 3.3.3.

[35]For example, Skt. *dr̥ś-* 'see' takes the shape *darś-* in *darśati*, *dadarśa*, but *drak-* in *adrākṣīt* (not **adārkṣīt*) and *drakṣyáti* (not **darkṣyáti*). Cf. Anttila 1969:52ff..

[36]See Jasanoff 2003:43ff. and Vijūnas 2006:90ff. for discussion with references.

[37]See Collinge 1985:127-31, Kobayashi 2001:94 and Fortson 2010:70-1.

[38]Kobayashi 2001:94, 2004:26-7

Attributing SL to the avoidance of a superheavy syllable has been proposed elsewhere,[39] though with no explanation as to why there even existed a superheavy syllable in these sequences. If PIE speakers wanted to avoid superheavy sequences, then why was */tert + i̯o/ even syllabified as *[tert]$_\sigma$[i̯o]$_\sigma$ in the first place? If we follow Ringe in assuming that SL was a postlexical process in PIE, then the answer is straightforward: when syllabified at the stem (and word) level, *[tert]$_\sigma$[i̯o]$_\sigma$, *[tēt]$_\sigma$[i̯o]$_\sigma$, etc. were the most optimal forms. Once these forms were fed into the postlexical level, however, there was a violation of a constraint that blocks superheavy syllables, prompting the epenthesis of a sonorant, whose quality is copied from the adjacent non-syllabic resonant.

4.3.1 Framework Used in Analysis.

Our formal analysis of Sievers' Law will employ a Stratal Optimality Theory framework, which assumes multiple stratified constraint systems in the grammar.[40] These strata are typically threefold and are arranged in the following order: STEM, WORD and POSTLEXICAL (or phrasal). By reintroducing the phonological cycle into the grammar, Stratal OT provides a response to the ever-vexing problem of opacity in OT. Unlike classical OT, where there is one level of constraint rankings, constraints may have different rankings at each phonological level in Stratal OT.[41]

The first stratum in the grammar, the stem level, is where roots combine with cer-

[39] Hoenigswald 1988:202, Fullerton 1992:85-6, Neri 2003:32[69].

[40] See Kiparsky 2000 and more recently Bermúdez-Otero, forthcoming. Stratal OT is by-and-large an OT adaptation of the earlier, rule-based Lexical Phonology; for discussion with references, see Kaisse and Shaw 1985.

[41] Though this analysis adopts Stratal OT as its framework of choice, the reader should bear in mind it may be conducted using a ruled-based framework as well, though certain elements of the analysis, such as the The Emergence of the Unmarked (TETU) phenomenon inferred at the postlexical level, would be lost (for which see section 4.3.3).

tain affixes, which are usually derivational.[42] The output produced at this level is not a morphological word and therefore is not uttered *per se*. To give a concrete example of a PIE form derived by Stratal OT, let us examine the hypothetical root **tert-* once again, which is derived as an adjective with the suffix **-i̯o-*.

(64) STEM LEVEL: **tert-* + **-i̯o-* → *[tert]$_\sigma$[i̯o]$_\sigma$

This form *[tert]$_\sigma$[i̯o]$_\sigma$ was never uttered by a PIE speaker, as additional overt inflectional morphology was required.

The second stratum in the grammar, the word level, will not be addressed directly in our SL analysis, though it will become relevant later on in the next chapter. Here, inflectional endings will be added and perhaps some derivational ones (see section 5.2.1):

(65) WORD LEVEL: *[tert]$_\sigma$[i̯o]$_\sigma$ + *-s* (nom.sg.) → *[tert]$_\sigma$[i̯os]$_\sigma$

It is this form that is fed into the postlexical grammar.

The last stratum in the grammar, the postlexical level, is where rules occur across the board with no regard to morphological category. It is here that SL is hypothesized to have occurred.

(66) POSTLEXICAL LEVEL: *[tert]$_\sigma$[i̯os]$_\sigma$ → *[ter]$_\sigma$[ti]$_\sigma$[i̯os]$_\sigma$

Note that at each level there is a syllabification cycle. At the stem level, the coalescence of the root and suffix produces an initial syllabification, which favors keeping morphemes syllabically distinct, if possible. At the word level the nominative singular case ending is added, which must be syllabified and therefore is adjoined to the nearest syllable [i̯o]$_\sigma$. Lastly, at the postlexical level, we find syllabic repartition as the result of SL, which is driven by the avoidance of the superheavy syllable *[tert]$_\sigma$.

[42] See Bermúdez-Otero, forthcoming for discussion with examples.

4.3.2 The Stem Level.

Let us now turn to the formal analysis of Sievers' Law. Strata that are relevant here are the stem and postlexical levels. The constraints used at the stem level are given in (67) below.

(67) CONSTRAINTS USED AT STEM LEVEL.

a. MAXST: The syllable in question cannot violate the PIE maximum syllable template CCVCC, where the coda cannot violate the Sonority Sequencing Principle. Assign one $*$ for each violation.

b. PARSE: Syllabify all segments. Assign one $*$ for each segment not syllabified.

c. DEP-V(OWEL): Don't insert a vowel. Assign one $*$ for each instance of vowel epenthesis.

d. MAX-C(ONSONANT): Don't delete a consonant. Assign one $*$ for each consonant lost.

e. ALIGN: For every morpheme boundary, there must be a syllable boundary. Assign one $*$ for each violation.

f. *SUPERHEAVY (*SPRHVY): No syllable may consist of three or more morae. Assign one $*$ for each superheavy syllable.

To conduct this analysis, we must rank these constraints in the grammar, postulating their positions as precisely as possible and providing external evidence whenever we can. This ranking is given in (68) below, with justifications presented in a footnote.[43]

[43] MAXST $\gg$ all, as there is no sequence reconstructable for PIE (at any phonological level) that violates the maximum syllable template. MAXST $\gg$ DEP-V as well, since *$ph_2trés$ → *[pəh$_2$]$_\sigma$[trés]$_\sigma$. Next, PARSE $\gg$ all, since all segments that are not extrasyllabic must be syllabified in the derivation: *$d^hugh_2trés$ → *[d^{h}uk]$_\sigma$[trés]$_\sigma$ and not **[d^{h}ug]$_\sigma$h$_2$[trés]$_\sigma$. As we saw in section 2.2.1, DEP-V $\gg$ MAX-C, because *$d^hugh_2trés$ → *[d^{h}uk]$_\sigma$[trés]$_\sigma$ and not **[d^{h}ug]$_\sigma$[h$_2$ə]$_\sigma$[trés]$_\sigma$ (or the like).

(68) STEM LEVEL CONSTRAINT RANKING.

MAXST, PARSE ≫ DEP-V ≫ MAX-C ≫ ALIGN ≫ *SUPERHEAVY.

Let us now proceed to examine the predictions of these constraint rankings in this analysis of SL, beginning with the syllabification of */tert+i̯o/ at the stem level. For reasons of brevity, only */tert+i̯o/ will be discussed, since */tert+i̯o/ and */tēt+i̯o/ behave in an identical fashion.

(69) STEM */tert+i̯o/

/tert+i̯o/	MAXST	DEP-V	MAX-C	ALIGN	*SPRHVY
a. ☞ [tert]$_\sigma$[i̯o]$_\sigma$					*
b. [ter]$_\sigma$[ti̯o]$_\sigma$				*!	
c. [ter]$_\sigma$[i̯o]$_\sigma$			*!		
d. [ter]$_\sigma$[ti]$_\sigma$[i̯o]$_\sigma$		*!		*	

Here we see that syllable and morpheme boundary are kept identical, due to the constraint ranking ALIGN ≫ *SUPERHEAVY.

(70) STEM */tets+i̯o/

/tets+i̯o/	MAXST	PARSE	DEP-V	MAX-C	ALIGN	*SPRHVY
a. [tets]$_\sigma$[i̯o]$_\sigma$	*!					*
b. [tet]$_\sigma$s[i̯o]$_\sigma$		*!			*	
c. ☞ [tet]$_\sigma$[si̯o]$_\sigma$					*	
d. [tet]$_\sigma$[i̯o]$_\sigma$				*!		
e. [tet]$_\sigma$[si]$_\sigma$[i̯o]$_\sigma$			*!		*	

Next, we should postulate that MAX-C ≫ ALIGN, because *d^hugh_2- + *-ter* → *[d^{h}ug]$_\sigma$[$h_2$ter]$_\sigma$, not **[duk]$_\sigma$[ter]$_\sigma$ and *h_2*uksto* → *[$h_2$uk]$_\sigma$[sto]$_\sigma$, not **[$h_2$uk]$_\sigma$[to]$_\sigma$. Lastly, ALIGN ≫ *SUPERHEAVY, based on the assumptions of this analysis.

Because /s/ violates the MAXST in the root **tets*, it cannot be syllabified in the same syllable as /tet/. The high-ranking nature of the constraint PARSE requires that this /s/ be syllabified, and therefore it is syllabified with the following suffix.

(71) STEM */tr̥k+tu̯o/

/tr̥k+tu̯o/		MAXST	DEP-V	MAX-C	ALIGN	*SPRHVY
a.	[tr̥kt]$_\sigma$[u̯o]$_\sigma$	*!			*	*
b. ☞	[tr̥k]$_\sigma$[tu̯o]$_\sigma$					
c.	[tr̥k]$_\sigma$[u̯o]$_\sigma$			*!		
d.	[tr̥k]$_\sigma$[tu]$_\sigma$[u̯o]$_\sigma$		*!			

Here, in our hypothetical absolutive form, the candidate *[tr̥k]$_\sigma$[tu̯o]$_\sigma$ is chosen, which violates neither ALIGN nor *SUPERHEAVY.

4.3.3 The Postlexical Level.

Let us now proceed to the postlexical level, for which we must make two assumptions. First, since these are phonological processes at the phrasal, or syntactic, level, morpheme boundaries become irrelevant.[44] For this reason, ALIGN is no longer a relevant constraint in the analysis. Second, the constraints MAX-C and DEP-V must be re-ranked, as this is required for vowel epenthesis (SL) to occur and not deletion — we do not find **[ter]$_\sigma$[i̯os]$_\sigma$ from *[tert]$_\sigma$[i̯os]$_\sigma$, etc. Furthermore, two additional constraints become relevant at the postlexical level.

(72) ADDITIONAL CONSTRAINTS AT THE POSTLEXICAL LEVEL.

a. FAITH(σ): Do not alter the syllabification of the base form. Assign one

[44] Known as 'Bracket Erasure'; see Kiparsky 1982:11.

∗ for every instance in the output a segment is syllabified in a syllable different from that of the input.[45]

b. *COMPLEXONSET: Onsets may not consist of more than one consonant in the output. Assign one ∗ for each violation.

The first constraint, FAITH(σ), is required to ensure that that winning candidates of the input forms *[tet]$_\sigma$[si̯os]$_\sigma$ & *[tr̥k]$_\sigma$[tu̯o]$_\sigma$, which satisfy the constraint *COMPLEXONSET, do not win.[46] The latter constraint, *COMPLEXONSET, is crucial in the choice of **tertii̯os*, and not **terti̯os*, as the winning candidate. The interaction of these newly added constraints is given in (73).

(73) POSTLEXICAL CONSTRAINT RANKING.[47]
MAXST, PARSE, MAX-C ≫ *SUPERHEAVY ≫ FAITH(σ) ≫ *COMPLEXONSET ≫ DEP-V.

The crucial constraint ranking is *SUPERHEAVY ≫ FAITH(σ) ≫ *COMPLEXONSET — their interaction is what drives Sievers' Law at the postlexical level.

[45]Or more precisely: "If x$_i$ belongs to σ_i in the input, and x$_i$ has an output correspondent x$_o$, x$_o$ must belong to a syllable σ_o that corresponds to σ_i." I thank Kie Zuraw for her clarification of this matter.

[46]Further evidence for FAITH(σ) may also be found in the 'exceptional' syllabification of nasal-infixed presents (e.g. **i̯ungénti* 'they yoke', not ***iu̯n̥génti*) and certain accusatives in **-m(s)* (e.g. **méntim* 'mind (acc.sg.)', not **ménti̯m̥* (contra Keydana, forthcoming b). See section 5.2, below.)

[47]The postlexical constraint ranking will be justified as follows. First, MAXST, PARSE and MAX-C ≫ all, since they are never violated in outputs produced at the postlexical level. MAX-C ≫ *SUPERHEAVY, as we find *[i̯eu̯k]$_\sigma$[trom]$_\sigma$ 'yoke', not **[i̯eu̯]$_\sigma$[trom]$_\sigma$. More generally, it may be said that the constraint MAX-IO (MAX-C, MAX-V) ≫ *SUPERHEAVY, since an "Osthoff's Law"-like process did not occur in PIE: **tērtrom* $\not>$ ***tertrom.* *SUPERHEAVY ≫ FAITH(σ), because *[tert]$_\sigma$[i̯os]$_\sigma$ → *[ter]$_\sigma$[ti]$_\sigma$[i̯os]$_\sigma$, not **[tert]$_\sigma$[i̯os]$_\sigma$. FAITH(σ) ≫ *COMPLEXONSET, because *[tet]$_\sigma$[si̯os]$_\sigma$ does not undergo SL. MAX-C ≫ DEP-V, because SL is a process of vowel epenthesis, and not consonant deletion (*[tert]$_\sigma$[i̯os]$_\sigma$ → *[ter]$_\sigma$[ti]$_\sigma$[i̯os]$_\sigma$, and not **[ter]$_\sigma$[i̯os]$_\sigma$). *COMPLEXONSET ≫ DEP-V, because *[ter]$_\sigma$[ti]$_\sigma$[i̯os]$_\sigma$, and not **[ter]$_\sigma$[ti̯os]$_\sigma$, is the winning output when *[tert]$_\sigma$[i̯os]$_\sigma$ is processed at the postlexical level. Lastly, *SUPERHEAVY ≫ *COMPLEXONSET for reasons of transitivity, because *SUPERHEAVY ≫ FAITH(σ).

Turning now to the tableaux, we find that for our hypothetical inputs *[tet]$_{\sigma}$[si̯os]$_{\sigma}$ and *[tr̥k]$_{\sigma}$[tu̯o]$_{\sigma}$, there is no change in syllabification at the postlexical level.

(74) POSTLEXICAL *[tet]$_{\sigma}$[si̯os]$_{\sigma}$

/[tet]$_{\sigma}$[si̯os]$_{\sigma}$/			MAXST	MAX-C	*SPRHVY	FAITH(σ)	*COMPONS	DEP-V
a.		[tets]$_{\sigma}$[i̯os]$_{\sigma}$	*!		*	*		
b.	☞	[tet]$_{\sigma}$[si̯os]$_{\sigma}$					*	
c.		[tet]$_{\sigma}$[i̯os]$_{\sigma}$		*!				
d.		[tet]$_{\sigma}$[si]$_{\sigma}$[i̯os]$_{\sigma}$				*!		*

(75) POSTLEXICAL *[tr̥k]$_{\sigma}$[tu̯o]$_{\sigma}$

/[tr̥k]$_{\sigma}$[two]$_{\sigma}$/			MAXST	MAX-C	*SPRHVY	FAITH(σ)	*COMPONS	DEP-V
a.		[tr̥kt]$_{\sigma}$[u̯o]$_{\sigma}$	*!		*	*		
b.	☞	[tr̥k]$_{\sigma}$[tu̯o]$_{\sigma}$					*	
c.		[tr̥k]$_{\sigma}$[u̯o]$_{\sigma}$		*!				
d.		[tr̥k]$_{\sigma}$[tu]$_{\sigma}$[u̯o]$_{\sigma}$				*!		*

The input *[tert]$_{\sigma}$[i̯os]$_{\sigma}$ is correctly realized as *[ter]$_{\sigma}$[ti]$_{\sigma}$[i̯os]$_{\sigma}$, as is given in (76).

(76) POSTLEXICAL *[tert]$_{\sigma}$[i̯os]$_{\sigma}$

/[tert]$_{\sigma}$[i̯os]$_{\sigma}$/			MAXST	MAX-C	*SPRHVY	FAITH(σ)	*COMPONS	DEP-V
a.		[tert]$_{\sigma}$[i̯os]$_{\sigma}$			*!			
b.		[ter]$_{\sigma}$[ti̯os]$_{\sigma}$				*	*!	
c.		[ter]$_{\sigma}$[i̯os]$_{\sigma}$		*!				
d.	☞	[ter]$_{\sigma}$[ti]$_{\sigma}$[i̯os]$_{\sigma}$				*		*

Here we find that the output **[tert]$_\sigma$[i̯os]$_\sigma$ is not chosen because the now highly-ranked constraint *SUPERHEAVY is violated. The candidate **[ter]$_\sigma$[ti̯os]$_\sigma$, with simple resyllabification, is avoided because of a violation of *COMPLEXONSET, or the avoidance of onsets consisting of more than one consonant. The most optimal form is [ter]$_\sigma$[ti]$_\sigma$[i̯os]$_\sigma$, with vowel epenthesis (Sievers' Law); this candidate avoids both a superheavy syllable and a complex onset in the output.[48]

Thus we find that SL is motivated by the avoidance of a superheavy syllable (*SUPERHEAVY), coupled with the desire of the PIE speaker to avoid complex onsets (*COMPLEXONSET) at the postlexical level. We are now in a position to address Collinge's cogent objection to the idea that SL is driven by an avoidance of a sequence $_\sigma$[TR- (or $_\sigma$[TU̯-): "But it is not totally clear why, if the first consonant of the cluster has become a syllable-coda, the sequence $ty- is then any less acceptable than the word initial ##ty-".[49] Sievers' Law is not driven by the avoidance of a complex onset in medial position; rather, Sievers' Law is driven by the avoidance of a superheavy syllable at the postlexical level, with the most optimal candidate avoiding a complex onset. This is a classic example of "The Emergence of the Unmarked" (TETU), a key tenet of OT: "[A] preference for some universally unmarked structure, such as syllables with onsets, can emerge under the right circumstances even if the language as a whole permits the corresponding marked structure." (McCarthy 2008:24-5)

[48]Should one prefer to restrict SL to particular sequences of consonant + sonorant (such as *TU̯-), (s)he could assume a wide distribution of epenthetic constraints, such as DEP-[i], DEP-[u], DEP-[r̥], etc. that are ranked accordingly at the postlexical level. So, in order to rule out *tert-ro-s → *tertr̥ros* by SL, we would need to rank DEP-[r̥] above *COMPLEXONSET, making *[ter]$_\sigma$[tros]$_\sigma$ a candidate more optimal than [ter]$_\sigma$[tr̥]$_\sigma$[ros]$_\sigma$. An alternate way would be to assume two constraints, DEP-V/C_ R "don't epenthesize a vowel in a cluster of consonant + *m*, *n*, *r*, *l*" and DEP-V/C_U̯ "don't epenthesize a vowel in a cluster of consonant + glide" (cf. Zuraw 2007:297, following Fleischhacker 2005), with the relevant constraint ranking in PIE: DEP-V/C_ R ≫ *COMPLEXONSET ≫ DEP-V/C_U̯.

[49]Collinge 1985:165.

4.3.4 Overgeneration.

Upon careful inspection, we find that the solution presented above is too powerful: it predicts vowel epenthesis to arise in environments where it never occurs. In fact, the present analysis demands that *every* superheavy syllable located at the postlexical level be 'fixed' with vowel epenthesis, due to the constraint ranking *SUPERHEAVY ≫ DEP-V. This, of course, is clearly false — a form such as **i̯eu̯ktrom* 'binding; cord' (Skt. *yoktra-*), which is composed of the root **i̯eu̯g-* 'yoke; join' plus the instrumental suffix *-tro-* plus acc.sg. *-m*, is predicted to undergo either schwa epenthesis (stem *[i̯eu̯k]$_\sigma$[trom]$_\sigma$ → postlexical *[i̯eu̯]$_\sigma$[kə]$_\sigma$[trom]$_\sigma$) or epenthesis of a corresponding syllabic sonorant (stem *[i̯eu̯k]$_\sigma$[trom]$_\sigma$ → postlexical *[i̯eu̯]$_\sigma$[ktr̥]$_\sigma$[rom]$_\sigma$). Both scenarios are conceivable within the framework of PIE syllabification proposed thus far.

(77) POSTLEXICAL *[i̯eu̯k]$_\sigma$[trom]$_\sigma$

/[i̯eu̯k]$_\sigma$[trom]$_\sigma$/	MAXST	MAX-C	*SPRHVY	FAITH(σ)	*COMPONS	DEP-V
a. ☹ [i̯eu̯k]$_\sigma$[trom]$_\sigma$			*!		*	
b. [i̯eu̯]$_\sigma$[trom]$_\sigma$		*!			*	
c. [i̯eu̯]$_\sigma$[ktr̥]$_\sigma$[rom]$_\sigma$				**!	*	*
d. ☞ [i̯eu̯]$_\sigma$[kə]$_\sigma$[trom]$_\sigma$				*	*	*

I presently see two possible routes that one may take to solve this problem of overgeneration. The first, and perhaps simplest, would be to abandon the requirement for syllable onsets at the postlexical level and to follow Sievers 1878 et al. in assuming SL to be a process of resyllabification, not epenthesis, as in (58) above: *[tert]$_\sigma$[i̯os]$_\sigma$ → *[ter]$_\sigma$[ti]$_\sigma$[os]$_\sigma$. In the OT analysis, we may simply replace the constraint that blocks vowel epenthesis (DEP-V) with one that blocks any change in the feature [syllabic] within the derivation.

(78) IDENT(σ): Do not alter the value of the feature [syllabic]. Assign one $*$ for every instance this feature has been altered.

To propose a constraint of this nature for PIE would not be controversial, since the feature [syllabic] alternated frequently, productively and cyclically in PIE sonorants, whose syllabicity (for the most part) depended on its surrounding phonological environment.[50] In addition, should we view SL as a process of resyllabification, the change of **terti̯os* to **tertios* (and not ***tertəos* or ***tertr̥os*) is a given, while the above analysis utilizing epenthesis does not provide a reason why **terti̯os* should epenthesize *[i] (→ **tertii̯os*) and not *[u], *[a], *[r̥], etc. Nevertheless, the assumption of SL as a process of resyllabification does lead to problems in the formal derivation.

(79) POSTLEXICAL *[i̯eu̯k]$_\sigma$[trom]$_\sigma$

/[i̯eu̯k]$_\sigma$[trom]$_\sigma$/	MAXST	MAX-C	*SPRHVY	FAITH(σ)	*COMPONS	IDENT(σ)
a. ☹ [i̯eu̯k]$_\sigma$[trom]$_\sigma$			$*!$		$*$	
b. [i̯eu̯]$_\sigma$[trom]$_\sigma$		$*!$			$*$	
c. ☞ [i̯eu̯]$_\sigma$[ktr̥]$_\sigma$[rom]$_\sigma$				$**$	$*$	$*$

In example (79) we see that the desired candidate *[i̯eu̯k]$_\sigma$[trom]$_\sigma$ loses, since superheavy syllables are avoided in all cases through sonorant resyllabification.

[50]See Kobayashi 2004:27 and section 1.5.

(80) POSTLEXICAL *[tert]$_{\sigma}$[i̯os]$_{\sigma}$

/[tert]$_{\sigma}$[i̯os]$_{\sigma}$/	MAXST	MAX-C	*SPRHVY	FAITH(σ)	*COMPONS	IDENT(σ)
a. [tert]$_{\sigma}$[i̯os]$_{\sigma}$			*!			
b. ☞ [ter]$_{\sigma}$[ti̯os]$_{\sigma}$				*	*	
c. [ter]$_{\sigma}$[i̯os]$_{\sigma}$		*!				
d. ☹ [ter]$_{\sigma}$[ti]$_{\sigma}$[os]$_{\sigma}$				*!*		*

In example (80) the candidate [ter]$_{\sigma}$[ti]$_{\sigma}$[os]$_{\sigma}$ is not chosen because it violates FAITH(σ) twice, making [ter]$_{\sigma}$[ti̯os]$_{\sigma}$ the most optimal candidate. Now, one might propose to switch the constraints FAITH(σ) and *COMPLEXONSET around (*COMPLEXONSET $\gg$ FAITH(σ)), thereby making *[ter]$_{\sigma}$[ti]$_{\sigma}$[os]$_{\sigma}$ the best candidate. This idea fails, however, since we need some mechanism to prevent words of the shape **tetsi̯os* and absolutives such as **tr̥ktu̯o-* from undergoing Sievers' Law, as seen in (81) below.

(81) *[tet]$_{\sigma}$[si̯os]$_{\sigma}$, *[tr̥k]$_{\sigma}$[tu̯o]$_{\sigma}$

/[tet]$_{\sigma}$[si̯os]$_{\sigma}$/, */[tr̥k]$_{\sigma}$[tu̯o]$_{\sigma}$/	MAXST	MAX-C	*SPRHVY	*COMPONS	FAITH(σ)	IDENT(σ)
a. ☹ [tet]$_{\sigma}$[si̯os]$_{\sigma}$				*!		
b. ☞ [tet]$_{\sigma}$[si]$_{\sigma}$[os]$_{\sigma}$					**	*
a. ☹ [tr̥k]$_{\sigma}$[tu̯o]$_{\sigma}$				*	*	
b. ☞ [tr̥k]$_{\sigma}$[tu]$_{\sigma}$[o]$_{\sigma}$					**	*

Thus it seems that SL should not be explained as a process of resyllabification and we should return to the original idea of epenthesis, with minor tweaks to the analysis. These tweaks should not only explain why the forms ***i̯eu̯kətrom* and ***i̯eu̯ktr̥rom* do not exist; it should also provide a reason for the choice in epenthetic segment in the output, such that **terti̯os* → **tertii̯os* and not ***tertu̯ios*, etc.

Perhaps a better solution to the problem would be to assume that SL is a process of vowel epenthesis, whose epenthetic segment has completely assimilated with an adjacent sonorant. Such a process of segment harmony is well attested typologically and in Indo-European, most often in partial assimilation. For example, in Sanskrit, sequences that continue *R̥̄ (< PIE *R̥H) develop an epenthetic vowel, whose quality is affected by adjacent segments: *ū* if following a labial or labiovelar consonant (*pūrṇá-* 'full' < *pl̥$h_1$nó-*), *ī* otherwise (*tīrṇá-* 'crossed' < PIE **tr̥$h_2$nó-*).[51] Another example may be found in Old Latin, where a short vowel produced by 'weakening' is rounded to [u] before a *pinguis*, or dark, *l* (**sikelos* > *Siculus* 'an inhabitant of Sicily') and before labial consonants (**pontifaks* 'high priest' > OLat. *pontufex*).[52] In both instances certain features of adjacent consonants have spread to the vowels in question.

To present formally this process of segment harmony in PIE we will need to assume two additional constraints for the postlexical level.

(82) ADDITIONAL CONSTRAINTS.

a. DEP[F]$_{IO}$: Every feature of the output has a correspondent in the input. Assign one ∗ for each instance a new feature is inserted.[53]

b. NO-SPREAD([F], seg): Feature-segment associations in the output must be reflected by the corresponding elements in the input. Assign one ∗ for each instance for each violation.[54]

The first constraint, DEP[F]$_{IO}$, requires that no additional feature be present in the surface form that is not underlying. The second, NO-SPREAD([F], seg), prevents the

[51] Fortson 2010:212.

[52] Weiss 2009:117-8. I refer the reader to Uffmann 2006:1080ff. for additional examples from non-IE languages.

[53] Struijke 2002:153. Note that I assume the default vowel of epenthesis, *[ə], to have had (a bundle of) features in PIE ([-high]? [-round]?).

[54] Struijke 2002:154, based on McCarthy 2000:159.

assimilation of features between segments. The necessary constraint ranking will be DEP[F]$_{IO}$ ≫ *SUPERHEAVY ≫ FAITH(σ) ≫ *COMPLEXONSET ≫ DEP-V, NO-SPREAD([F], seg). This will ensure that any segment epenthesized at the postlexical level will have assimilated as much as possible with its surrounding segments.

(83) POSTLEXICAL *[ter]$_\sigma$[ti̯os]$_\sigma$

/[tert]$_\sigma$[i̯os]$_\sigma$/	DEP(F)	*SPRHVY	FAITH(σ)	*COMPONS	DEP-V	NO-SPREAD
a. ☞ [ter]$_\sigma$[ti]$_\sigma$[i̯os]$_\sigma$			*		*	*
b. [ter]$_\sigma$[ti̯os]$_\sigma$			*	*!		
c. [ter]$_\sigma$[tə]$_\sigma$[i̯os]$_\sigma$	*!		*		*	
d. [tert]$_\sigma$[i̯os]$_\sigma$		*!				

(84) POSTLEXICAL *[i̯eu̯k]$_\sigma$[trom]$_\sigma$[55]

/[i̯eu̯k]$_\sigma$[trom]$_\sigma$/	DEP(F)	*SPRHVY	FAITH(σ)	*COMPONS	DEP-V	NO-SPREAD
a. ☞ [i̯eu̯k]$_\sigma$[trom]$_\sigma$		*		*		
b. [i̯eu̯]$_\sigma$[kə]$_\sigma$[trom]$_\sigma$	*!			*	*	

4.4 Consequences of Analysis.

Having fixed the problem of overgeneration with the introduction of the constraints DEP[F]$_{IO}$ and NO-SPREAD([F], seg), let's now turn to the consequences of our analysis, of which there are many.

[55] At the moment it is not entirely clear to me why the candidate *[i̯eu̯]$_\sigma$[ktr̥]$_\sigma$[rom]$_\sigma$ is not found as a Sievers variant. Perhaps this may be attributed to the markedness of the sequence $PPR̥- in PIE, as is seen in the simplification of the sequence $PPN̥-: *#d^{h}g^{h}m̥-* → *#g^{h}m̥-* (Mayrhofer 1986:117-8).

4.4.1 Advantages.

There are a number of advantages to this analysis of SL. First, by assuming that SL was driven by the desire to avoid superheavy syllables at the postlexical level, we have provided a motivation that is well attested both in Indo-European and cross-linguistically. Unlike many studies of SL in the past, our analysis is not circular, since we have not based the syllabification rules of PIE upon our analysis of Sievers' Law itself. Rather, we have based them on the phonotactic analysis of edge consonant clusters, as I have discussed in the previous two chapters. The adoption of the MAXST as the largest possible syllable shape in PIE neatly explains Schindler's two exceptions to SL discussed above: the absolutives and words of the shape **tets-i̯os*. SL does not occur in the absolutive **tr̥k-tu̯o-* ([tr̥k]$_\sigma$[tu̯o]$_\sigma$) because a superheavy coda never existed, as one was never created by the morphology. SL did not occur in words of the shape **tetsi̯os*, since the PIE MAXST did not permit it to be syllabified as [tets]$_\sigma$[i̯os]$_\sigma$ at the stem, word or postlexical levels, because an SSP violation would have resulted in the coda. In both instances Sievers' Law was never triggered at the postlexical level because a superheavy syllable was never created at any point in the phonological derivation.[56]

4.4.2 Disadvantages.

To my knowledge, the sole downside to the above analysis is that it requires SL and Lindeman's Law (LL) to have been separate phonological processes in PIE. Whereas SL targets syllables of the shape **-RV-* that immediately follow a superheavy syllable such as [tert]$_\sigma$, LL targets the onsets of monosyllabic words of the shape **#TRV-*:

[56]As Kie Zuraw points out to me, the above analysis of Sievers' Law, if correct, argues fairly strongly in favor of phonological strata within the grammar, since one phonological structure is shown to be 'fixed' at the stem and word level (**tetsi̯os* → *[tet]$_\sigma$[si̯os]$_\sigma$, to satisfy the MAXST), while another structure, *[tert]$_\sigma$[i̯os]$_\sigma$, is 'fixed' at the postlexical level (*[tert]$_\sigma$[i̯os]$_\sigma$ → *[ter]$_\sigma$[ti]$_\sigma$[i̯os]$_\sigma$, to satisfy *SUPERHEAVY).

**di̯ḗu̯s* 'sky (god)' → **dii̯ḗu̯s*. Their structural descriptions are fundamentally different, and therefore the collapse of SL and LL as a unitary process is a mirage.

As discussed above, Schindler (1977a:64) had assumed that both processes were restricted to the onset of the final syllable of a word. Of course, since LL only applied within monosyllables, his claim here is irrefutable. However, we do find instances of SL outside of this particular configuration, such as in the Ved. hapax legomenon *poṣyā́vant-* 'creating property' as *poṣiyā́vant-* (Sihler 2006:186). If this form is truly archaic, Schindler's formulation of SL cannot explain this form, whereas the above analysis can in a straightforward fashion: the verbal suffix *-ya-* → *-iya-* due to the preceding superheavy syllable *poṣ-* (from earlier **pauṣ-*).[57] Moreover, while Schindler (1977a:62-3) must assume paradigmatic leveling as the source of SL in forms such as *kā́viyasya* 'prophetic; of a kavi (gen.sg.)', the present analysis predicts both *kā́viyas* 'prophetic; of a kavi (nom.sg.)' and *kā́viyasya* 'prophetic; of a kavi (gen.sg.)' to have been created equally by SL, prompted by the avoidance of a superheavy syllable at the postlexical level at the juncture of a superheavy root (*kāv-*) + a suffix of the shape *-RV-* (*-ya-*). As to why SL occurs so frequently in the onset of the final syllable of a word: this must be due to the fact that this position is the most common environment for the juxtaposition of a superheavy syllable (which is always a root) before a syllable of the shape **-RV-* (which is always a suffix).[58]

There is, in fact, another and perhaps more compelling reason to view SL and LL as separate phonological processes, one which is not theory-internal. As Craig Melchert has pointed out to me, should we continue to view SL and LL as the same process,

---

[57]However, as Brent Vine points out to me, this form may simply be a later derivative of *poṣ(i)ya-* and therefore would provide no evidence of SL occurring in a non-final syllable. Further investigation of the application of SL in non-final syllables is required.

[58]So also Peter Barber (p.c.). Unfortunately at this time I do not have a solution for why SL is frequently unattested in vr̥ddhi formations such as *vāiśvānará-* (Schindler 1977:62), and must assume for the time being that this non-application is a later (Indic?) innovation.

we must accept one of two views: either 1) SL applied to *all* resonants in PIE or 2) LL applied to *only* glides. Should we reject both positions, it would be impossible to view the two rules in question as a unitary phenomenon, since they would not target the same segments. Though some scholars (beginning with Osthoff and followed most notably by Edgerton 1934:257ff.) have proposed that SL extended to all sonorants,[59] most of the recent IE handbooks and phonological treatments,[60] such as Mayrhofer 1986:167 (which strongly reflects Schindler's views), insist that SL applied almost entirely to glides. This problem deserves a thorough examination, as was recognized by Schindler himself (1977:64^6).[61]

The second alternative, restricting LL to glides, would be even more problematic. In addition to cases of LL occurring in clusters of the shape C + U̯, such as **d(i)i̯ḗus* 'sky (god)', **d(u)u̯oh*$_x$ 'two' and **k̂(u)u̯ō* 'dog', there are many well-established examples of C + non-glide, such as Lat. *homō*, Goth. *guma* 'man' from **(d*^{h}*)ĝ*^{h}*m̥mō* 'earthling',[62] and Boeot. *banā́*, Gk. *gunḗ* 'woman' from **g*^{w}*n̥nah*$_2$,[63] to name a few. Eschewing solid examples of LL such as these for the sake of collapsing two phonological rules together is in my opinion (and likely in the opinion of many within the IE scholarly community) not the best route to take. Of course, if we accept the findings of my analysis above, which postulates that SL and LL could not have been the same process since Schindler's syllabifications of forms of the shape **tetsi̯os* were in fact

[59] See most recently Ruijgh 1996 for Greek and Praust 2000 (see above).

[60] Meier-Brügger 2003:90-1; 141-2, Sihler 2006:183, Weiss 2009:39-40, Fortson 2010:72.

[61] The IE facts aside, the restriction of SL to glides is understandable, since, as Fleischhacker 2005 shows, *TU̯-* clusters (*pi̯a, ku̯a*, etc.) are more 'splittable' than *TR-* clusters (*pra, tma, kla*, etc.); in other words, *CUU̯-* and *CU̯-* are more perceptually similar to each other than are *CR̥R-* and *CR-* (cf. Sihler 2006:181). In fact, Fleischhacker has demonstrated there to be a gradient scale of 'splittability': in order of least splittable to most, CT → Cm → Cn → Cl → Cr → CU̯ (cf. Zuraw 2007:284). This raises the question, if non-glides *did* participate in SL, can any gradience be found, such that PIE suffixes of the shape **-rV-* were more likely to undergo SL than, say, those of the shape **-lV-*, **-nV-* and **-mV-*?

[62] Weiss 2009:105.

[63] Cf. Vine 1999:560ff.

impossible, such problems become irrelevant.[64]

4.4.3 Predictions.

The above analysis provides us with a straightforward definition of Sievers' Law in PIE and makes clear predictions of what should be attested in the Indo-European languages. It proposes that SL originated as a phonological process that altered suffixes of the shape **-RV-*, and therefore all instances of SL occurring in suffixes of the shape **-CRV-* must be secondary. This fact may explain why certain language groups, such as Germanic, Tocharian, Italic and Baltic, do not provide evidence for SL in suffixes of the shape **-CRV-*. Furthermore, it would explain why instances of certain suffixes such as Ved. *-bhiyas* are actually rarer after a heavy syllable (CVC, C$\bar{\text{V}}$) than are instances of *-bhyas*. Moreover, should one extend SL to apply to all sonorants, it would provide an answer to why no SL variants of the instrumental suffixes **-tro-*, **-tlo-*, **-*d^h*ro-* and **-*d^h*lo-* are ever attested, despite the fact that there are certain cases of the perfect ending *-ire* in Sanskrit which may have originated as Sievers' variants to *-re*, as Praust (2000) has argued. Of course, if we follow Schindler in assuming that SL is to be explained as the avoidance of a complex onset **TR-* in medial position, there is in principle no reason why **-tro-*, etc. should behave any differently from **-ro-*, etc. in this regard.[65]

---

[64]The problem of LL should be approached in much the same fashion as we have been done for SL, attempting to answer the following questions: 1) What was the motivation for LL? Was it driven by an avoidance of a complex onset of the shape **TR-*, and if so, why did it only occur in monosyllables? 2) If LL did create sandhi variants within a particular higher level constituent (intonational phrase? utterance?), can the exact conditions be ascertained? 3) Can this phonological process be connected to broader prosodic phenomena in PIE such as foot structure or a minimal word requirement?

[65]An additional prediction is that words of the shape **testi̯o-* and **teh*$_x$*ti̯o-* should in fact have undergone SL in PIE, since fricatives were likely more sonorous than stops and therefore the coda -FP]$_\sigma$ would not have violated the SSP. In other words, **testi̯o-* and **teh*$_x$*ti̯o-* would have been syllabified as *[test]$_\sigma$[i̯o]$_\sigma$ and *[teh$_x$t]$_\sigma$[i̯o]$_\sigma$, respectively, at the stem level and thus would have undergone SL at the postlexical level. This perhaps would explain the form *gabhastyos* 'hands (gen.abl.du.)', which is attested twenty times in the Rigveda and is always to be scanned as *gabhastiyos* (Sihler 2006:187-8).

4.5 Summary and Conclusions.

In this chapter, I have proposed that SL was driven in PIE by the well-attested Indo-European tendency to avoid superheavy syllables in medial position, providing a straightforward motivation to Sievers' Law. The solution is non-circular, as the rules of PIE syllabification used in the analysis are not based on our analysis of SL itself, but rather are based on the independent phonotactic analysis of the PIE phonology as a whole. The solution addresses why SL does not occur in words of the shape **tetsi̯o-* (Ved. *matsya-*) and in the absolutive (Ved. *yuktvā́*), since a superheavy syllable never existed at any point within the derivation. Lastly, the solution predicts all instances of SL occurring in suffixes of the shape *-CRV-* to be secondary, a prediction corroborated by the sparse and irregular attestation of SL within suffixes of this shape throughout the Indo-European languages. In the next chapter I will apply the findings of my analysis of SL to the problem of IE syllabification in general, addressing the remaining two unexplained exceptions to Schindler's rule of syllabification, the nasal-infix presents and the accusatives of sonorant-final stems, and reconciling Kobayashi's view of ONSET MAXIMIZATION with Hermann's parsing of the dimorphemic sequence VCCV as VC.CV.

I leave this matter open for future investigation.

CHAPTER 5

Concluding Remarks.

5.1 Overview of Findings.

We have now reached a point in our discussion where we are able to present a better overall formulation of PIE syllabification. This reformulation will accept Kobayashi's ONSET MAXIMIZATION principle, which was argued to be an improvement over Schindler's iterative 'right-to-left' algorithm in section 1.5. In addition, our reformulation should be able to explain the remaining two exceptions given in Schindler 1977a, as well as incorporate Hermann's findings of bipartite cluster division. However, before doing so it would first be beneficial to summarize the two main insights of the preceding chapters.

First, the analysis of reconstructable PIE edge sequences and the assumption of extrasyllabic consonants have enabled us to reconstruct a MAXIMUM SYLLABLE TEMPLATE (MAXST) for PIE. The MAXST was an undominated constraint in the PIE grammar that restricted possible phonotactic sequences in word-medial position. It has been defined as follows: "The maximum PIE syllable consists of two consonants in the onset and two consonants in the coda. The onset may violate the SONORITY SEQUENCING PRINCIPLE; the coda may not." Any sequence that violated the MAXST resulted in consonant deletion or vowel epenthesis, depending on whether the resulting cluster was legal in PIE. Many edge sequences that violated the MAXST contained extrasyllabic segments, consonants that can only be syllabified at word's edge since they may attach to a higher prosodic category. Consequently, this template is able to explain a curious feature discovered by Schindler (1977a) regarding Sievers' Law – the non-application within medial sequences of the shape *-OOR- (or *-OOU̯-), such as is found in Ved. *matsya-* 'fish' and PGk. *dik^ht^hi̯o-* 'double' (> Ion. *diksós*), since *-OO$]_\sigma$ was not a legal medial coda in PIE.[1]

The second main finding of this dissertation is the identification of a number of phonological conspiracies in PIE, or sets of phonological rules that serve the same basic purpose in the PIE grammar. In addition to the well-understood phenomenon of geminate avoidance (formally expressed as the constraint OCP in section 1.3.1), I have identified at least three such constraints reconstructable for PIE.

(85) PIE CONSPIRACIES.

a. MAXST

As previously mentioned, the MAXST was an undominated constraint in PIE, which required that word-medial consonant sequences be of a particular shape. Through it we have explained the PH.CC > P.CC rule, the avoidance of *s-epenthesis in the *métron* rule and the deletion of *t in PInd. *$h_x o\hat{k}th_x t$í-* (> Skt. *asītí-* 'eighty'). Instances of vowel epenthesis, schwa primum (PIE *d^hh_1só-* 'divine' → *$[\mathrm{d^h}\text{ə}\mathrm{h_1}]_\sigma[\text{só}]_\sigma$) and schwa secundum (*$d^h\hat{g}^h$més* → *$[\mathrm{d^h}\text{ə}\mathrm{\hat{g}^h}]_\sigma[\text{més}]_\sigma$), are very likely to have been prompted by violations of the MAXST as well.

b. *SUPERHEAVY

This constraint blocks syllables containing three or more morae. In chapter 3 I proposed that compensatory lengthening does not occur in word-medial position through avoidance of a *SUPERHEAVY violation

[1]Where *-OO$]_\sigma$ stands for *-PP$]_\sigma$ and *-PF$]_\sigma$, thereby resulting in a violation of the SSP.

(*$\hat{g}enh_1mnah_2$ → *[ĝen]$_\sigma$[mnah$_2$]$_\sigma$). In chapter 4 it was claimed that Sievers' Law was driven by the high ranking of *SUPERHEAVY at the postlexical level. It is also likely that certain cases of Schwebeablaut and the analogical replacement of weak full-grade forms with the zero-grade in roots of the shape *CeR(C) were motivated by violations of this constraint (at least partially so).

c. *RF]$_\sigma$

The coda sequence sonorant + fricative, while legal in late PIE, was blocked at an earlier stage of the proto-language. In addition to the well-established cases of fricative deletion in Szemerényi's law in **ptérs* 'father (nom.sg.)' (→ *[ptḗr]) and PIE *$\hat{u}édorh_2$ 'water (nom./acc. pl.) (→ *[u̯é]$_\sigma$[dōr]$_\sigma$), I utilized the constraint *RF]$_\sigma$ to explain laryngeal deletion in the medial sequence *RHCC (e.g. *$\hat{g}enh_1mnah_2$ → *[ĝen]$_\sigma$[mnah$_2$]$_\sigma$).

5.2 The Rules of Proto-Indo-European Syllabification Redefined.

Throughout this dissertation we have followed Kobayashi's sensible assumption of ONSET MAXIMIZATION as the underlying principle of syllable division in PIE: maximize onsets, minimize codas. Our study of Sievers' Law in chapter 4, however, has shown that this characterization is too simplistic overall and must be refined to explain the attested data. At the stem level, we find that onsets are maximized, though only within a given morpheme, in order to obey the highly-ranked ALIGN constraint. The syllabification generated at the stem level may then be modified at a higher stratum (word or postlexical) should the re-ranking of certain markedness constraints in the grammar demand it.

(86) GENERAL PRINCIPLE OF PIE SYLLABIFICATION.

Maximize onsets within a given morpheme at the stem level. Avoidance of marked sequences may change this formulation later in the derivation.

In this section we will examine three additional cases, which confirm the rule given in (86): nasal-infix presents, accusatives of the shape **-im(s)*, **-um(s)* and **-r̥m(s)* and the syllable division VC.CV within VCCV sequences, even when *CC- was a legal syllable onset. The first two are the remaining 'exceptional' syllabifications, as first identified by Schindler (1977a) and briefly discussed in section 1.5. We will see that their exceptionality results from a compliance with the constraint FAITH(σ) at the word level. The third case (VCCV → VC.CV) is especially problematic should we assume the principle of ONSET MAXIMIZATION applied to PIE. We will see that it may ultimately be attributed to a stem level satisfaction of the constraint ALIGN. Note that each case discussed is an example of the syllabification being affected by the morphology.

5.2.1 FAITH(σ) at the Word Level.

Chapter 4's analysis of Sievers' Law introduced the constraint FAITH(σ) at the postlexical level. This constraint requires that the output candidate not alter the syllabification of the input form – i.e., the syllabification produced at the word level. In this section I will assume that this constraint is also relevant at the word level, where I will postulate the nasal-infix morpheme and the accusative morphemes **-m(s)* to be added.[2] Therefore, at both the word and postlexical levels the high-ranking constraint FAITH(σ) requires that output candidates continue the syllabic parsing established at the previous level.

[2] At this point I have no independent evidence for these assumptions, though I am not aware of any counter-argument that suggests otherwise.

5.2.1.1 Nasal-infix Presents.

The nasal-infix presents form a verbal class which is continued to a greater or lesser extent by most attested IE languages.[3] Well-attested roots that form present stems with this infix include, for example, **i̯eu̯g-* 'join' (Eng. *yoke*) and **b^{h}ei̯d-* 'split' (Eng. *bite*). The suffix takes two ablauting shapes: **-né-* and **-n-*. While the strong form **-né-* has posed no problem to any previous (or present) conception of PIE syllabification (**i̯u̯-né-g-ti* → *[i̯u]$_\sigma$[nék]$_\sigma$[ti]$_\sigma$ '(s)he joins'; **b^{h}i̯-né-d-ti* → *[b^{h}i]$_\sigma$[nét]$_\sigma$[sti]$_\sigma$ '(s)he splits'), the weak form in **-n-* has been especially problematic. Should we follow Schindler's right-to-left syllabification algorithm or Kobayashi's principle of ONSET MAXIMIZATION, the infixation of a nasal into these roots should produce the forms ***iu̯n̥g-* and ***b^{h}i̯n̥d-*, respectively. However, this is not what we find: PIE */i̯u̯ngénti/ 'they yoke' →*[i̯un]$_\sigma$[gén]$_\sigma$[ti]$_\sigma$[4] (not **[i]$_\sigma$[u̯n̥]$_\sigma$[gén]$_\sigma$[ti]$_\sigma$) and PIE */b^{h}i̯ndénti/ 'they split' → *[b^{h}in]$_\sigma$[dén]$_\sigma$[ti]$_\sigma$ (not **[b^{h}i̯n̥]$_\sigma$[dén]$_\sigma$[ti]$_\sigma$).[5]

In the following analysis I will assume the nasal-infix suffix to be added at the word level, unlike the suffixes **-i̯o-* and **-tro-* discussed above in chapter 4. Though it is unclear exactly why certain derivational suffixes were added at the stem level vs. the word level, it is crucial to assume that the nasal infix entered the derivation at the latter. This is because if the suffix in question were added at the stem level, I see no way to explain its 'exceptional' syllabification within the current framework. I'll also tentatively assume that the first cycle of syllabification at the stem level produces *[i̯ug]$_\sigma$ and *[b^{h}id]$_\sigma$ for the nasal-infix roots 'join' and 'split', respectively.[6]

---

[3] See Fortson 2010:97 for a brief introduction.

[4] The weak stem is continued by Skt. *yuñjánti*, YAv. *yuṇjiṇti*, Lat. *iungō*, Lith. *jùngiu*, etc.

[5] The weak stem is continued by Ved. *bhindánti*, Lat. *findō*, etc.

[6] Kessler, n.d. has made a similar suggestion.

(87) STEM LEVEL: /i̯u̯g/

/i̯u̯g/	PARSE	ALIGN	*COMPCODA	*CODA	*COMPONS
a. ☞ [i̯ug]$_\sigma$				*	
b. [iu̯g]$_\sigma$			*!		
b. [i̯u]$_\sigma$g	*!	*			

The nasal infix is then added at the word level, where FAITH(σ) is ranked high in the grammar. This constraint requires that the stem syllabification of *[i̯ug]$_\sigma$ and *[b^{h}id]$_\sigma$ to remain as such in the output of the word level.

(88) WORD LEVEL: /[i̯ug]$_\sigma$ + -n-/

/[i̯ug]$_\sigma$ + n/	MAXST	FAITH(σ)	DEP-V	MAX-C	ALIGN	*SPRHVY
a. ☞ [i̯ung]$_\sigma$					*	*
b. [i]$_\sigma$[u̯n̥g]$_\sigma$		*!			*	
c. [i̯ug]$_\sigma$				*!		
d. [i̯u]$_\sigma$[nəg]$_\sigma$			*!		*	
e. [i̯ugn]$_\sigma$	*!				*	*

In this way we can motivate the actually occurring stem *[i̯ung]$_\sigma$, with a nucleus *[u] and not *[n̥], as being driven by faithfulness to the syllabification of an earlier stage of derivation, the stem level. The true syllabification, of course, was not *[i̯ung]$_\sigma$[én]$_\sigma$[ti]$_\sigma$, but *[i̯un]$_\sigma$[gén]$_\sigma$[ti]$_\sigma$. This satisfaction of the constraint ONSET could have occurred at either the word or postlexical level in the grammar, and even indirectly so to satisfy the high-ranking *SUPERHEAVY at the postlexical level. I leave this question open for future research.

There is another root structure that behaves differently with regard to the nasal-infixed presents: the class of disyllabic roots,[7] or roots of the shape *CePH-, such as **ma/eth$_2$-* ‘tear off’. In this class the nasal infix forms a syllable nucleus in the weak stems, such that */CeCH/ + /n/ → *[CeC]$_\sigma$[n̥H]$_\sigma$.[8] For example, the root **g^{h}r̥b^{h}h$_2$-* ‘grab’ formed a weak nasal-infixed stem *[g^{h}r̥b^h]$_\sigma$[n̥h$_2$]$_\sigma$, continued by Ved. *gr̥bhāyati* and YAv. *gəuruuaiieiti* (cf. OPers. *agarbāya*).[9] From what is now known of PIE syllabification through the MAXST, it is clear that *H cannot be syllabified with its preceding root segments in roots of the shape *CETH-. The reason /n/, when infixed, is syllabified in the output form is twofold. First, it syllabifies in order to avoid a MAXST violation: **[g^{h}r̥b^{h}nh$_2$]$_\sigma$. Second, its syllabification incurs no violation of FAITH(σ), because *H was never co-syllabified with the other root segments at the stem level.

(89) WORD LEVEL: /[g^{h}r̥b^h]$_\sigma$h$_2$+ n/

/[g^{h}r̥b^h]$_\sigma$h$_2$ + n/	MAXST	FAITH(σ)	DEP-V	MAX-C
a. ☞ [g^{h}r̥b^h]$_\sigma$[n̥h$_2$]$_\sigma$				
b. [g^{h}r̥b^{h}nh$_2$]$_\sigma$	*!*			
c. [g^{h}r̥nb^h]$_\sigma$				*!
d. [g^{h}r̥b^h]$_\sigma$[nəh$_2$]$_\sigma$			*!	

In short, the avoidance of a MAXST violation at the stem level for disyllabic roots is

[7]So-called because roots like **ma/eth$_2$-* become disyllabic **mat^{h}i-* in Sanskrit, etc. More broadly we may say that every PIE root of the shape CVC$_1$C$_2$was disyllabic, where C$_1$ was of an equal or less sonority than C$_2$, since C$_2$ could not be syllabified with the preceding segments of the root.

[8]For morphological discussion, see Jasanoff 2003:123ff., following Saussure 1879:241ff. Once again, I of course admit the actual syllabification to have been *[Ce]$_\sigma$[Cn̥H]$_\sigma$, though it must be established at what point in the derivation the constraint ONSET must be satisfied; that is, when does ONSET $\gg$ FAITH(σ)?

[9]Consequently, the weak stem of class IX verbs such as Ved. *gr̥bhnā́ti*, *gr̥bhnīté* (RV) must be viewed as secondary. The expected form is ****gr̥bhāté* (< [g^{h}r̥]$_\sigma$[b^{h}n̥h$_2$]$_\sigma$[tór]$_\sigma$).

what leads to the different syllabification treatments in the weak forms of the nasal-infixed presents.

5.2.1.2 Accusatives in **-im(s), *-um(s), *-r̥m(s)*.

The assumption of a high-ranking FAITH(σ) provides us with a mechanism to explain a second type of PIE syllabification that has been problematic for previous theories: the accusative singular and plural of acrostatic and proterokinetic **-i-*, **-u-* and **-r-* stems.[10] Examples such as **méntim* 'mind', **s(e)uh$_x$num(s)* 'son(s)' and **p(h$_2$)tr̥m(s)* 'fathers' defy both Schindler's and Kobayashi's systems of IE syllabification, which both predict ***méntim̥*, ***s(e)uh$_x$nu̯m̥(s)* and ***p(h$_2$)trm̥s*, respectively. However, we may approach these accusative forms in the same manner as we have done for the nasal-infix presents: assume that adherence to a high-ranking FAITH(σ) is driving the violation of the expected maximization of onsets.

(90) STEM LEVEL: **men-* + **-ti-* → *[men]$_\sigma$[ti]$_\sigma$

At the stem level, onsets are maximized within their given morphemes.

(91) WORD LEVEL: [men]$_\sigma$[ti]$_\sigma$ + *-m*[11]

/[men]$_\sigma$[ti]$_\sigma$ + -m/	MAXST	FAITH(σ)	DEP-V	MAX-C	ONSET
a. ☞ [men]$_\sigma$[tim]$_\sigma$					
b. [men]$_\sigma$[ti̯m̥]$_\sigma$		*!			
c. [men]$_\sigma$[ti]$_\sigma$[m̥]$_\sigma$					*!

[10] Schindler 1977a:57. See Fortson 2010:119ff. for discussion of the acrostatic and proterokinetic ablaut classes, with examples.

[11] The constraint ranking DEP-V, MAX-C ≫ ONSET is taken from Keydana 2004:168-9.

At the word level, the form **méntim* 'mind (acc.sg.)' is chosen over ***ménti̯m̥* because the latter violates FAITH(σ).

Keydana (forthcoming b) prefers a phonological motivation for the 'exceptional' syllabifications of accusatives to proterokinetic **i-*, **u-* and **r*-stems. He hypothesizes that these accusative forms were not in fact exceptional but rather follow the normal rules of syllabification. Keydana suggests that cases of apparent onset maximization are in fact driven by a high-ranking constraint *R/C 'Don't have a coronal sonorant in the coda'.[12] However, if we accept Keydana's analysis, then the nasal-infix presents become quite surprising – why would a speaker of PIE syllabify **i̯u̯ng+V* as *[i̯un]$_\sigma$[g+V]$_\sigma$ if coronal codas were so marked? In fact, Keydana is required to assume that the syllabification of nasal-infix presents must be lexicalized.[13] The analysis given in section 5.2.1 above is simpler, as it utilizes one constraint to explain two separate 'exceptional' cases of syllabification, whereas Keydana must assume one constraint to handle the syllabification of the accusatives in question, but lexicalization to explain the nasal-infix presents.[14] Moreover, we have seen in section 1.5 that the constraint rankings Keydana has assumed for PIE simply do not explain all of the forms reconstructable for PIE, as they do not explain ONSET MAXIMIZATION in instances where the rightmost non-coronals are selected as the syllable nucleus, as may be seen in **k̂li̯tós* → *[k̂li][tós].

[12]Cf. section 1.5.

[13]"Zweitens ist CU*n*C dann lizenziert, wenn -*n*-Morphemstatus hat. So wird das präsensstammbildende *n*-Infix immer nichtsilbisch realisiert: *sunve* (: *SAV*), *pinvate* (: *PAY*^{1}) etc. Diese Besonderheit des -*n*- Infixes muß – da es allein betroffen ist – lexikalisch sein."

[14]Brent Vine kindly points out to me that analogy to the strong stem *[i̯u]$_\sigma$[nég]$_\sigma$ is, in fact, a very sensible and straightforward way to produce syllabifications of the type *[i̯un]$_\sigma$[gén]$_\sigma$[ti]$_\sigma$. In a sense, one may view the explanations given in this and the preceding section as analogical, with morphological factors driving the syllabifications in both types.

5.2.2 Why should VCCV be syllabified as VC.CV?

Lastly, the principle given in (86), whereby onsets are maximized within their given morpheme, may also be able to explain a curious fact of PIE syllabification identified by Hermann (1923:351ff.) — the division of medial VCCV sequences as VC.CV even when CC was a legal onset.[15] Cf. *$[h_2a\hat{k}]_\sigma[ro]_\sigma$ 'high' (Gk. *ákros*, OIr. *ēr*), *$[put]_\sigma[lo]_\sigma$ 'little one' (Ved. *putrá-* 'son', Osc. **puklum** 'id.'), etc.

As we saw in section 1.5, this syllable division is not predicted by a system of IE syllabification that maximizes all onsets with no consideration of morpheme boundaries, as suggested by Kobayashi. Aside from the Lallwörter **atta* 'daddy' and **kakka* 'poo-poo', disyllabic sequences of the shape VCCV were overwhelmingly dimorphemic in PIE.[16] Moreover, since the minimal root structure was of the shape (C)VC in PIE,[17] it follows that a sequence VCCV would frequently consist of a root ending in VC plus a suffix beginning in CV. Following the principle of PIE syllabification given in (86), each such sequence would have been syllabified as VC.CV:[18]

There are two other possible permutations of a dimorphemic VCCV, both of which are attested: VCC + V and V + CCV. The first, VCC (root) + V (suffix) must always be parsed as VC.CV: PIE **b^hei̯d-* 'split' + *e.ti* → **b^hei̯.de.ti*, not ***b^hei̯d.e.ti*. This may be attributed to a requirement that a syllable onset be filled at the stem, word or postlexical level (ONSET)[19] or to the avoidance of a superheavy syllable at the postlexical level, as discussed in detail above. The second, V (root) + CCV, (suffix) is trickier. Were words like **h_1i-tro-* (> OIr. *ethar* 'ferry boat') and **tn̥-tlo-* (> Lith. *tiñklas* 'net') syllabified

[15] See also Marchand 1958:77ff. and Keydana 2004:173.

[16] Trimorphemic and tetramorphemic sequences will not be discussed at this time.

[17] Cf. Benveniste 1935:149ff.

[18] Weiss 2009:280.

[19] See Keydana 2004.

as *$[h_1i]_\sigma[tro]_\sigma$ and *$[t\d{n}]_\sigma[tlo]_\sigma$, respectively, in PIE?[20] And if so, could there have existed a syllabic contrast between a hypothetical **u̯r̥-tro-* 'repellent' (*$[$u̯r̥$]_\sigma[tro]_\sigma$)[21] and a hypothetical **u̯r̥t-ro-* 'turned' (*$[$u̯r̥t$]_\sigma[ro]_\sigma$)? This analysis predicts that there would have been such a contrast.[22] A claim of this nature for PIE must remain highly speculative for the time being, though I hope to pursue such an account in further detail at a later date.

5.3 Future Directions.

This dissertation has been devoted to the problem of reconstructing PIE syllabification. In it I hope to have given solutions that are grounded in typology, phonological theory and, most importantly, data that are securely reconstructable through careful philological analysis. I have claimed that a non-circular reconstruction of syllabification for a proto-language is possible if one accepts the well-established hypothesis that medial codas and onsets are affected by what types of clusters exist at word's edge (the DECOMPOSITION THEOREM). By doing so I believe that we may unify a number of phonological rules as sequence-specific responses to violations of a single phonological constraint, the MAXIMUM SYLLABLE TEMPLATE (MAXST). Lastly, I have posited that morphology plays a role in the process of PIE syllabification, realized through the constraints ALIGN and FAITH(σ). I trust the veracity of each of these

[20]Forms taken from Wodtko et al. 2008.

[21]Cf. **u̯ertrom-* > Skt. *vártram* 'protective dam, pond', Av. *varəθra-* 'shield', Middle Welsh *gwerthyr* 'fortress' (Olsen 1988:7).

[22]Though such cases are rare in English (as they probably would have been in PIE), note the difference in syllabification between Eng. *mistake* 'error' ([mɪsteɪk]) and *mis-take* 'accidentally pick up' ([mɪst^{h}eɪk]). The latter must be syllabified as $[$mɪs$]_\sigma[$t^{h}eɪk$]_\sigma$, due to the aspiration of /t/ (cf. Davidson-Nielsen 1974:38ff.). Elsewhere in Indo-European we find similar constrasts in syllabification driven by morphology. In Greek, cf. *ek.lúei* '(s)he loosens' (Euripides, *Phoenissae* 695) vs. *é.kluon* 'they heard' (Euripides, *Phoenissae* 919); in Latin, cf. *ab.rumpō* 'I break off' vs. *tene.brae* 'shadows, darkness' (Plautus); see Devine & Stephens 1994:35 for further discussion.

claims will be vetted by scholars in the not-too-distant future.

In the meantime, I would like to suggest just a few topics potentially interesting for future research. Of course, these are not the only unanswered questions to arise in this dissertation, as many may be found scattered in the footnotes of the preceding pages.

- Cross-linguistically, it is very common for unstressed vowels and nuclei of lower sonority to license fewer consonants in the syllable margins. Could this have also been the case in PIE? That is to say, would it be possible to reconstruct a REDUCED SYLLABLE TEMPLATE for PIE? If so, perhaps we may identify the rule *#CCN̥ → *#CN̥- (PIE *$d^h\hat{g}^h$*m̥(m)-* > Gk. k^h*amaí* 'on the ground') as an instance of STRAY ERASURE, given the low sonority of the nucleus **m̥*.

- The reconstruction of a PIE MAXIMUM SYLLABLE TEMPLATE has neatly explained a number of syllable-driven phonological processes of vowel epenthesis and consonant deletion. Can we identify similar processes within the prehistories of the individual IE daughter languages? For exemplary studies of the syllable structures of ancient IE languages, I refer the reader to Steriade 1982 (primarily Greek), Devine & Stephens 1994 (Greek) and Kobayashi 2004 (Sanskrit).

- Lastly, the work done in theoretical linguistics has made it possible for us to reconstruct PIE syllabification in a non-circular way. I believe that this knowledge of the IE syllable may be applied to the study of syllabification in general, especially the MAXST. The oddity that onsets – but not codas – could violate the SSP should be examined from a typological point of view, and it may have ramifications for the synchronic analyses of the syllable structure of many modern IE languages spoken today.

APPENDIX A

Glossary of Concepts and Constraints

ALIGN	For every morpheme boundary, there must be a syllable boundary.
ALIGNNUC	ALIGN(Nucleus, R, σ, R): Align the right edge of a syllable nucleus with the right edge of a syllable.
*CHCC	The sequence *CHCC is prohibited in the output.
*COMPCODA	*COMPLEXCODA. Codas may not contain more than one consonant in the output.
*COMPONS	*COMPLEXONSET. Onsets may not contain more than one consonant in the output.
DEP[F]$_{IO}$	Every feature of the output has a correspondent in the input.
DEP-[s]	Every **s* in the output has a correspondent in the input.
DEP-V	Every vowel in the output has a correspondent in the input.
DT	DECOMPOSITION THEOREM. Medial consonant clusters are decomposable into a sequence consisting of a coda plus onset, whose syllable division is produced by the interaction of a speakerUs knowledge of consonant sequences at wordUs edge and syllable markedness constraints.
FAITH(σ)	Do not alter the syllabification of the base form.
*FINALVOICE	No surface form may contain a final voiced obstruent.

HNUC	When there is more than one segment which can become the nucleus of a syllable, the nucleus is assigned to the one with the highest sonority.
IDENT(σ)	Corresponding input and output segments have identical values for the feature [syllabic].
IDENT(VOICE)	Corresponding input and output segments have identical values for the feature [voice].
INTACT	Do not delete the root in its entirety.
*K^wu̯/*u̯K^w	A labiovelar may not be directly adjacent to a *u̯ in the output.
LICENSE	Consonants must be properly licensed.
MAX-C	Every non-laryngeal consonant in the input has a correspondent in the output.
MAX-H	Every laryngeal consonant in the input has a correspondent in the output.
MAX-/s, R/	Every **s* and sonorant in the input has a correspondent in the output.
MAX-T	Every dental stop in the input has a correspondent in the output.
MAX-μ	Every mora in the input has a correspondent in the output.
MAXST	MAXIMUM SYLLABLE TEMPLATE. The maximum PIE syllable consists of two consonants in the onset and two consonants in the coda. The onset may violate the SSP; the coda may not.
NOCODA	No syllable may have any consonants in the coda.

NO-SPREAD	NO-SPREAD([F], seg). Feature-segment associations in the output must be reflected by the corresponding elements in the input.
OCP	THE OBLIGATORY CONTOUR PRINCIPLE. Two identical (heteromorphemic) segments may not be adjacent to each other.
ONSET	A syllable must have an onset.
PARSE	Syllabify all segments.
*$_{\sigma}$[PPR	The sequence two stops + sonorant (*r, *l, *m, *n, *u̯) is prohibited in the onset of the output.
*RH$	The output may not have the sequence sonorant + laryngeal immediately preceding a syllable boundary.
SSP	SONORITY SEQUENCING PRINCIPLE. Between any member of a syllable and the syllable peak, only sounds of higher sonority rank are permitted.
*SUPERHEAVY	No medial syllable may consist of three or more morae.
W-S(I)	For any I, a syllable-initial segment, there is a word such that its initial segment is identical to I.
W-S(F)	For any F, a syllable-final segment, there is a word such that its final segment is identical to F.
*VCHCV	The sequence *VCHCV is prohibited in the output.

APPENDIX B

Proto-Indo-European Edge Phonotactics

Directly attested, unless otherwise noted; (i) = Indirectly attested; (e/s) = Reconstructed for etymological / structural reasons[1]

WORD-INITIAL

BIPARTITE

#PN-:

*pn-, *dm-, *d^hm-, *ĝn-, *$\hat{g}^h$n-, *gn-, *g^wn-, *g^{wh}n-.[2]

#PL-:

*pr-, *pl-, *tr-, *ḱr-, *ḱl-, *kr-, *kl-, *k^wr-;[3]

*dr-, *dl-, *ĝr-, *ĝl-, *gr-, *gl-, *g^wr-;[4]

*b^hr-, *b^hl-, *d^hr-, $d^{h\text{-}}$l-, *g^hr-, *g^hl-, *g^{wh}r-.[5]

#PU̯-:

*ti̯-, *di̯-, *d^hi̯- (i), *ḱi̯-, *ĝi̯-, *$\hat{g}^h$i̯-, *ki̯-, *g^wi̯-;[6]

*tu̯-, *du̯-, *d^hu̯-, *ĝu̯-, *ḱu̯-, *$\hat{g}^h$u̯-.[7]

#RR-: Only attested in *mn-, *mr-, *ml-, *mi̯-, *u̯r-, *u̯l-, *u̯i̯-.[8]

#FR-:

#sR-: *sm-, *sn-, *sr-, *sl-, *su̯-, *si̯-;[9]

#h_1R- (i): *h_1n-, *h_1r-, *h_1l-, *h_1u̯-;[10]

#h_2R- (i): *h_2m-, *h_2n-, *h_2r-, *h_2l-, *$h_2\underaccent{\frown}{u}$-;[11]

#h_3R- (i): *h_3m-, *h_3n-, *h_3r-, *h_3l-, *$h_3\underaccent{\frown}{u}$-;[12]

#$h_x\underaccent{\frown}{i}$- (i): well-attested.[13]

#FF-:

#sh_x- (i): *sh_2-;[14]

#h_xs- (i): *h_1s-, *h_2s-.[15]

#FP-:

#sP-: *sp-, *st-, *$s\hat{k}$-, *sk-, *sk^w-;[16]

#zP-: *zg^w-;[17]

#z^hP-: *z^hb^h-, *z^hd^h-, *z^hg^{wh}-;[18]

#h_1P- (i): *h_1d-, *h_1g-, *h_1g^{wh}-;[19]

#h_2P- (i): *h_2t-, *$h_2\hat{k}$-;[20]

#h_3P- (i): *h_3p-, *h_3k^w-.[21]

#PF-:

#Ps-: *ps-, *ks-, *k^ws- (i);[22]

#PH- (i): *th_2-, *kh_2-.[23]

#PP-:

#TK-: *$t\hat{k}$-, *tk-, *$d^h\hat{g}^h$-, *d^hg^{wh}-;[24]

Other #PP: *$p\hat{k}$-, *pt-.[25]

TRIPARTITE

#FPR-:

#sPR-: *spr-, *spl-, *str-, *skr-;[26]

#h_xPR-: only attested in *$h_3b^hruh_x$- 'brow' (i).[27]

#OsC-:

#h_xsC- (i): *h_1st-, *$h_1z^hd^h$-, *h_2st-, *h_3sl-;[28]

#PsC-: *pst-, *bzd- (e/c), *ksn-, *ksu̯-, *k^wsp- (e/c).[29]

#PPC-: Only reconstructable in *$\hat{g}^h d^h$i̯- (i).[30]

WORD-FINAL

BIPARTITE

-RO#: *-ms, *-nd, *-ns, *-nh_2 (i), *-rd, *-ls, *-u̯d, *-u̯s, *-i̯d.[31]

-OT#: *-pt, *-$\hat{k}$t, *-st.[32]

-OF#:

-Os#: *-ps, *-ts, *-$\hat{k}$s, *-ks, *-k^ws;[33]

-Oh_x# (i): *-$d^h h_2$, *-$\hat{g}h_2$.[34]

TRIPARTITE

-CCT#: *-$\hat{k}$st, *-rst, *-u̯h_2t;[35]

-RCs#: *-nts, *-nks, *-rks, *-lks, *-u̯ks;[36]

-OCs#: *-sts, *-k^wts;[37]

-CCH# (i): *-$zd^h h_2$.[38]

Notes

[1]All clusters inferred and all representative examples containing said clusters have been taken from IEW, LIV, NIL, EWAia and Schindler 1972, unless otherwise noted. Given the limited number of references, this appendix of edge clusters should not be considered exhaustive. Questionable reconstructions will be marked with a (?). Note that a number of sequences containing laryngeals have been omitted, if they are not attested in languages in which laryngeals are not vocalized. For example, at this point it is difficult to decide whether *kréu̯h_2s 'flesh', continued solely by Ved. kravís- and Gk. kréas, was realized as such (with a tripartite coda) or as *kréu̯h_2əs, though perhaps the loss of laryngeal in Av. mraot̰ 'spoke' speaks in favor of a tripartite coda.

[2]*pneu̯- 'breathe', *dmah_2- 'build', *d^hmeh_x- 'blow', *ĝnoh_3- 'know', *ĝ^{h}neu̯- 'be ashamed' (?), *gnet- 'knead', *g^wnah_2- 'woman', *g^{wh}nént 'they slayed'.

[3]*prek̂- 'ask', *pleh_1- 'fill', *tréi̯es 'three', *k̂reth_2- 'loosen, slacken', *k̂lei̯- 'lean', *kreu̯h_2s 'flesh', *klep- 'steal', *k^wrei̯h_2- 'buy'.

[4]*drák̂u̯r̥ 'tear', *dlong^{h}o- 'long', *ĝrei̯- 'extend', *ĝlei̯- 'rush, attack' (?), *gres- 'devour', *glei̯h_x- 'coat, paste', *g^wrih_x- 'heavy'.

[5]*b^hrā́ter- 'brother', *b^hleg- 'shine', *d^hreg^h- 'run', *d^hlas- 'squeeze' (?), *g^hrebh_2- 'grab', *g^hlend^h- 'look; shine', *g^{wh}ren- 'diaphragm'.

[6]*ti̯eg^w- 'withdraw oneself', *di̯eu̯- 'heaven', *d^hi̯a$h_2$mn̥ 'sign' (?), *k̂i̯eh_1- 'dark-colored' (?), *ĝi̯eu̯h_x- 'chew', *ĝ^{h}i̯em- 'winter', *ki̯eh_1- 'move', *g^wi̯oh_3- 'live'.

[7]*tu̯erk- 'cut', *du̯oh_x 'two', *d^hu̯or- 'door', *k̂u̯on- 'dog', *ĝu̯erh_x- 'be hot', *ĝ^{h}u̯er- 'be crooked'.

[8]*mnah_2- 'think about', *mreĝ^{h}u- 'short', *mleu̯h_2- 'speak', *mi̯eu̯h_1- 'move', *u̯reh_1ĝ- 'break', *u̯lei̯k^w- 'moisten', *u̯i̯eh_1- 'wrap'.

[9]*smer- 'think about', *snei̯g^{wh}- 'snow', *sreu̯- 'flow', *slei̯ĝ- 'smear', *su̯ah_2d- 'sweet', *si̯eu̯h_x- 'sew'.

[10]*$h_1$neu̯n̥ 'nine', *$h_1$rem- 'become quiet', *$h_1$leud^h- 'climb; grow', *h_1u̯eg^{wh}- 'speak solemnly'.

[11]*$h_2$meh_1- 'mow', *$h_2$nek̂- 'obtain; arrive', *$h_2$reu̯i- 'sun', *$h_2$leg- 'attend to, worry', *h_2u̯eks- 'grow'.

[12]*$h_3$mei̯ĝ^h- 'urinate', *$h_3$nei̯d- 'abuse', *$h_3$reĝ- 'rule; stretch', *$h_3$lei̯sd^h- 'slide' (?), *h_3u̯ath_2- 'in-

jure'.

[13]*h_xi̯ó- 'which', *h_xi̯aĝ- 'revere', etc. Exact laryngeal to be determined.

[14]*s$h_2$om- 'song'.

[15]*$h_1$sénti 'they are', *$h_2$seu̯s- 'become dry'.

[16]*spek̂- 'look at', *stah_2- 'stand', *sk̂eh_x(i̯)- 'shine', *skab^h- 'shave', *sk^{w}al- 'big fish'.

[17]*sg^{w}es- 'extinguish'. Cf. Southern 2000:36.

[18]Or #sP^h- by Siebs' Law. *sb^{h}eng- 'shine', *sd^{h}erb^h- 'become fixed', *sg^{wh}al- 'make a mistake' (?).

[19]*$h_1$dont- 'tooth', *$h_1$ger- 'wake up', *h_1g^{wh}el- 'wish'.

[20]*$h_2$teu̯g- 'terrify', *h_2k̂ou̯s- 'hear'.

[21]*$h_3$pus- 'copulate; marry' (?), *h_3k^{w}V- 'eye'.

[22]*pseh_x- 'tear; rub', *kseu- 'shave', *k^{w}sep- 'dark'.

[23]*t$h_2$au̯s- 'be quiet' (?), *k$h_2$aid- 'hit' (?).

[24]Perhaps #TsK-. *t(s)k̂ei̯- 'inhabit', *t(s)keh_1- 'acquire', *d^h(z^h)ĝ^{h}ŏm- 'earth', *d^h(z^h)g^{wh}ei̯- 'be destroyed'.

[25]*pk̂u- 'cattle', *ptero- 'feather'.

[26]*spreg- 'crackle', *splend- 'shine', *streu̯- 'strew', *skreb- 'scrape'.

[27]The oblique of *$h_1$étmō 'breath, soul' (Ved. ātmā́), is often reconstructed as *$h_1$tmen- (cf. Ved. tmánā 'soul (instr.sg.)'), though the only evidence is of structural/etymological nature.

[28]*$h_1$stí- 'existence', *$h_1$zd^{h}i 'be!', *$h_2$ster- 'star', *$h_3$slei̯d^h- 'slide' (?).

[29]*psten- 'breast, nipple', *bzdV- 'fart lightly', *ksneu̯- 'sharpen', *ksu̯ei̯b^h- 'throw, jerk' (?), *k^{w}spent- 'night'.

[30]*ĝ^{h}d^{h}i̯és 'yesterday'. Perhaps also *tk̂i̯- if the reconstruction *tk̂i̯a$h_2$ino- 'bird of prey' is correct (see Clackson 1994:143-4 and EWAia II 662 for discussion, with references).

[31]*ĝ^{h}i̯éms 'winter', *-nt '3rd pl. secondary ending', *-Vns 'masc.acc.pl.', *g^{w}énh_2 'woman', *k^{w}ért 'cut (3 sg. impfct.)', sáls 'salt', *di̯ḗu̯s 'sky (god)', *stḗu̯d 'praised (3 sg. impfct.)', *$h_1$éi̯d 'went (3 sg. impfct.)'.

[32]*$h_1$épt or *h_1ḗpt 'seized (3 sg. impfct.)', *u̯ḗk̂t 'wished (3 sg. impfct.)', *$h_1$ést 'was (3 sg. impfct.)'.

[33] *$h_2$áps ‘water’, *póts ‘foot’, *h_3rḗk̂s ‘king’, *i̯uks ‘joining, harnessing’, *u̯ṓk^{w}s ‘voice’.

[34] *-med$^h h_2$ ‘1st pl. M/P secondary ending’, *méĝh_2 ‘great’.

[35] *h_2u̯ḗk̂st ‘carried (3 sg. aor.)’, *d^{h}ḗrst ‘fastened (3 sg. aor.)’, *mleu̯h_2t ‘spoke (3 sg. impfct.)’.

[36] *$h_1$dó/énts ‘tooth’, *strínks ‘owl’ (?), *i̯órks ‘doe’, *$h_2$álks ‘courage’, *lé/óu̯ks ‘light’.

[37] *Kó/ésts ‘hunger’ (Vijūnas 2006:92), *nék^{w}ts ‘night (gen.sg.)’.

[38] *-mezd$^h h_2$ ‘1st pl. M/P secondary ending’ (?), if not *-mezd$^h h_2$a.

References

Adamczyk, Elżbieta. 2001. Old English Reflexes of Sievers' Law. *Studia Anglica Posnaniensia* 36.61–72.

Albright, Adam. 2008. Inflectional paradigms have bases too: evidence from Yiddish. In *The Bases of Inflectional Identity*, 271–312. Oxford & New York: Oxford University Press.

Anttila, Raimo. 1969. *Proto-Indo-European Schwebeablaut*. Berkeley & Los Angeles: University of California Press.

Bagemihl, Bruce. 1991. Syllable structure in Bella Coola. *Linguistic Inquiry* 22:4. 589–646.

Bailey, Todd M. and Ulrike Hahn. 2001. Determinants of Wordlikeness: Phonotactics or Lexical Neighborhoods? *Journal of Memory and Language* 44.568–91.

Barlow, Jessica A. 2001. Individual differences in the production of initial consonant sequences in Pig Latin. *Lingua* 111.667–696.

Beekes, Robert S.P. 1969. *The Development of the Proto-Indo-European Laryngeals in Greek*. The Hague & Paris: Mouton.

———. 1988a. *A grammar of Gatha-Avestan*. New York & Leiden: Brill.

———. 1988b. Laryngeal Developments: A Survey. In *Die Laryngaltheorie und die Rekonstruktion des indogermanischen Laut- und Formensystems*, 59–105. Heidelberg: C. Winter.

———. 1995. *Comparative Indo-European Linguistics: An Introduction*. Amsterdam & Philadelphia: John Benjamins Publishing Company.

Benveniste, Émile. 1935. *Origines de la formation des noms en indo-européen*. Paris: Maisonneuve.

Bermúdez-Otero, Ricardo. Forthcoming. *Stratal Optimality Theory (Oxford Studies in Theoretical Linguistics)*. Oxford: Oxford University Press.

Bessell, Nicola J. 1993. *Towards a phonetic and phonological typology of post-velar articulation*. PhD dissertation, University of British Columbia.

Blevins, Juliette. 1995. The Syllable in Phonological Theory. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, 206–244. Cambridge, Mass. and Oxford: Blackwell.

Blevins, Juliette and Andrew Wedel. 2009. Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica* 26/2.143–83.

Breen, Gavan and Rob Pensalfini. 1999. Arrernte: A Language with No Syllable Onsets. *Linguistic Inquiry* 30/1.1–25.

Brennan, Jon. 2006. A Weight Problem: Vowel Length in Classical Latin. Paper presented at the 8th SUNY/CUNY/NYU Miniconference, November 18th, 2006.

Broselow, Ellen. 1976. *The Phonology of Egyptian Arabic*. PhD dissertation, University of Massachusetts, Amherst.

Brugmann, Karl & Delbrück, Berthold. 1897-1916. *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen I/1-2, II/1-3*. Berlin: Strassburg.

Byrd, Andrew Miles. Forthcoming a. Deriving Dreams from the Divine: Hittite *tesḫa-/zasḫ(a)i-*. *Historische Sprachforschung*.

———. Forthcoming b. Motivating Sievers' Law. In Stephanie W. Jamison, H. Craig Melchert, and Brent Vine (eds.), *Proceedings of the 21st Annual UCLA Indo European Conference, October 30th - 31th, 2009*.

———. Forthcoming c. Predicting Indo-European Syllabification through Phonotactic Analysis. In *The Sound of Indo-European - selected papers from the conference, held in Copenhagen 16-19 April 2009*.

Calabrese, Andrea. 1994. Sievers' Law in Gothic: a Synchronic Analysis and Some Speculations on its Historical Development. *The Linguistic Review* 11.149–94.

———. 1999. Sievers' Law in Vedic. In H. van der Hulst and N. A. Ritter (eds.), *The Syllable: View and Facts*, 673–751. Berlin: Mouton de Gruyter.

Carlson, Katy. 1997. Sonority and reduplication in Nakanai and Nuxalk (Bella Coola). Ms., University of Massachusetts, Amherst.

Carlton, Terence R. 1990. *Introduction to the Phonological History of the Slavic Languages*. Columbus: Slavica Publishers, Inc.

Charpentier, Jarl. 1929. Some Remarks on Pashto Etymology. *Acta Orientalia* 7. 180–200.

Cho, Young-Mee and Tracy Holloway King. 2003. Semisyllables and universal syllabification. In Caroline Féry and Ruben van de Vijver (eds.), *The syllable in Optimality Theory*, 183–212. Cambridge: Cambridge University Press.

Clackson, James. 1994. *The Linguistic Relationship between Armenian and Greek*. Oxford & Cambridge: Blackwell.

Collinge, N.E. 1985. *The Laws of Indo-European*. Amsterdam & Philadelphia: John Benjamins Publishing Company.

Côté, Marie-Hélène. 2000. *Consonant cluster phonotactics: A perception-based approach*. PhD dissertation, MIT.

Dahl, Ivar. 1938. *Substantival inflexion in Early Old English. (Lund studies in English, 7)*. Lund: Gleerup.

Davidsen-Nielsen, Niels. 1974. Syllabification in English words with medial *sp, st, sk*. *Journal of Phonetics* 2.15–45.

de Vaan, Michiel. 2008. *Etymological Dictionary of Latin and the other Italic Languages*. Leiden & Boston: Brill.

Dell, François and Mohamed Elmedlaoui. 1985. Syllabic Consonants and Syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics* 7.105–30.

Derksen, Rick. 2008. *Etymological Dictionary of the Slavic Inherited Lexicon*. Leiden & Boston: Brill.

Derwing, Bruce L. 1992. A 'pause-break' task for eliciting syllable-boundary judgments from literate and illiterate speakers: Preliminary results for five diverse languages. *Language and Speech* 35.219–235.

Devine, A.M. and Laurence D. Stephens. 1994. *The Prosody of Greek Speech*. Oxford & New York: Oxford University Press.

Dubiel, S. 1994. Syllables and syllabification: a case study of Polish. Master's thesis, the Lublin Institute of English Philology.

Edgerton, Franklin. 1934. Sievers's Law and IE. Weak-Grade Vocalism. *Language* 10/3.235–65.

Edgren, A. Hjalmar. 1885. On the Relation in the Rig-Veda between the Palatal and Labial Vowels (i, î, u, û) and Their Corresponding Semivowels (y, v). *Journal of the American Oriental Society* 11.67–88.

Elbourne, Paul. 2000. Plain voiceless stop plus laryngeal in Indo-European. *Historische Sprachforschung* 113.2–28.

Fleischhacker, Heidi. 2005. *Similarity in phonology: Evidence from reduplication and loan adaptation*. PhD dissertation, University of California, Los Angeles dissertation.

Fortson, Benjamin. 2010. *Indo-European Language and Culture: An Introduction* (2nd ed.). Chichester, U.K. & Malden, Mass.: Wiley-Blackwell.

Frazier, Melissa. 2006. Accent in Proto-Indo-European Athematic Nouns: Antifaithfulness in Inflectional Paradigms. Master's thesis, University of North Carolina at Chapel Hill.

Fullerton, G. Lee. 1992. Proto-Indo-European (PIE) syllabification and Germanic nominal inflection. In Irmengard Rauch, Gerald Carr, and Robert Kyes (eds.), *On Germanic Linguistics: Issues and Method*, 81–96. Berlin & New York: Mouton de Gruyter.

Garrett, Andrew. 1998. Adjarian's law, the glottalic theory, and the position of Armenian. In Benjamin K. Bergen, Madelaine C. Plauché, and Ashlee C. Bailey (eds.), *Proceedings of the Twenty-Fourth Annual Meeting of the Berkeley Linguistics Society, February 14-16, 1998: Special Session on Indo-European Subgrouping and Internal Relations*, 12–23. Berkeley: Berkeley Linguistics Society.

Greenberg, Joseph H. 1978. Some Generalizations Concerning Initial and Final Consonant Clusters. In *Universals of Human Languages, Vol. 2: Phonology*, 243–79. Stanford, CA: Stanford University Press.

Hackstein, Olav. 1995. *Untersuchungen zu den sigmatischen Präsensstammbildungen des Tocharischen*. Göttingen.

———. 2002. Uridg. CH.CC > C.CC. *Historische Sprachforschung* 115.1–22.

Halle, Morris. 1959. *The Sound Pattern of Russian: a Linguistic and Acoustical Investigation*. The Hague: Mouton & Co.

Hammond, Michael. 1999. *The Phonology of English: a Prosodic Optimality-Theoretic Approach*. Oxford & New York: Oxford University Press.

Hamp, Eric. 1970. Sanskrit *duhitā*, Armenian *dustr*, and IE Internal Schwa. *Journal of the American Oriental Society* 90/2.228–231.

Harris, James W. 1983. *Syllable structure and stress in Spanish: A nonlinear analysis*. Linguistic Inquiry Monograph 8. Cambridge, MA & London: MIT Press.

Hdouch, Youcef Y. 2008. *Extraprosodicity and Syllable Structure in Berber: an Optimality-Theoretic Approach*. Muenchen: Lincom Europa.

Hermann, Eduard. 1923. *Silbenbildung im Griechen und in den andern indogermanischen Sprachen*. Göttingen: Vandenhoeck & Ruprecht.

Hill, Eugen. 2003. *Untersuchungen zum inneren Sandhi des Indogermanischen: der Zusammenstoss von Dentalplosiven im Indoiranischen, Germanishen, Italischen und Keltischen*. Bremen: Hempen Verlag.

Hinskens, Frans. 1996. *Dialect Levelling in Limburg: Structural and sociolinguistic aspects*. Tübingen: Max Niemeyer Verlag.

Hirt, Hermann. 1921-37. *Indogermanische Grammatik, I-VII*. Heidelberg.

Hoenigswald, Henry. 1988. A note on semivowel behavior and its implications for the laryngeals. In A. Bammesberger (ed.), *Die Laryngaltheorie*, 199–211. Heidelberg: Carl Winter.

Hoffmann, Karl and Bernhard Forssman. 1996. *Avestische Laut- und Flexionslehre*. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

Hoffner, Harry A. and H. Craig Melchert. 2008. *A Grammar of the Hittite Language, Part 1: Reference Grammar*. Winona Lake: Eisenbrauns.

Horowitz, Franklin E. 1974. *Sievers' law and the evidence of the Rigveda*. The Hague: Mouton.

Insler, Stanley. 1971. Some Problems of Indo-European *ə in Avestan. *Language* 47 .573–585.

Itô, Junko. 1988. *Syllable theory in prosodic phonology*. New York & London: Garland.

Jasanoff, Jay. 1978. *Stative and Middle in Indo-European*. IBS 23. Innsbruck: Institut für Sprachwissenschaft.

———. 1988. The Sigmatic Aorist in Tocharian and Indo-European. *Tocharian and Indo-European Studies* 2.52–76.

———. 1989. Old Irish *bé* 'woman'. *Ériu* 40.135–41.

———. 2002. The Vedic imperatives yódhi 'fight' and bodhi 'heed'. *Journal of the American Oriental Society* 122.2.290–295.

———. 2003. *Hittite and the Indo-European Verb*. New York: Oxford.

———. 2008. *g^wes-, *(z)g^wes- or *(s)g^wesh_2-? The PIE root for 'extinguish/go out'. In Claire Bowern, Bethwyn Evans, and Luisa Miceli (eds.), *Morphology and Language History. In honour of Harold Koch*, Current Issues in Linguistic Theory 298, 155–166. Amsterdam & Philadelphia : John Benjamins.

Joseph, Lionel S. 1982. The treatment of *CR̥H- and the origin of CaRa- in Celtic. *Ériu* 33.31–57.

Juret, C. 1913. *Dominance et résistance dans la phonétique latine*. Heidelberg: Winter.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Kahn, Daniel. 1980. *Syllable-Based Generalizations in English Phonology*. New York & London: Garland Publishing, Inc.

Kaisse, Ellen M. and Patricia A. Shaw. 1985. On the theory of Lexical Phonology. *Phonology Yearbook* 2.1–30.

Kapović, Mate. 2006. *Reconstruction of Balto-Slavic Personal Pronouns, with emphasis on accentuation*. PhD dissertation, University of Zagreb.

Kavitskaya, Darya. 2002. *Compensatory Lengthening: Phonetics, Phonology, Diachrony*. New York & London: Routledge.

Kenstowicz, Michael and Charles Kisseberth. 1979. *Generative Phonology: Description and Theory*. New York: Academic Press.

Kent, Roland G. 1953. *Old Persian* (2nd ed.). New Haven: American Oriental Society.

Kessler, Brett. n.d. On the phonological nature of the Proto-Indo-European laryngeals. http://spell.psychology.wustl.edu/~bkessler/.

Keydana, Götz. 2004. Silbenstruktur und Phonotaktik im Indogermanischen. In Maria Kozianka, Rosemarie Lühr, and Susanne Zeilfelder (eds.), *Indogermanistik – Germanistik – Linguistik, Akten der Arbeitstagung der Indogermanischen Gesellschaft, Jena 18.-20.09.2002*, 163–192. Hamburg: Dr. Kovač.

———. 2006. Die Indogermanische Perfektreduplikation. *Folia Linguistica Historica* 27.61–116.

———. Forthcoming a. Evidence for non-linear phonological structure in Indo-European: The case of fricative clusters. In *The Sound of Indo-European - selected papers from the conference, held in Copenhagen 16-19 April 2009.*

———. Forthcoming b. Hohe Sonorität in der Koda: Indo-Germanische Lösungen für ein phonetisches Problem. *Historische Sprachforschung*.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In I.-S. Yang (ed.), *Linguistics in the Morning Calm*, 3–91.

———. 1998. Sievers' Law as Prosodic Optimization. In Jay Jasanoff, H. Craig Melchert, and Lisi Oliver (eds.), *Mir Curad: Studies in Honor of Calvert Watkins*, 345–60. Innsbruck.

———. 2000. Opacity and Cyclicity. *Linguistic Review* 17.351–366.

Kisseberth, Charles W. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1.291–306.

Klingenschmitt, Gert. 1982. *Das altarmenische Verbum*. Wiesbaden: L. Reichert.

Kloekhorst, Alwin. 2008. *Etymological Dictionary of the Hittite Inherited Lexicon*. Leiden & Boston: Brill.

Kluge, Friedrich. 1891. Geschichte der gotischen Sprache. In Hermann Paul (ed.), *Grundriss der germanischen Philologie*, 497–517. Strassburg: Trübner.

Kobayashi, Masato. 2001. Syllable Rimes in Old Indo-Aryan and Dravidian. *Journal of Asian and African Studies* 62.91–106.

———. 2004. *Historical Phonology of Old Indo-Aryan Consonants*. Tokyo: Fujiwara.

Kortlandt, Frederik. 1982. IE **pt* in Slavic. *Folia Linguistica Historica* III/I.25–8.

Kümmel, Martin. 2000. *Das Perfekt im Indoiranischen. Eine Untersuchung der Form und Funktion einer ererbten Kategorie des Verbums in den altindoiranischen Sprachen*. Wiesbaden.

Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Oxford & Cambridge: Blackwell.

Ladefoged, Peter. 1967. *Three Areas of Experimental Phonetics*. Oxford: Oxford University Press.

Ladefoged, Peter and Ian Maddieson. 1990. Vowels of the World's Languages. *Journal of Phonetics* 18.93–122.

Leumann, Manu. 1977. *Lateinische Laut- und Formenlehre*, Volume 1 of *Lateinische Grammatik* by Manu Leumann, J. B. Hofmann, and Anton Szantyr. Munich: Beck.

Lindeman, Fredrik Otto. 1965. La loi de Sievers et le début du mot en indoeuropéen. *Norsk Tidsskrift for Sprogvidenskap* 20.38–108.

Lunden, Anya. 2006. *Weight, Final Lengthening and Stress: a Phonetic and Phonological Case Study of Norwegian*. PhD dissertation, University of California, Santa Cruz dissertation.

Malzahn, Melanie. 2009. *The Tocharian Verbal System*. Leiden & Boston: Brill.

Marchand, James W. 1958. Sievers' Law and a Rule of Indo-European Syllable Formation. *General Linguistics* III/2.73–84.

Martirosyan, Hrach K. 2010. *Etymological Dictionary of the Armenian Inherited Lexicon*. Leiden & Boston: Brill.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.), *Consonant Types & Tone. Southern California Occasional Papers in Linguistics No. 1*, 71–95. Los Angeles: Linguistics Program, University of Southern California.

Mayrhofer, Manfred. 1981. Laryngalreflexe im Indo-Iranischen. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 34.427–438.

———. 1985-2001. *EWAia: Etymologisches Wörterbuch des Altindoarischen*. Heidelberg: C. Winter.

———. 1986. *Indogermanische Grammatik: I-2 Lautlehre*. Heidelberg: Winter.

McCarthy, John J. 1994. The phonetics and phonology of Semitic pharyngeals. In Patricia Keating (ed.), *Papers in Laboratory Phonology III: Phonological Structure and Phonetic Form*, 191–233. Cambridge: Cambridge University Press.

———. 2000. Faithfulness and prosodic circumscription. In Joost Dekkers, Frank van der Leeuw, and Jeroen van de Weijer (eds.), *Optimality Theory : phonology, syntax, and acquisition*, 151–189. Oxford: Oxford University Press.

———. 2008. *Doing Optimality Theory: Applying Theory to Data*. Oxford: Blackwell.

McCarthy, John J. and Alan Prince. 1986. Prosodic morphology. Updated 1996.

McCrary, Kristie. 2004. *Reassessing the Role of the Syllable in Italian Phonology: An Experimental Study of Consonant Cluster Syllabification, Definite Article Allomorphy and Segment Duration*. PhD dissertation, University of California, Los Angeles.

Meier-Brügger, Michael. 2003. *Indo-European Linguistics*. Berlin & New York: Walter de Gruyter.

Meillet, Antoine. 1923. *Les origines indo-européennes des mètres grecs*. Paris: Presses Universitaires de France.

———. 1934. *Introduction á l'étude comparative des langues indo-européennes* (7th ed.). Paris: Hachette.

Meillet, Antoine and Joseph Vendryes. 1968. *Traité de grammaire comparée des langues classiques*. Paris: Champion.

Meiser, Gerhard. 1998. *Historische Laut- und Formenlehre der Lateinischen Sprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Melchert, H. Craig. 1977. Tocharian verb stems in *-tk-*. *Kuhns' Zeitschrift* 91.92–130.

———. 1987. Proto-Indo-European Velars in Luvian. In Calvert Watkins (ed.), *Studies in Memory of Warren Cowgill*, 182–204. Berlin & New York: de Gruyter.

———. 1989. New Luvo-Lycian isoglosses. *Historische Sprachforschung* 102.23–45.

———. 1994. *AHP = Anatolian Historical Phonology*. Amsterdam and Atlanta: Rodopi.

———. 1997. Luvian /tāna-/ 'sanctified, inviolable'. *Historische Sprachforschung* 110.47–51.

———. 2003. PIE “thorn” in Cuneiform Luvian? In Karlene Jones-Bley, Martin E. Huld, Angela Della Volpe, and Miriam Robbins Dexter (eds.), *Proceedings of the Fourteenth Annual UCLA Indo-European Conference*, 145–61. Washington, D.C.: Institute for the Study of Man.

Merlingen, Weriand. 1957. Idg. “þ” und Verwandtes. In Heinz Kronasser (ed.), ΜΝΗΜΗΣ ΧΑΡΙΝ. *Gedenkschrift Paul Kretschmer*, Volume 2, 49–61. Wiener Sprachgesellschaft: Vienna.

Mester, R. Armin. 1994. The Quantitative Trochee in Latin. *Natural Language and Linguistic Theory* 12.1–61.

Morelli, Frida. 1999. *The Phonotactics and Phonology of Obstruent Clusters in Optimality Theory*. PhD dissertation, University of Maryland.

Morgenstierne, Georg. 1929. Notes on Professor Charpentier’s article. *Acta Orientalia* 7.198–200.

Muller, Gerard. 1984. Notes on Four Old Armenian Etymologies. *Arbejdspapirer udsendt af Institut for Lingvistik, Københavns Universitet* 4.95–101.

Narten, Johanna. 1964. *Die sigmatischen Aoriste im Veda*. Wiesbaden: Harrassowitz.

Neri, Sergio. 2003. *I sostantivi in -u del Gotico. Morfologia e Preistoria*. Innsbruck.

Niepokuj, Mary. 1997. *The Development of Verbal Reduplication in Indo-European*. Washington, D.C.: Institute for the Study of Man.

Nishimura, Kanehiro. 2008. *Vowel Reduction and Deletion in Italic: Effects of Stress*. PhD dissertation, University of California, Los Angeles.

Nussbaum, Alan J. 1986. *Head and Horn in Indo-European*. Berlin & New York: de Gruyter.

———. 1997. The "Saussure Effect" in Latin and Italic. In Alexander Lubotsky (ed.), *Sound Law and Analogy: Papers in honor of Robert S.P. Beekes on the occasion of his 60th birthday*. Amsterdam & Atlanta, GA: Rodopi.

Oettinger, Norbert. 1979. *Die Stammbildung des hethitischen Verbums*. Nuremberg: Hans Carl.

Olsen, Birgit Anette. 1988. *The Proto-Indo-European Instrument Noun Suffix *-tlom and its Variants*. Historisk-filosofiske Meddelelser 55. Det Kongelige Danske Videnskabernes Selskab: Copenhagen.

———. 1999. *The Noun in Biblical Armenian*. Berlin & New York: Mouton de Gruyter.

Orel, Vladimir E. 1998. *Albanian etymological dictionary*. Leiden & Boston: Brill.

Osthoff, Hermann and Karl Brugmann. 1879. *Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen, Part II*. Leipzig: S. Hirzel.

Peters, Martin. 1980. *Untersuchungen zur Vertretung der indogermanischen Laryngale im Griechischen*. Vienna: Österreichische Akademie der Wissenschaften.

———. 1999. Ein tiefes Problem. In H. Eichner and H. C. Luschützky (eds.), *Compositiones Indogermanicae in Memoriam Jochem Schindler*, 447–456. Prague: Enigma.

Pierrehumbert, Janet. 1994. Syllable Structure and Word Structure. In *Papers in Laboratory Phonology III*, 168–188. Cambridge: Cambridge Univ. Press.

Pinault, Georges-Jean. 2005 [2006]. Analyse étymologique d'un Nom de Parenté Indo-Européen. In *Indogermanica: Festscrift Gert Klingenschmitt. Indische, Iranische und Indogermanische Studien dem verehrten Jubilar dargebracht zu seinem fünfundsechzigsten Gebürtstag*. Taimering: Schweiger-VWT-Verlag.

———. 2008. *Chrestomathie tokharienne: textes et grammaire*. Leuven: Peeters.

Pokorny, Julius. 2005. *IEW: Indogermanisches Etymologisches Wörterbuch* (5th ed.). Tübingen & Basil: Francke Verlag.

Praust, Karl. 2000. Altindisch *dr̥-/dr̥̄-*: seṭ oder aniṭ? In Bernhard Forssman and Robert Plath (eds.), *Indoarisch, Iranisch, und die Indogermanistik, Arbeitstagung der Indogermanischen Gesellschaft vom 2. bis 5. Oktober 1997 in Erlangen*, 425–441. Wiesbaden: Reichert Verlag.

Prince, Alan and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. MS, Rutgers University and University of Colorado, Boulder.

Pulgram, Ernst. 1970. *Syllable, Word, Nexus, Cursus*. The Hague & Paris: Mouton.

Rasmussen, Jens. 1999. *Selected Papers on Indo-European Linguistics, Part 2*. Copenhagen.

Rau, Jeremy. 2003 [2009]. Vedic Sanskrit *aśītí-* '80'. *Münchener Studien zur Sprachwissenschaft* 63.139–144.

Rehg, Kenneth L. and Damian G. Sohl. 1981. *Ponapean Reference Grammar*. Honolulu: University of Hawaii Press.

Reynolds, Elinor, Paula West, and John Coleman. 2000. Proto-Indo-European 'laryngeals' were vocalic. *Diachronica* 17.351–87.

Ringe, Donald. 1991. Laryngeals and Sievers' law in Tocharian. *Münchener Studien zur Sprachwissenschaft* 52.137–168.

———. 2006. *From Proto-Indo-European to Proto-Germanic*. Oxford & New York: Oxford University Press.

Rix, Helmut. 1992. *Historische Grammatik des Griechischen: Laut und Formenlehre*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Rix, Helmut, Martin Kümmel, Thomas Zehnder, Reiner Lipp, and Brigitte Schirmer. 2001. *LIV: Lexikon des Indogermanischen Verben* (2nd ed.). Wiesbaden: Dr. Ludwig Reichert Verlag.

Ruijgh, C.J. 1996. Chronologie relative: le grec. Sur les traitements préhistoriques des sonantes. In A. Rijksbaron and F.M.J. Waanders (eds.), *Scripta Minora ad linguam Graecam pertinentia II*, 353–77. Amsterdam & Philadelphia: John Benjamins Publishing Company. = *Rekonstruktion und Relative Chronologie*, Innsbruck 1992, 75-99.

Saussure, Ferdinand de. 1879. *Mémoire sur le système primitif des voyelles dans les langues indo-européenes*. Leipzig: Teubner.

———. 1885. Sur un point de la phonétique des consonnes en indo-européen. *Mémoires de la Société de Linguistique de Paris* 6.246–57.

Schindler, Jochem. 1972. *Das Wurzelnomen im Arischen und Griechischen*. PhD dissertation, Julius-Maximilians-Universität at Würzburg.

———. 1977a. Notizen zum Sieversschen Gesetz. *Die Sprache* 23.56–65.

———. 1977b. A Thorny Problem. *Die Sprache* 23.25–35.

Schmidt, Gernot. 1973. Die iranischen Wörter für "Tochter" und "Vater" und die Reflexe des interkonsonantischen *H (ə)* in den idg. Sprachen. *Kuhns' Zeitschrift* 87.36–83.

Schmidt, Johannes. 1895. *Kritik der Sonantentheorie*. Weimar: Hermann Böhlaus Nachfolger.

Schrijver, Peter. 1991. *The Reflexes of the Proto-Indo-European Laryngeals in Latin*. Amsterdam & Atlanta, GA: Rodopi.

———. 1995. *Studies In British Celtic Historical Phonology*. Amsterdam and Atlanta: Rodopi.

Schumacher, Stefan. 2000. *The Historical Morphology of the Welsh Verbal Noun*. Maynooth: Dept. of Old Irish, National University of Ireland, Maynooth.

Seebold, Elmar. 1972. *Das System der indogermanischen Halbvokale*. Heidelberg: Winter.

Serbat, Guy. 1975. *Les dérivés nominaux latins à suffixe mediatif*. Paris: Belles Lettres.

Sherer, Tim D. 1994. *Prosodic Phonotactics*. PhD dissertation, University of Massachusetts, Amherst.

Siebs, Theodor. 1904. Anlautstudien. *Kuhns' Zeitschrift* 37.277–324.

Sievers, Eduard. 1878. Zur Accent- und Lautlehre der germanischen Sprachen. *Beiträge zur Geschichte der deutschen Sprache und Literatur* 5.63–163.

———. 1881. *Grundzüge der Phonetik*. Leipzig: Breitkopf and Härtel.

Sihler, Andrew L. 1995. *New Comparative Grammar of Greek and Latin*. New York & Oxford: Oxford University Press.

———. 2006. *Edgerton's Law: The Phantom Evidence*. Heidelberg: Winter.

Southern, Mark R.V. 1999. *Sub-Grammatical Survival: Indo-European s-mobile and its Regeneration in Germanic*. Washington, D.C.: Institute for the Study of Man.

Steriade, Donca. 1982. *Greek Prosodies and the Nature of Syllabification*. PhD dissertation, MIT.

———. 1999. Alternatives to the syllabic interpretation of consonantal phonotactics. In Bohumil Palek, Osamu Fujimura, and Brian Joseph (eds.), *Proceedings of the 1998 Linguistics and Phonetics Conference*, 205–42. Prague: The Karolinum Press.

Struijke, Caro. 2002. *Existential Faithfulness: A Study of Reduplicative TETU, Feature Movement, and Dissimilation*. New York & London: Routledge.

Stüber, Karin. 1998. *The Historical Morphology of n-Stems in Celtic*. Maynooth: Dept. of Old Irish, National University of Ireland, Maynooth.

Szemerényi, Oswald. 1970. *Einführung in die Vergleichende Sprachwissenschaft*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Thurneysen, Rudolf. 1946. *A Grammar of Old Irish*. Translated by D.A. Binchy and Osborn Bergin. Dublin: DIAS; reprinted 1980.

Treiman, Rebecca and Catalina Davis. 1988. Syllabification of intervocalic consonants. *Journal of Memory and Language* 27.87–104.

Turner, Ralph Lilley. 1966. *A Comparative Dictionary of the Indo-Aryan Languages*. London: Oxford University Press.

Uffmann, Christian. 2006. Epenthetic vowel quality in loanwords: Empirical and formal issues. *Lingua* 116.1079–1111.

Untermann, Jürgen. 2000. *Wörterbuch des Oskisch-Umbrischen*. Heidelberg: C. Winter.

Vaux, Bert and Andrew Wolfe. 2009. The Appendix. In *Contemporary Views on Architecture and Representations in Phonology*, 101–143. Cambridge, MA & London: MIT Press.

Vennemann, Theo. 1972. On the Theory of Syllabic Phonology. *Linguistische Berichte* 18.1–18.

———. 1985. *Preference Laws for Syllable Structure and the Explanation of Sound Change, with Special Reference to German, Germanic, Italian and Latin*. Berlin / New York / Amsterdam : Mouton de Gruyter.

Vijūnas, Aurelijus. 2006. *The Indo-European Primary t-Stems*. PhD dissertation, University of California, Los Angeles dissertation.

Vine, Brent. 1999. On "Cowgill's Law" in Greek. In Heiner Eichner and Hans Christian Luschützky (eds.), *Compositiones Indogermanicae in memoriam Jochem Schindler*, 555–600. Prague: Enigma Corporation.

———. 2008. Back to the Future: 'Yesterday' (and 'Tomorrow') in Indo-European. *Indo-European Studies Bulletin* 12/2.13–23.

Walde, Alois and J.B. Hofmann. 1982. *Lateinisches etymologisches Wörterbuch* (5th ed.). Heidelberg: C. Winter.

Walker, Rachel. 2000. Nasal reduplication in Mbe affixation. *Phonology* 17/1.65–115.

Watkins, Calvert. 1982. A Greco-Hittite Etymology. In *Serta Indogermanica : Festschrift für Günter Neumann zum 60. Geburtstag*, 455–7. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.

———. 2000. *The American Heritage Dictionary of Indo-European Roots* (2nd ed.). Boston & New York: Houghton Mifflin Co.

———. 2008. "Hermit Crabs," or New Wine in Old Bottles: Anatolian-Hellenic Connections from Homer and Before to Antiochus I of Commagene and After. In Mary R. Bachvarova Billie Jean Collins and Ian C. Rutherford (eds.), *Anatolian Interfaces: Hittites, Greeks and Their Neighbours, Proceedings of an International Conference on Cross-Cultural Interaction, September 17-19, 2004, Emory University, Atlanta, GA*, 135–41. Oxford: Oxbow Books.

Watson, Janet C. E.. 2002. *The Phonology and Morphology of Arabic*. Oxford & New York: Oxford University Press.

Weiss, Michael. 1995. Life Everlasting: Latin *iūgis* "everflowing," Greek ὑγιής "healthy," Gothic *ajukdūs* "eternity," and Avestan *yauuaējī-* "living forever". *Münchener Studien zur Sprachwissenschaft* 55.131–156.

———. 2009. *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor & New York: Beech Stave Press.

Weitenberg, J.J.S. 1984. *Die hethitischen u-Stämme*. Amsterdam: Rodopi.

Winter, Werner. 1965. Tocharian evidence. In Werner Winter (ed.), *Evidence for Laryngeals*. London / The Hague / Paris: Mouton & Co.

Wodtko, Dagmar S., Britta Irslinger, and Carolin Schneider. 2008. *NIL: Nomina im Indogermanischen Lexikon*. Heidelberg: Winter.

Wolff, Ludwig. 1921. *Studien über die Dreikonsonanz in den germanischen Sprachen*. Berlin: Ebering.

Yamazaki, Yoko. 2009. The Saussure Effect in Lithuanian. *The Journal of Indo-European Studies* 37/3&4.430–61.

Yu, Alan C. L. 2004. Explaining Final Obstruent Voicing in Lezgian: Phonetics and History. *Language* 80/1.73–97.

Zair, Nicholas. 2009. A New Environment for Laryngeal Loss in Proto-Celtic. Paper presented at the Sound of Indo-European - phonetics, phonology and morphophonology, University of Copenhagen, 16-19 April 2009.

Zec, Draga. 1995. Sonority Constraints on Syllable Structure. *Phonology* 12/1.85–129.

Zuraw, Kie. 2007. The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language* 83/2.277–316.