Check for updates

RESEARCH ARTICLE

REVISED The rise of gemination in Celtic

[version 2; peer review: 4 approved]

David Stifter

Early Irish, Maynooth University, Maynooth, Kildare, W23 F2H6, Ireland

v2 **First published:** 02 Feb 2023, **3**:24
https://doi.org/10.12688/openreseurope.15400.1
Latest published: 08 Feb 2024, **3**:24
https://doi.org/10.12688/openreseurope.15400.2

Abstract

This study investigates systematically the emergence and establishment of geminate consonants as a phonological class in the Celtic branch of Indo-European. The approach of this study is comparative historical linguistics, drawing on diachronic structuralism combined with aspects of language contact studies and functional approaches to language usage. This study traces the development of geminates from Proto-Indo-European (fourth millennium B.C.), which did not allow geminate consonants, to the Common Celtic period (first millennium B.C.), when almost every consonant could occur as a singleton or as a geminate, and on to the earliest attested stages of the Insular Celtic languages (first millennium A.D.). Although they were prominent in the phonology of Proto- and Ancient Celtic (Gaulish, Celtiberian), ultimately geminates were gotten rid of as a phonological class in the individual Insular Celtic languages. This is probably due to the fact that the contrast between lenited and unlenited sounds took on a central role in Insular Celtic phonology, making gemination a phonetically redundant category.

Most instances of geminate consonants in Celtic can be explained by regular sound change operating on inherited clusters of consonants. Each sound change will be discussed in a separate section in a rough chronological order. Effectively, gemination is largely a strategy to reduce the number of allowed consonant combinations. To a limited degree, gemination also had a morphological function, especially in the formation of personal names and in the creation of adjectival neologisms. However, there is a residue of words, especially nouns, in the Insular Celtic languages that defy any attempt at etymologising. They are prime suspects of having been borrowed from prehistoric, substratal languages.

Plain language summary

Geminate, *i.e.,* 'double' or 'long', consonants were very common in Proto- and Ancient Celtic languages, such as Gaulish or Celtiberian of

Open Peer Review

Approval Status ✓ ✓ ✓ ✓

	1	2	3	4
version 2 (revision) 08 Feb 2024			✓ view	
version 1 02 Feb 2023	✓ view	✓ view	✓ view	✓ view

1. **Martin Joachim Kümmel**, Friedrich-Schiller-Universität Jena, Jena, Germany
2. **Karin Stüber**, University of Zurich, Zurich, Switzerland
3. **Anders Jørgensen**, Uppsala University, Uppsala, Sweden
4. **Elisa Roma**, Università degli Studi di Pavia, Pavia, Italy
   Gianguido Manzelli, Universita degli Studi di Pavia, Pavia, Italy

Any reports and responses or comments on the article can be found at the end of the article.

the first millennium B.C. and earlier. They were also very prominent in the prehistory of the Insular Celtic languages, *e.g.* Irish, Welsh or Breton, but they were abandoned as a class of sounds shortly before the attestation of those languages due to other developments in those languages, especially the rise of lenited sounds as a grammatically very important class. This important role of geminates in Celtic contrasts with the situation in its ancestor language, reconstructed Proto-Indo-European (ca. middle fourth millennium B.C.), which effectively disallowed geminate consonants. This article explains how geminate consonants arose step by step in the prehistory and the early history of Celtic, mostly by regular sound change operating on inherited words. In addition, gemination became prominent in the formation of personal names and in the creation of new adjectives. However, a group of nouns with geminates finds no explanation within the traditional framework of historical linguistics. It is suggested that they are due to borrowings from prehistoric, lost languages in the west of Europe.

Keywords
Indo-European linguistics, Celtic linguistics, historical linguistcs, geminate consonants, substrate linguistics

This article is included in the Languages and Literature gateway.

This article is included in the Horizon 2020 gateway.

This article is included in the European Research Council (ERC) gateway.

This article is included in the Linguistic Diversity collection.

Corresponding author: David Stifter (david.stifter@mu.ie)

Author roles: Stifter D: Conceptualization, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme(s) (grant agreement No(s) 647351 & 716732).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2024 Stifter D. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Stifter D. **The rise of gemination in Celtic [version 2; peer review: 4 approved]** Open Research Europe 2024, **3**:24 https://doi.org/10.12688/openreseurope.15400.2

First published: 02 Feb 2023, **3**:24 https://doi.org/10.12688/openreseurope.15400.1

REVISED **Amendments from Version 1**

In the revised version of this article, I have taken on board the extensive feedback received from the four very helpful peer reviews. I have applied almost all recommendations, except for a very small number where my opinion diverges strongly from that of the reviewers.

The revision has changed nothing in the overall structure, argument and conclusions of the article. To the largest extent, the changes have added precision and clarity to my arguments, and additional material to support my proposals. Furthermore, around ten new lexical items have been added to the collection. For Cornish and Breton, the cited material has been verified with more specialised sources (George, 2020; Graves, 1962; Menard, 2016; Nance, 1938).

Otherwise, what can be regarded as susbtantial changes are of an outwardly subtle nature. Several changes concern the phonological systems of the various stages of the Celtic languages. Following suggestions by Martin Kümmel, the sound *$\underset{\smile}{u}$ (= [w]), which traditionally is classified as a bilabial, has been grouped with labiovelars in those stages of the languages, in which labiovelars form a distinctive class. 'Tau Gallicum' is now treated as a bimoraic element *ts, not as monomoraic *t^s. And, in a very radical step, I have added *pp, i.e. geminate *p, as a possibly separate phoneme for the early stages of the Celtic languages, a phoneme that was potentially not affected by the loss of single Indo-European *p in Celtic. As regards geminate voiced stops in the British Celtic languages, I now consider the possibility more strongly that their differential treatment reflects different chronological stages and different morphological contexts. Finally, while I was sceptical of the change *rl > *ll in the first version of the article, I now view it more positively (section 3.7.).

Any further responses from the reviewers can be found at the end of the article

Abbreviations

acc.	accusative
Bret.	Breton
Cib.	Celtiberian, c. 150–0 B.C.
CIIC	*Corpus Inscriptionum Insularum Celticarum, i.e.* Macalister, 1945
CisGaul.	Cisalpine Gaulish, second–first centuries B.C.
Corn.	Cornish
dat.	dative
eDIL	*electronic Dictionary of the Irish Language, i.e.* Toner *et al.*, 2019
fem.	feminine
Fr.	French
Gall.-Lat.	Gallo-Latin (Latin spoken in Gaul)
Gaul.	Gaulish, second century B.C.–c. fourth century A.D.
gen.	genitive
Germ.	German
GOI	*A Grammar of Old Irish,* i.e. Thurneysen, 1946
GPC	*Geiriadur Prifysgol Cymru* (Dictionary of the University of Wales)
Gr.	Ancient Greek
Hisp.-Lat.	Hispanic Latin, Latin in the Iberian Peninsula
IEW	*Indogermanisches Etymologisches Wörterbuch, i.e.* Pokorny, 1959
Ital.	Italian
Lat.	Latin
Latv.	Latvian
LCorn.	Late Cornish, early eighteenth century
LEIA	*Lexique Étymologique de l'Irlandais Ancien,* i.e. Vendryes *et al.*, 1959–
LGaul.	Late Gaulish, c. second–fourth centuries A.D.
Lith.	Lithuanian
LIV	*Lexikon der Indogermanischen Verben*^{2}, i.e. Kümmel & Rix, 2001
masc.	masculine
MBret.	Middle Breton, c. 1100–1650
MCorn.	Middle Cornish, c. 1200–1600
MHG	Middle High German, c. 1050–1500
MIr.	Middle Irish, c. 900–1200
MLat.	Middle Latin, Latin in the Middle Ages
MLG	Middle Low German, c. 1200–1600
ModIr.	Modern Irish, c. 1200–present
MW	Middle Welsh, c. 1100–1400
NBret.	New (or Modern) Breton, c. 1650–present
neut.	neuter
NIL	*Nomina im Indogermanischen Lexikon*, i.e. Wodtko, Irslinger & Schneider, 2008
nom.	nominative
OBret.	Old Breton, c. 800–1100
OCorn.	Old Cornish, c. 800–1200
OCS	Old Church Slavonic, c. 860–1100
OEngl.	Old English, c. mid-seventh–eleventh centuries
OFr.	Old French, c. ninth–fourteenth centuries
OHG	Old High German, c. 750–1050
OIr.	Old Irish, c. 700–900

ON	Old Norse, c. 800–1350
OW	Old Welsh, c. 800–1100
PC	Proto-Celtic, c. late second millennium B.C.
Pict.	Pictish
PIE	Proto-Indo-European, c. middle of the fourth millennium B.C.
pl.	plural
pres.	present
PrimIr.	Primitive Irish, c. fourth–sixth centuries A.D.
qPC	quasi-Proto-Celtic
qPIE	quasi-Proto-Indo-European
RIIG	*Recueil informatisé des inscriptions gauloises, i.e.* Ruiz Darasse et al., 2022
Sc. Gael.	Scottish Gaelic
sg.	singular
subj.	subjunctive
VKG	*Vergleichende keltische Grammatik, i.e.* Pedersen, 1909–1913
VLat.	Vulgar Latin
VN	verbal noun
voc.	vocative
W	Welsh (any stage of the language, but especially the modern variant after 1400)
1/2/3sg./pl.	first, second, third person singular or plural

Preliminaries

This study investigates systematically the emergence and establishment of geminate consonants as a phonological class in the Celtic branch of Indo-European[1]. Geminates were absent from Proto-Indo-European, so their presence as a phonological class in its descendant languages such as Celtic requires an explanation. Geminates may have arisen either through regular phonological changes and/or through other, non-regular processes. *Mutatis mutandis*, the same holds true for other branches of Indo-European that developed geminate sounds, for which the methodological approach of this article can serve as a model. The aim of this article is to come closer to a general theory of this phenomenon in Celtic, in order to permit inferences about the etymology of words that contain geminate consonants in the attested Celtic languages. Geminate sounds have sometimes been used as an argument for identifying borrowings from unknown substrate languages. However, before their precise diachronic and synchronic status within the phonological system of Proto-Celtic and the individual languages hasn't been determined, conclusions about layers of loanwords based on geminates are circular.

The approach of this study is diachronic structuralism combined with aspects of language contact studies and of functional approaches to language usage. The working hypothesis is that the inexorable, albeit gradual, rise of geminate consonants as a phonological class across all modes of articulation in the older stages of Celtic was a multicausal process that was fed by a combination of language-internal developments in Celtic and of language-external factors. The focus of this article is on the period up to the early part of the 1st millennium A.D. Evidently younger sound changes that occurred in the documented histories of Irish or British and that created further instances of geminates or, at that stage, fortes sounds (for example, assimilations that postdate syncope such as *nl* > *ll* in OIr. (Old Irish) *tenlach* > *tellach* 'hearth', *ld* > *ll* in *acaldam* > *acallam* 'conversation, dialogue', or MacNeill's Law in Irish, or provection after syncope in British, or other patently late sound changes) are not treated here. Reference to them will only be made when they are relevant to clarify points in the prehistoric developments.

The terminology in this article follows the traditional practice in Celtic historical phonology. The term 'geminate consonant' will be used synonymously with 'long consonant'. In accordance with the traditional practice, geminate consonants are written double in reconstructions, *e.g.*, **ballo-*, which is meant as equivalent to a phonetic analysis as [balːo]-. The main contrast between the two classes of Celtic stops is considered to be between 'voiced' (= D) and 'unvoiced/voiceless' (= T) consonants. Phonetically, the contrast between the D-series and the T-series may rather have been that between 'lenis' voiceless consonants and 'fortis' aspirated voiceless consonants (cf. Stifter, 2017: 1191; similar Van Sluis, 2019: 3–36; sceptical Martinet, 1952: 201) or, in Eska's (2018) framework of Laryngeal Realism, a contrast in the feature [spread glottis]. Since the sound changes in this article are understood as arithmetic abstractions, notations of reconstructions can usually be transferred easily between alternative frames of references. The phonetic details of these variant descriptions do not make a practical difference for the question at hand. What is curcial, however, is that the contrast between those two classes of Celtic stops does not imply a concomitant phonological contrast in length. In addition, it is assumed that intervocalic voiced stops developed non-contrastive, lenited, *i.e.*, fricative allophones already at a very early stage (cf. Schrijver, 2016: 497–499; Stifter, 2017: 1189–1190). These fricative allophones are not indicated in the reconstructions.

[1] Part of this study was originally written as a digression within a very different article (Hayden & Stifter, 2022), but ultimately went beyond the scope of that article and was therefore removed from it. The arguments in some sections have profited substantially from comments received from anonymous reviewers of that original article. I thank Joseph Eska, Corinna Salomon, Michael Weiss and in particular Paulus van Sluis for invaluable support, advice and suggestions. My special gratitude extends to the reviewers and readers who have commented on the first published version of this article, Anders Jørgensen, Martin Kümmel, Gianguido Manzelli, Elisa Roma, and Karin Stüber, and to Stefan Höfler, Lionel Joseph, Guto Rhys, and Stefan Schaffner. All disclaimers apply.

The label 'PIE' (Proto-Indo-European) will be reserved for reconstructions that can be securely set up for the protolanguage, while 'qPIE' (quasi-Proto-Indo-European) is an umbrella label for any *voreinzelsprachlich*, pre-Celtic reconstruction that may have arisen in the long period between the break-up of Proto-Indo-European and the emergence of Proto-Celtic (PC), for instance during an extended Western Indo-European period. qPIE reconstructions will be given in standard PIE phonology, even where this may be anachronistic. *Mutatis mutandis*, similar considerations apply to the label 'qPC' (quasi-Proto-Celtic). The label 'quasi-' is intentionally used in a very vague sense and fulfils mainly a heuristic, not a chronological function.

Examples are cited from the standard lexicographic collections. As a rule, no special reference to these sources is made and etymologies are only discussed in cases where it is necessary. This study aims to be a representative, not an exhaustive collection of relevant words in the Celtic languages; systematic searches, especially in the Insular Celtic languages, would doubtless result in more examples. The reference points for Proto-Indo-European are *Lexikon der Indogermanischen Verben* (LIV = Kümmel & Rix, 2001), *Nomina im Indogermanischen Lexikon* (NIL = Wodtko, Irslinger & Schneider, 2008) and *Indogermanisches Etymologisches Wörterbuch* (IEW = Pokorny, 1959). For Proto-Celtic, the standard handbooks are Matasović's *Etymological Dictionary of Proto-Celtic* (2009) and Schumacher's (2004) lexicon of primary verbs, *Die keltischen Primärverben*. Old Irish words are cited from the *electronic Dictionary of the Irish Language* (eDIL = Toner *et al*., 2019; albeit with occasional spelling normalisation to an idealised Old Irish standard; for criticism of the inconsistent spelling of headwords in eDIL see Griffith *et al*., 2018: 7), Welsh from *Geiriadur Prifysgol Cymru* (GPC Online), Old Cornish from Graves (1962), Cornish of later periods from Nance (1938) and George (2020), and Breton from Favereau (2016) and *Devri* (Menard, 2016). Gaulish data is taken from Delamarre (2003) and from Delamarre (2007) for personal names; Celtiberian data from Jordán Cólera (2019). Lepontic and Cisalpine Celtic examples are taken from *Lexicon Leponticum* (Stifter *et al*., 2009). Ancient Celtic personal names are cited either as abstract stems or in an idealised nominative singular, even if that form is unattested. Since this article is concerned with the predesinential parts of words, this procedure has no consequence for the main argument. Due to the nature of the sources, the semantics of ancient Celtic cognates is occasionally uncertain. The position adopted here is that the choice of ancient cognates is guided by plausibility, *i.e.*, by their formal phonological and morphological correspondence with words in the younger Insular Celtic languages. If new evidence should emerge that casts doubt on these equations, the items would have to be removed.

This study consists of two big conceptual parts. Chapters 2–8 are devoted to etymological gemination, *i.e.*, types of gemination that arose by regular sound change operating on inherited words. The sections within this first part are arranged in an approximate, but not strict, chronological order. The principle of thematic coherence occasionally overrides chronology. For instance, morphological processes such as gemination arising from inflection, derivation and compounding are treated in single sections, even if they may relate to extended periods of time. The second part of the study is concerned with non-etymological gemination. 'Non-etymological' means that such types or instances of gemination cannot be described by regular sound laws, but they arose non-predictably from pragmatic contexts such as addressing, through sound-symbolic neologisms, or are due to borrowing.

Previous research

The first dedicated study of geminates in a Celtic language was in response to an observation in Germanic linguistics, namely that stops before **n* seem to have become geminates in Proto-Germanic. This is known as 'Kluge's Law' (proposed by Kluge, 1884; for a modern take, see Kroonen, 2013: xxxiv–xxxv). Stokes (1891; revised in 1891–3) took his inspiration from this recently postulated Germanic rule and tried to apply it to Old Irish. It is evident from Stokes' discussion that he had no clear understanding of the phonetic reality of Old Irish double consonant spellings such as *cc*, and many of his explanations no longer stand up to a modern understanding of Celtic historical phonology. From an early date, the notion of the operation in Celtic of a sound change comparable to Kluge's Law was met with scepticism (*e.g.*, GOI (*Grammar of Old Irish* = Thurneysen, 1946) 92–93; Sjoestedt, 1926: 19–20; Martinet, 1952: 197–198).

Zupitza (1900) meant progress over Stokes' attempts insofar as he also recognised other sources for Irish geminates. Building on a very thorough study of spellings in the major Old and Middle Irish sources, he rightly identified many instances of geminates as the products of assimilation across morphemic boundaries, but he also saw a major source of geminate stops in clusters of stops + **n*. Like Stokes', many of his etymologies have been rejected in the meantime. In a similar vein, Pedersen (VKG (*Vergleichende keltische Grammatik*) i 158–161, 476–477) distinguishes between double consonants arising from assimilation in word formation, while also being inclined to operate with a Celtic version of Kluge's Law. In addition, he recognises doubling of consonants in terms of endearment.

Sommerfelt (1954) considered substratal influence responsible for establishing the contrast between short and long consonants in Irish, as well as that between non-palatalised and palatalised consonants. Kuryłowicz (1957) adopted a very different strategy in order to explain those instances of gemination in Celtic and Germanic that are not caused by assimilation. Critical of the notion of Kluge's Law, he saw a different, single principle behind the phenomenon. Starting from a core where pairs of simple *vs*. geminate consonants were the product of regular sound change (*l* : *ll*, *n* : *nn*), Kuryłowicz thinks that this opposition was extended as a morphological marker to other sounds as well.

New life was breathed into the hypothesis of the geminating force of **n* by Lühr (1985) who adduced further instances of geminate stops as evidence for that rule. She had to explain the numerous counterexamples through analogy (Lühr, 1985: 345). In addition, Lühr's hypothesis allows for geminates that are due to expressivity. The idea of a Celtic version of Kluge's Law

was also advocated by Bammesberger (1998). Like a century before, most scholars remained sceptical towards the usefulness of such a law in Celtic.

De Bernardo Stempel (1999: 508–522) provides a useful list of Old Irish words with gemination. She distinguishes between inherited geminates, *i.e.*, those geminates that arose through regular sound change, and other cases, which she largely ascribes to expressiveness. Her discussion also includes a detailed refutation of Lühr (1985). In a 2010 article, De Bernardo Stempel turns her attention to geminates in the ancient Continental Celtic languages, in particular to the possibility of gemination resulting from accentuation. Like in the case of Old Irish, she distinguishes between etymological and non-etymological gemination. This useful dichotomy between etymological and non-etymological gemination will inform the present article, too. McCone (2005: 406–407) touches briefly on the subject of geminates in the context of substratal influence on Insular Celtic. His remarks are specifically directed at word-initial geminates and their role in the emergence of initial mutations. He maintains that, except perhaps for *rr*, *ll*, *nn*, the status of geminated initial consonants in the historically attested Insular Celtic languages is dubious, and he downplays their significance as an indicator of substrate influence.

The status of gemination in the phonology and phonetics of the Insular Celtic languages and the part it played, or did not play, in establishing the system of initial mutations, especially in British Celtic, was the topic of a heated debate that extended for more than five decades, from the middle of the 20th until the early 21st century (see the more detailed summary in Van Sluis, 2019: 15–24). The debate grew out of an exchange between Jackson (1953: 473–480, 545–548, 565–573, 634–638; 1960; 1967: 307–308, 317–323) and Greene (1956; 1966); cf. also Kortlandt (1982), Feuth (1983), Koch (1989). A satisfactory solution to the problem was proposed by Harvey (1984) and Russell (1985) that removed the notion of gemination as a genuine mutation in the history of Irish and British. Afterwards, the debate shifted away from gemination towards the related question of the relationship between the spirant mutation and nasalisation in the British Celtic languages, with notable contributions by Thomas (1990), Sims-Williams (1990, 2008), McCone (1996; 92–96), Schrijver (1999), and Isaac (2004; 2008).

Although it can be seen from this brief historical sketch that many scholars have contributed to the study of Celtic gemination or to aspects of it in the past, a comprehensive modern account of the phenomenon in a diachronic perspective is lacking. The following study aims at filling this gap, by conducting a systematic analysis of all the phonological sources and internal and external contexts in which geminate sounds arose in the early history of the Celtic languages.

1. Proto-Indo-European and Pre-Celtic

Celtic geminate stops cannot have been inherited from Proto-Indo-European (PIE), since the Indo-European protolanguage famously did not have geminate consonants as a phonological class. This prohibition went so far that even accidental geminates across the word boundary were prone to simplification. It is believed that the so-called 'mobile *s*' of Indo-European arose in that way (Mayrhofer, 1986: 120–121). However, this is not the whole story, because geminate sounds may not have been entirely absent from some registers. For a small number of strongly affective words an argument can or could be made that they were articulated with geminates already in the protolanguage. In any case, the continuations of these words in the Celtic languages, again used in strongly affective contexts, display gemination.

(1) PC **atta* 'daddy (or some other older, male relative)' probably underlies numerous Gaulish names such as *Atta*, *Attū*, *Attios etc*. How OIr. *aite* 'fosterfather' belongs here, is unclear. Its modern continuation, ModIr. *oide* 'fosterfather, tutor', has a /d/, not a /t/, as would be expected from PC **tt*. If this item can be projected back to Proto-Indo-European at all is uncertain. Martin Kümmel (in peer-review) makes the valid point that some of the parallels with **tt* in other Indo-European languages are either only superficial (*e.g.*, Hittite *atta-*) or could have arisen language-internally (*e.g.*, Germanic from **at-n-* via Kluge's Law).

(2) PC **kakko-* 'shit, excrement' > MIr. *cacc*, W *cach*, Corn. *caugh*, MBret. *cauch*, NBret. *kac'h*. It is rather unlikely that various Gaulish names beginning with *Cac-* and *Cacc-* contain this etymon.

(3) PC **mamma* 'mum' can underlie W *mam*, MBret. *mam*, OCorn. *mam* 'mother', and perhaps OIr. *muimme* 'fostermother, nurse' and Gaulish names such as *Mammios*.

After some far-reaching simplifications of the phoneme inventory, especially the merger of the Indo-European velar and palatal series of stops, the merger of voiced and voiced aspirate stops (traditionally called mediae and mediae aspiratae; Schrijver, 2016: 497–499; Stifter, 2017: 1189–1190) and the loss of laryngeals, the Pre-Celtic system of consonants was reduced to a relatively simple 15-phoneme inventory that probably did not yet distinguish geminates as a phonological category. All sound changes discussed in the following chapters operate on the basis of this post-PIE, pre-Proto-Celtic sound inventory in Table 1.

Table 1. the pre-Proto-Celtic sound system.

	stop	nasal	fricative	glide	liquid
bilabial	p b	m			
dental	t d	n			l r
alveolar			s (z)		
palatal				j	
velar	k g				
labiovelar	k^w g^w			w	

A major Proto-Celtic change was the fricativisation of *p > *φ, which then underwent various further changes in Celtic, depending on phonological context. Marginal phonemes, *i.e.*, phonemes that occur only in very restricted contexts and which are allophones of core sounds of the system, are put in brackets here and in later sections. In keeping with the traditional practice in comparative Indo-European and Celtic studies, [j] and [w] will be represented by the symbols *i̯* and *u̯* in the main part of this article. In a deviation from the traditional practice, I follow a suggestion by Martin Kümmel (personal communication) and classify [w], which phonetically is a labial-velar glide, as a labiovelar sound for language stages that possess labialised velars.

From their virtual lack in the ancestral stages of the language, it follows that geminates must have been acquired as a phonological class after the break-up of the Proto-Indo-European parent language. It will become clear from this study that the emergence of geminates occurred gradually. There are around twenty regular, phonological sources for geminate consonants of Proto- and Common Celtic origin, plus one that is exclusive to Irish. In addition, a number of non-phonological sources (sound symbolism, loans) will be identified.

2. Etymological gemination: Proto-Celtic and Common Celtic

The phonetic class among which geminates started to arise earliest in Celtic are resonants. The starting point are several inherited consonantal clusters of common occurrence. They provided the input for assimilatory processes that created long, *i.e.*, geminate resonants either already in Proto-Celtic or so early after the split into individual branches that the changes spread across the entire speech community, leading to identical outcomes everywhere.

2.1. *ln > *ll

The assimilation of *ln > *ll is carried through in all attested Celtic languages and can therefore be securely assigned to Proto-Celtic, bolstered by a large number of convincing etymologies. The input for this change is predominantly morphological, *i.e.*, nominal or verbal derivatives where suffixes with *n-* attached to roots ending in *-l*.

Note that, if the rule formulated by Hill (2012), whereby *$Cl̥n$ resulted in PC *$Clin$, is correct, several of the explanations offered below cannot be upheld as formulated here (especially items 2, 4, 5, 7, 11–14). Either alternative explanations have to be sought, or intra-paradigmatic influence from forms with a vowel before the *l has to be assumed, i.e. from forms where *$CVln$ developed regularly to *$CVll$.

(1) qPIE *h_1elneh_2- > *$ellā$- > OIr. *ell* ‘herd, flock’.

(2) qPIE *$h_2ebl̥neh_2$- > *$aballā$ > Gaul. *aballo-*, *auallo-*, OIr. *aball*, W *afall(en)*, OBret. *aballenn*, Bret. *avalenn* ‘apple-tree’. The vocalisation can be influenced by *$abal$ ‘apple’ < *$h_2ebl̥$; cf. W *afal*.

(3) PIE *h_2elno- or *$alno$- (?) > PC *$allo$- > Cib. *Allus*, *All[on]is* (?), Gaul., W *all-* ‘other, different’, perhaps W *arall*, OIr. neut. *alall* ‘other’ (cf. Dunkel, 2014: 19 fn. 7, 22 fn. 9).

(4) PIE *$b^hl̥no$- ‘inflated thing’ > PC *$ballo$- > Gaul. *ballo-*, OIr. *ball* ‘member’.

(5) The reconstruction PIE *$d^hu̯l̥no$- > PC *$du̯allo$- > Gaul. *dallo-*, OIr., W, Corn., Bret. *dall* ‘blind’ is somewhat doubtful since PC *$du̯$- would be expected to be retained in Gaulish.

(6) qPIE *k^uelno- ‘distant’ > PC *k^uello- > W, Corn., Bret. *pell* ‘far’. The reconstruction *k^uelso- is also possible.

(7) qPIE *$ml̥no$- ‘hesitant’ (?) > *$mallo$- > Gaul. *mallo-*, OIr. *mall* ‘slow’, W *mall* ‘evil, rotten, bad’. The reconstruction *$ml̥so$- is also possible.

(8) qPIE *$polno$- ‘full’ > *$\varphi olno$- or *$(h_2{}^{\text{ʔ}})olno$- > *$ollo$- > Gaul. *ollon* ‘big?; whole?’, OIr. *oll* ‘great, ample’, W *oll*, *holl*, Corn., Bret. *oll* ‘all’ (cf. Dunkel, 2014: 19, 593).

(9) qPIE *sh_2lno- ‘salted’ > *$sallo$- > OIr. *sall* ($\bar{a}$, f) ‘bacon < *salted meat’, Bret. *sall* ‘salted’. Alternatively, the reconstruction *sh_2ldeh_2- is also conceivable (see 3.8. (3)). The Old Irish verb *saillid* ‘to salt’, on the other hand, is synchronically derived from the noun *salann* ‘salt’, namely *$saleu̯n$-$\bar{\imath}$- > *$sal^{j\,\dagger}n^j\bar{\imath}$- > *$sailnid$ > *saillid* (the dagger † indicates a syllable that has undergone syncope in the prehistory of Irish). Its geminate is due to the later, Old Irish assimilation of *ln > *ll*. In contrast, MBret. *sallaff* ‘to salt’ is probably directly derived from *sall* ‘salted’.

The change *ln > *ll is also witnessed by a group of Insular Celtic verbs that are based on old nasal-infix formations (cf. McCone, 1991: 11–23). They usually lack parallels in the Continental Celtic languages, but because of their archaic morphology they can nevertheless be securely projected back to Proto-Indo-European or at least to Proto-Celtic:

(10) PIE *$d^hh_2l̥nh_1$- or *d^halnh_1- ‘to emanate, sprout’ > *$daln$- → *$dallu̯e$- > MW *deilliaw* ‘to emanate, come out’ (Schumacher, 2004: 257–259).

(11) PIE *$gl̥$-*ne-H*- ‘to get power’ > *$galnV$- > *$gallV$- > W *gallu*, Corn. *gallos*, Bret. *gallout* ‘to be able, have power’. The vocalism is probably influenced by *$galā$- ‘valour, fury; steam’ < *$gl̥H$-*eh*$_2$-.

(12) PIE *$g^ul̥neh_1$- ‘to throw’ > *$balni$- > *$balli$- > OIr. *at·baill* ‘to die’.

(13) PIE *$k^ul̥$-*ne-h*$_1$- ‘to turn’ >> *$k^ue/alnV$- ‘to go in a circle’ > *$k^ue/allV$- > OIr. *·cella* ‘to turn, go in a circle’, W *pallu* ‘to fail, weaken, cease, refuse, deny’.

(14) PIE *$pl̥$-*ne-h*$_2$- ‘to approach’ >> *$\varphi e/alnā$- > *$\varphi e/allā$- > OIr. *ad·ella* ‘to visit’.

(15) PIE *$u̯elnH$-/*$u̯l̥nH$- ‘to be powerful’ > *$u̯ella$-/*$u̯alla$- > OIr. *follaithir* ‘to rule’ and *follnaithir* with re-insertion of the nasal. The verbal root *$u̯aln$- can also be assumed indirectly for Gaulish, since it underlies the widely attested nominal formation *$u̯ellamno$-/*$u̯allamno$- ‘ruler’ found in the onomastic element Gaul. *uallaunos*, OIr. *Folloman*, W *Cad-wallon*. The vocalism is probably influenced by *$u̯alo$- ‘ruler’ < *$u̯l̥H$-*o*-.

Other possible examples are uncertain.

(16) Delamarre (2003: 99; but cf. IEW 524) suggests that Gaul. *callio-* 'hoof' and W *caill*, Bret. *kell* 'testicle' continue **kalno-*, the Indo-European preform of which, perhaps **kHl-no-*, is uncertain.

(17) OIr. *cellach* 'strife, contention' and *Sucellus*, the Gaulish name of a hammer-wielding god, perhaps 'good-striker', can be combined under the reconstruction **kel-n-(h₂)o-*, a nasal-infix formation from the root **kelh₂-* 'to strike'. However, it is far from certain that they belong together, and alternative individual explanations exist (Matasović, 2009: 199; Zair, 2012: 183, 204).

(18) If OIr. *fillid* 'to return' continues the Indo-European root **u̯el-* 'to turn, roll', the starting point is probably a nasal extension **u̯el-n-*.

Finally, two important morphological processes in ancient Celtic languages must be discussed separately. One of them may have arisen, and the other certainly arose as a consequence of the assimilation **ln* > *ll*.

In Celtiberian, this sound change resulted in a productive inflectional and derivational pattern (Jordán Cólera, 2019: 611–614). The Celtic suffix *-on-*, particularly common with an individualising function among personal names, inflected originally with full paradigmatic ablaut, *i.e.*, lengthened grade **-ū* < **-ō(n)* in the nominative and full grade **-on-* and zero grade **-n-* in other cases. In the majority of cases, ablaut was eradicated in Celtiberian by generalising the lengthened grade across the oblique cases as well. However, the original state of affairs is still preserved in words with roots or presuffixal stems ending in *l*. The Celtiberian script does not write geminate consonants, but Motta (1981) convincingly argued that written nom. *abulu*, gen. *abulos* or nom. *statulu*, gen. *statulos* hide phonological /abulū abullos/ and /statulū statullos/. The nominative continues PC **°lū* < PIE **°lō(n)*, while the genitive is the regular outcome of the zero-grade in the oblique stem, namely **°lnos*. Other Celtiberian words may also hide original **ln* behind a written single *l*. This has, for instance, been proposed for *kelaunikui*, which Jordán Cólera (2019: 690) explains as **kelnH-mno-*.

On a more speculative note, the same process can be invoked to explain the emergence of the common hypocoristic suffix **-llo-* for names in the ancient Celtic languages. I illustrate my ideas with Gaulish, but the basic steps are valid generally. It is well known that the morphology of shortened names often does not observe meaningful morpheme boundaries (Schmitt, 1995: 424), but that dithematic compound names can be truncated in the middle of the second element, irrespective of the meaning and the transparency of the formation. Stüber *et al.* (2009: 256) cite *Adnema* as a shortened form of a compound with **nemes-* 'sky' or **nemeto-* 'sanctuary' as second element, and *Verca* as a shortening of a compound of **u̯er-* 'on, upon' and a second element starting with **k-*. At the same time, shortened names often appear as *on*-stems (see Stüber, 2004 for that suffix in general). For example, the compound name *Boudilatis* (perhaps 'having the fury/heat of booty/victory') consists of the two lexemes **bou̯di-* 'victory, booty' and **lāti-* 'warrior fury'. This could be truncated to the short name **bou̯dilo-* (not attested), which could in turn be 'individualised' as **bou̯dil-on-*. The suffix, which originally was fully ablauting, would have led to a paradigm nom. sg. **bou̯dilū*, acc. sg. **bou̯dilonam* (perhaps attested as *Bodilo*) with full and lengthened grades in the strong cases, and an oblique stem **bou̯dill-* <**bou̯dilØn-* with zero-grade suffix and regular assimilation. The oblique stem then provided the springboard for newly thematised **bou̯dillo-* (attested as *Boudillus*), whence **-llo-* could be reanalysed as a suffix in its own right added to the *i*-stem **bou̯di-*. The notional connection with **bou̯dilāti-*, or a compound with any other second element starting in **l-*, had been lost at that stage. The derivation was felt to operate directly on the first element alone, which could be extended to nouns other than *i*-stems. An alternative, slightly variant explanation of the suffix **-llo-* starts from monothematic, originally adjectival names in **-lo-*, which were likewise 'individualised' by the addition of **-on-*, *e.g.*, **kamulo-* 'servant' or 'champion', attested in OIr. *cumal* 'female slave' → 'individualised' **kamulon-*, **kamuln-* → **kamullo-*, attested as Gaul. *Camullus*, OIr. *Cumall*.

In a further step, the presence of gemination in a suffix with *-l-* provided the starting point for the introduction of consonantal gemination into other suffixes via proportional analogy. In particular, in analogy to doublets where forms with **-lo-* and **-llo-* existed side by side with each other, doublets in **-kko-* could be created beside the very common suffixes in **-ko-*. This explains the observation in section 9.2.4. that some geminate sounds show a notable propensity to occur in suffixes.

2.2. *sm > *mm

The Proto-Celtic language showed a tendency towards weakening of word-internal *s* (Stifter, 2012: 541–542; Stifter, 2017: 1192), so much so that in various word-internal clusters with resonants, it was prone to disappear through assimilation to the resonant or, in other words, with compensatory lengthening of the resonant. However, only one such development can lay a secure claim on Proto-Celtic age by virtue of being attested for Celtiberian, namely **sm* > **mm*.

(1) The PIE pronominal dative singulars **Hi̯osmōi̯* 'to whom' and **tosmōi̯* 'to him' show up, *via* PC **i̯ommūi̯* and **sommūi̯* (with generalisation of the stem allomorph **so-*), as *iomui* and *somui* in Celtiberian.

(2) PIE 1sg. **h₁esmi* 'I am' > **emmi* > Gaul. *immi*, OIr. *am*.

It is interesting to note that while Celtic **s* is weak and feeble in sonorant contexts, it is strong and 'aggressive' in contexts with stops. That is to say, especially when following a stop, that sound is 'weakened' and eventually completely assimilated to the **s*. However, as soon as *m* is involved as well, it is the one that finally prevails over everything else. In less flowery language, **Tsm* (where *T* is any stop) appears to have become **mm* already in Proto-Celtic. Neuter verbal abstracts in **-s-man-* are one class of words where this can be observed in a number of items (for the intricate question of **-s-man-* as a variant of the suffix **-man-*, see Stüber, 1998: 52–58). One

example illustrates the pan-Celtic treatment of such complex clusters:

(3) qPIE *$ǵ^{h}$ŋg^{h}-s-men- 'act of stepping, striding' → pre-Celtic (pre-Clt.) *kangsman- > *kanχsman- > *kãmman- > OIr. *céimm*, W, Corn., MBret. *cam* 'stepping, step'; thematic derivatives are attested in Cib. *kamanom* 'path' and Gaul. *cammīnos*, the latter presupposed by Ital. *cammino*, Fr. *chemin*, *etc*. 'path, way' (Hamp, 1974).

Another example is supported by three languages:

(4) PC *garsman 'shout, cry' > *garmman > OIr. *gairm*, W, Corn., Bret. *garm*, and perhaps Gaul. *garma[n]*. For this explanation, I assume that *s in the cluster *rsm assimilated to the *m, and that the resulting *rmm contrasted with more common *rm to such an extent that it did not undergo the spirantisation to *rṽ in British. If, however, the *s assimilated to the preceding *r rather than to the following sound and the resulting *rrm merged with *rm before British spirantisation, unspirantised Brit. *garm* may be a borrowing from Irish (personal communication Paulus van Sluis).

Numerous other examples of neuter verbal abstracts of this type are conveniently collected in Stüber (1998: 45–83; 2015: 114–115) and need not be repeated here.

(5) A special case is PC *amman- 'time', attested in Gaul. *amman*, OIr. *amm*, and in the derivatives OIr. *aimser*, W *amser*, OCorn. *anser* (for *amser*), Bret. *amzer* 'time, weather'. It can either continue PIE *$h_2$et-s-men- or maybe 'asigmatic' *$h_2$et-men- 'the act of going around' (Stifter, 2017b).

(6) Structurally similar are masculine and feminine agent nouns with the suffix *-s-mon-, *e.g.*, OIr. *rúam* 'red dye' < *roṷdsmon- < qPIE *$h_1$reṷd^{h}-s-mon-, or OI. *femm*, *femmain*, W *gwymon*, LCorn. *gṷbman*, Corn. *goemmon*, Bret. *goumon*, Fr. *goémon* 'edible sea-weed' < *ṷimmon-, perhaps from qPIE *ṷip-s-mon- 'swayer' (Stifter, 1998: 204–207).

2.3. *ndn, nnd (and *nkn) > *nn

Triple clusters of two *n and a single *d, either in the order *ndn or *nnd, appear to have been simplified to a geminate *nn already in Proto-Celtic. Two or perhaps three nouns fall in the first group:

(1) qPIE *bŋ$d^{(h)}$-no/eh_2- > *bandno/ā- > Gaul. *banno-*, *benno-*, OIr. *benn*, W *ban*, Bret. *bann* 'top, tip, summit'. The Germanic cognates, OEngl. *pintel* 'penis' or MHG *pinz* 'awl', seem to point to a preform with *b-, so the word is perhaps a substratal loanword; for the nature of the dental, see Kümmel's (2016) remarks on the next item.

(2) qPIE *glŋ$d^{(h)}$-neh$_2$- > *glandnā- > *glannā- > Gaul., OBrit. *glanna*, W, Corn. *glan*, Bret. *glann* 'river-bank'. OIr. *glenn* 'valley' is a hybrid of this *glandnā- and *glendes- (W *glyn*, Bret. *glenn* 'glen, valley'). The cognate MLG *klint* 'hill' ostensibly suggests a preform with two unaspirated voiced stops in the pre-Proto-Germanic form. However, Kümmel (2016) offers an inner-Germanic explanation for the cluster *nt* in OEngl. *pintel* and MLG *klint* that would allow them to continue *nd^{h}* with an aspirated stop. Alternatively, an explanation through Kluge's Law would be possible, namely *glend^{h}nó- > Germanic *klentta- > *klinta-. Formally, a connection with the verbal root PIE *g^{h}lend^{h}- 'to look, shine' is possible, although semantically not immediately attractive.

(3) qPIE *$g^{(h)}$rŋ$d^{(h)}$no- > *grandno- > Gaul. *grannus*, OIr. *grenn* 'beard, bristles', W. *gran* 'cheek, jaw, beard', Bret. *gour(r)enn* 'eye-lash'; Provençal *gren* 'moustache' and OFr. *grenon* 'beard' are perhaps borrowed from Gaulish. The vocalism across the Celtic languages, which is parallel to that of the preceding two items, suggests a structure similar to those. IEW 440 thinks of a connection with the root *g^{h}er- or *g^{h}reh$_1$- (in modern notation) meaning 'to sprout, stick out'.

(4) Van Sluis *et al*. (2023: 226) propose that qPIE *kentno-, a thematic derivative of the *n*-stem *kenton- 'membrane, skin', was simplified to *kenno- > OIr. *cenn*, W *cen*, OCorn. [*c*]*he*[*n*], Corn. *kenn*, OBret. *cennen* 'skin'.

(5) The reconstruction PC *tundnā- is a formal possibility for OIr. *tonn*, W, Corn. *ton*, Bret. *tonn* 'wave'. This could be analysed as a formation from the PIE root *(s)teṷd- 'to push' with the nasal infix taken over from the verbal inflection, *i.e.*, structurally qPIE *tu-n-d-neh$_2$- (Stifter, 2023a: 179; other possible cases of nasal infixes transferred from verbal stems into the nominal inflection in Celtic are discussed in Stifter, 2018: 38). See 3.1. (10) for an alternative reconstruction of *tonn*.

Clusters with two nasals are particularly common in the present stem formation of verbs.

(6) OIr. *ro·finnadar* 'to find out' ultimately continues the PIE root *ṷeid- 'to see'. First, the *n* of the inherited nasal infix formation *ṷi-n(e)-d- became fossilised as *ṷind-, and then another nasal suffix was added to this neo-root, *i.e.*, PC *ṷind-nu-, which directly underlies the Old Irish present stem.

(7) Structurally similar, although involving a velar instead of a dental, is OIr. *srennaid* 'to snore', probably a denominal verb from the noun *srenn* 'snoring' < *srenk-no- or *sreng^{h}-no- (the precise character of the velar is difficult to determine).

McCone (1998) sets up several Celtic verbal stem formations as 'double nasal presents' with the sequence *nnd (differently now Jasanoff, 2022). The envisaged development is illustrated by the Indo-European root *g^{h}ed- 'to grab'. This had a nasal present in Indo-European, from which the secondary root *g^{h}end-/*g^{h}ŋd- was extrapolated in Celtic. This then received a new nasal infix, *i.e.*, *gan-n-d-e/o- 'to have room', reflected by OIr. *ro·geinn*, W *genni* (Schumacher, 2004: 330). Most of the pertinent roots already contained a fixed *n* in Proto-Indo-European. Schumacher (2004) includes the following verbs with this structure in his collection of Celtic primary

verbs: **bran-n-d-e/o-* ‘to well out’ < **b^{h}rend$^{(i)}$-* (233–234) **gan-n-de/o-* ‘to find room’ < **g^{h}end-* ←**g^{h}ed-* (330–331); **glan-n-d-e/o-* ‘to bring to light’ < **g^{h}lend^{h}-* (334–337); **skan-n-d-e/o-* ‘to jump’ < **skend-* (574–575); **tan-n-d-e/o-* (or **tend-e/o-*) ‘to break, cut’ < **tend-* (614–615). In the case of **gri-n-d-e/o-* ‘to drive, impel’ < **g^{h}re̯id^{h}-*, the change to OIr. *·greinn* is possibly analogical and may have occurred only during the historical period (McCone, 1998: 473; Schumacher, 2004: 354).

In addition to these relatively common sources for the geminate resonants *ll*, *mm* and *nn* mentioned in the foregoing sections, a number of rarer consonant clusters may have fed into the creation of marginal geminate stops through regular phonological change.

2.4. *dk > *kk

Martin Kümmel (in peer-review) points out that inherited PIE **dk* was apparently not assimilated in Celtic, but just simplified to **k* in the formation of the decades, such as PC **u̯ikantī* ‘20’, **trikont-* ‘30’ < (q)PIE **(h$_1$)u̯i-dḱm̥tih$_1$*, **tri-dḱomt-*, where other Indo-European languages show different treatments of the cluster, namely **dḱ* > **nḱ* in Indo-Iranian (Sanskrit *viṃśatí-*, *triṃśát-*; Ossetic Digoron *insæj*, *ærtin*) or loss of **d* with compensatory lengthening of the preceding vowel in Latin or Greek (e.g., Latin *uīgintī*, Doric Greek *u̯īkati*). The following words must therefore belong to a younger layer of formation. A handful of Insular Celtic examples illustrate the assimilation of **dk* > **kk* at the end of root syllables. Since this assimilation is also found across the composition boundary in Gaulish (see section 6.2.), it is likely that the change is already Proto-Celtic.

(1) qPC **ad-ki̯o/ā-* ‘nearness (< *at-ness)’ > OIr. *aicce* ‘proximity’, W *ach* ‘beside, near’.

(2) qPIE **h$_1$rud-ko-* ‘red’ → **rud-ki̯o-* ‘redness’ > **rukki̯o-* > OIr. *ruccae* ‘shame’.

(3) qPIE **pr̥d-keh$_2$-* > **φridkā-* > **φrikkā-* > W *rhech* ‘fart’.

(4) qPIE **slHd-kV-* ‘slay-thing’ > **slad-kV-* > **slakkV-* > OIr. hapax *slacc* ‘sword’, ModIr. *slacán* ‘bat’ (Schumacher, 2004: 585).

A fortiori one might expect that **kk* would also continue **gk* and **tk*, but unequivocal examples are difficult to identify. A number of possible instances are discussed in section 12.

(5) In the case of inherited PIE **h$_2$ŗtḱo-* ‘bear’, the outcome is Gaul. *arto-*, OIr. *art*, W *arth* < PC **arto-* < **arχto-* < **artko-*. One way of accounting for the different treatment from the foregoing examples is to assume that sequences of dental and velar underwent metathesis at the syllable onset, i.e. **ar.tko-* > **ar.kto-* etc., while no metathesis occurred across the syllable boundary, e.g. **rud.ko-* > **rut.ko-* > **rukko-* etc. The metathesis also occurred in **d^{h}g^{h}omi̯o-* ‘earthling, human’ > **dgoni̯o-* > **gdoni̯o-* > Gaul. *-χtonion*, OIr. *duine*, W *dyn*, Corn., Bret. *den* ‘person’.

2.5. ‘Geminating’ *p?

It is possible that the weakening and loss of IE **p* after resonants led to a marginal increase of geminates. The best examples involve the cluster **rp*:

(1) PIE **sterp-* or **stirp-* → **sti/erφākо-* > OIr. *serrach* ‘colt, foal; young animal’. For the root, cf. Lat. *stirps* ‘stem, stump, stock, race’, and Lith. *stir̃pti* ‘to grow up’ and *ster̃ptis* ‘to stiffen’.

(2) PIE **serpeh$_2$-* ‘sickle’ > **serφā-* > **serrā-* > OIr. *serr*, MW, OBret. *ser* ‘sickle’. However, a loan from Lat. *serra* ‘saw’ < **sersah$_2$-* is not excluded, especially in view of the retention of initial *s-* in British instead of its regular development to *h-*. Finally, the Celtic word could even be cognate with the Latin.

The cases of **lp* are more problematic:

(3) The root PIE **telp-* ‘to make room’ is probably continued by OIr. *do·alla*, *·talla*, occasionally *do·ella* ‘there is room for; to find room’. At first glance, this verb seems to support the notion of the change **lp* > **ll*, but perhaps the **ll* is rather due to a nasal-infix formation, *i.e.*, **tl̥-n-p-* > **taln-(φ)-* > **tall-*. The deuterotonic form *do·alla* must be analogical.

(4) Pedersen (VKG i 78–79) compares OIr. *sell* ‘iris’, *sellaid* ‘to see, perceive’, MW *syllu* ‘to gaze’, Bret. *sellout* ‘to watch’ with Gr. στιλπνός ‘brilliant’, στίλβω ‘to shine, be brilliant’. Even if that comparison were correct, the geminate *ll* could be due to the **n* of the suffix, not the **p*, *i.e.*, **stilpno-* > **stilφno-* > **stilno-* > **stillo-*. For an alternative explanation of PC **s(t)ill-* see section 3.7.

(5) pre-Celtic **kelpurno-* > OIr. *cilorn*, W *celwrn*, OBret. *chilorn*, Bret. *kelorn* ‘vessel’, perhaps also in the Old British placename *Cilurnum*, speaks clearly against the notion of such a sound change. The word, which has cognates in Lat. *calpar* and Gr. κάλπις ‘vessel, pitcher’, is possibly a substratal loan (cf. Van Sluis forthc.).

(6) OIr. *col*, W *cwl*, OBret. *col* ‘wrong, sin’ could be related to Lat. *culpa* ‘blame, guilt’, in which case they would continue PE **kulφo-* < **kulpo-*, but alternative explanations are possible.

Finally, it is conceivable that instances of the very rare inherited cluster **pk* led to geminate **kk*. Before a voiceless dental, PIE **p* became PC **χ via* the intermediate stage **φ*, *e.g.*, **septm̥* ‘7’ > **seφtam* > PC **seχtam*, cf. OIr. *secht^{N}*, MW, LCorn., OBret. *seith*, Bret. *seizh*. This means that it fell together with any kind of tectal sound in this position. It is conceivable that, in a parallel manner, **pk* became **χk via *φk*, which was then further assimilated to **kk* (Testen, 1999). Possible examples are, by necessity, few, but this explanation could account for the apparent suffix in some words for ‘swine’, *e.g.*, **su(H)-pḱ-u-* ‘swine-stock’ > **sukku-* → W *hwch*, OCorn. *hoch*, Bret. *houc’h* ‘pig’, OIr. *socc* ‘ploughshare, snout’, or **mo(H)-pḱ-u-* ‘big (?)

stock' > **mokku-* > Gaul. *Moccus*, OIr. *mucc*, W *moch*, Corn. *mogh*, Bret. *moc'h* 'pig'. The suffixal element would be the zero-grade of PIE **peḱu-* 'small livestock' (Testen, 1999). Perhaps this suffix is also contained in **brokkV-* 'badger' > Gaul. *brocco-*, OIr. *brocc*, W, OCorn. *broch*, Bret. *broc'h*, although the word is more likely a substratal loan (see 13.1. (5)). In the British Celtic languages, this suffixoid was perhaps also extended to the word for 'cow, cattle', PC **bou̯-*, to create **bou̯kko-* > W *buch*, *buwch*, OCorn. *buch*, Bret. *buc'h*.

3. Etymological gemination: developments across branches with identical or similar outcome

Further assimilations, predominantly involving clusters with resonants, took place in a staggered fashion after the end of the Proto-Celtic linguistic unity. From what the fragmentary documentation of the ancient Celtic languages allows us to see, none of these changes were carried through in all Celtic languages. Gaulish and, more importantly, Celtiberian sometimes reflect a more conservative stage. In the Insular Celtic languages, however, the changes discussed in this chapter are completely carried through. This is not a proof for a genetically close relationship between Goidelic and British, but it may be a function of their comparatively late attestation, which is to say that tendencies towards assimilation, which had their kernel much earlier, had enough time to unfold completely in these languages.

What unites these assimilations and those of the previous section, is that they occur word-internally, but not word-initially, and that they happened after the vocalisation of syllabic resonants. There is no single overarching rule that accounts for all contexts.

3.1. *sn > *nn

Two forms in Celtiberian appear to indicate that, unlike **sm*, **sn* had not assimilated to **nn* already in Proto-Celtic, even though there is ample evidence from the three other branches that the change was carried through before the historical period. The two items that show the conservative behaviour of Celtiberian are:

(1) PIE **pr̥h₃-sneh₂-* > **φrasnā* > **φrannā* > OIr. *rann*, W *rhan*, Corn. *ran*, Bret. *rann* 'part, share'. If Cib. *arznas* in Botorrita I derives from the same preform (Jordán Cólera, 2019: 107, 121), it would demonstrate that the assimilation to **nn* was not of Proto-Celtic age; however, the divergent syllabification of the initial resonant is also to be noted.

(2) The comparison of OIr. *trén*, ogam TRENA- 'strong', which continues PC **treχsno-*, with Gaul. *Trenos* indicates that clusters of the type **-Csn-* were subject to an early reduction to simple **-n-* at least at the Core Celtic stage. W *tren* 'fierce, rapid, powerful' is only attested from the very end of the 18th century and is therefore too uncertain to provide any evidence. Cib. *masnai*, also in Botorrita I, has been suggested to continue either **mak-snā-* 'enclosure' or **mad-snā-* 'breaking' (see Wodtko, 2000: 244). In either case, the treatment of the triple consonant cluster deviates from that in **treχsno* and the cluster *-sn-* is evidently maintained.

Strong evidence for the change **sn* > **nn* in the languages apart from Celtiberian comes from the following items:

(3) The Old Irish paradigm of nominative (nom.) *brú* 'bosom' < PC **brusū*, oblique stem *bronn-* < **brunn-* < **brusn-* is a morphological fossile that proves that the suffixal ablaut of PIE **bʰrusō(n)*, **bʰrusn-* had been inherited into Proto-Celtic. The weak stem **bʰrusn-* also underlies **brunn̥io-*, *i.e.*, OIr. *bruinne* 'bosom, breast', W *bryn* 'hill, mound, prominence', Corn. *brenn* 'hill', and was also borrowed into Proto-Germanic **brunjōn-* 'breastplate'.

(4) qPIE **dʰusno-* 'smoky' > **dunno-*, Gaul. *dunno-*, OIr. *donn* 'brown', W *dwn* 'some shade of brown'.

(5) PIE **h₁osno-* > **onno-* > Gaul. *onno-*, OIr. *onn* and, with further suffixation, *uinnius*, W *onn*, OCorn. *onnen*, Bret. *ounn* 'ash-tree'.

(6) PIE **h₁u̯h₂sno-* 'empty' > **u̯anno-* > OIr. *fann*, W *gwan*, OBret. pl. *guenion*, Bret. *gwan*, OCorn. *guan* 'weak, feeble'. For the problems with the PIE reconstruction, see Zair (2012: 46–47).

(7) qPIE **kₔsn-* or **kasn-* → **kannīnā-* > OIr. *cainnenn*, W *cennin*, OCorn. *kenin*, Bret. *kinnen* 'onion, leek'. This etymon is perhaps also contained in the ancient Germanic ethnonym *Canninefates*, if it means 'leek lords'.

(8) qPIE **ḱₔsnih₂-* or **ḱasnih₂-* 'female grey one' > **kasnī-* > **kannī-* → W *ceinach* 'hare'. This etymon may also underly the OIr. personal name *Cainnech*, possibly from **kannīko-*; this name is unlikely to be derived from *cainnenn* 'onion, leek', since that word did not syncopate when an extra syllable was added, as the genitive *cainninne* shows. In view of the reconstruction in (7) above, this lack of syncope must be secondary.

(9) pre-Celtic **kᵘresno-* (perhaps a substratal loanword) > **kᵘrenno-* > LGaul. *prenne*, OIr. *crann*, W, OCorn. *pren*, Bret. *prenn* 'tree, wood'; the vocalism of OIr. *crann* may be due to some sort of ablaut. For the underlying root **kᵘres-*, cf. **kᵘrestio-* > W *prys*, Sc. Gael. *preas* 'copse, grove', and **kᵘerstV-* > OIr. *cert* 'name of a letter', W *perth*, Corn. *perth* 'hedge, bush'. It has been suggested that the same etymon be contained in OIr. *currach*, *cuirrech* 'swamp, bog' (*Léxique Étymologique de l'Irlandais Ancien* (LEIA) C-278; see 3.2. (5)).

(10) qPIE **tuh₂sneh₂-* 'swelling (?)' > **tusnā-* > **tunnā-* > OIr. *tonn*, W, Corn. *ton*, Bret. *tonn* 'wave' (Stifter, 2023a: 179; Zair, 2012: 155). For an alternative explanation, see 2.3. (5).

(11) qPIE **u̯lHsno-* 'wounded (?)' > **u̯lanno-* > OIr. *flann* '(blood-)red' (Zair, 2012: 73).

(12) I have suggested a pertinent etymology for Gaul. **dannos* 'title of a public or religious official', namely *via *dasno-* from

PIE *$d^{h}h_1sno$- ‘pertaining to the religious sphere’, cf. Lat. *fānum* ‘sanctuary’ (Stifter, 2011: 166–167).

(13) Matasović (2009: 416; 2020: 337–338) sets up *$u̯esnālā$-, derived from PIE *$u̯esr/n$- ‘spring’, as the preform for OIr. *fannall*, W *gwennol*, Bret. *gwennel* ‘swallow’. I have instead proposed a substratal loan from a preform *$u̯annell$- (Stifter, 2010; but cf. the critical comments by Egurtzegi & Ariztimuño, 2013).

(14) Repanšek (2016: 96–97) suggests that the toponymic suffix *-*ennā* (e.g. in *Arduenna silva* etc.) continue *-*es-nā*-, i.e. a derivative in *-*nā*- from neutral *s*-stems. As a pertinent example, Stefan Schaffner (personal communication) cites the modern place name *Rovenna* (a quarter of Cernobbio at Lake Como) < *$reu̯h_1es$-nah_2- ‘place in the plain’, built on the neuter *s*-stem PIE *$réu̯h_1es$- ‘plain’ (cf. OIr. *roë* ‘level piece of ground’ < PC *$rou̯esi̯ā$- < qPIE *$reu̯h_1es$-$i̯eh_2$-).

3.2. *rs > *rr

A series of words, not always of clear etymological analysis, suggest that clusters of *r + *s were preserved in some form in Celtiberian (Jordán Cólera, 2019: 121). Instances of *r* + sibilant are Cib. *arznas* (already discussed in 3.1. (1)), *tikerzeboz*, *uerzoniti*, *uerzaizokum*. *Uerzaizokum* has been connected with the root *$u̯ers$- ‘better, elevated’; for *uerzoniti*, an analysis as *$uper$-$sonh_2$-$ei̯e$-*ti* has been proposed (cf. Jordán Cólera, 2019: 202, 208). However, the explanation of none of these words is beyond doubt. Other examples such as the coin legends *arzakoz*, *arzakozon* and *arzaoz* (Jordán Cólera, 2019: 279–348) are too unclear to be used as evidence. They could reflect Iberian names. A single Hispano-Celtic example has been adduced as positive proof for the assimilation of *rs > rr, namely the placename adjective *erredicis*, which Prósper (2002: 318) analyses as *per-sed-$i̯o$-. In view of the apparent retention of *rs otherwise in Celtiberian, this analysis must remain doubtful.

The Gaulish corpus also contains several examples with the surface sequence *-rs-*. Whatever is the precise analysis of the names Ουερσικνος, *Versenus* and *Versinius*, it can be argued that they are productively formed compounds with the intensive prefix *$u̯er$- ‘super’ + lexemes beginning with *s*-, *e.g.*, *$seno$- ‘old’. They are therefore not reliable examples for the treatment of inherited *rs in Gaulish. Several Gaulish names are derived from the etymologically obscure stem *darso*-. Delamarre (2003: 136) tentatively compares this with Late Lat. *darsus*, the name of a fish (Fr. *dard*, Engl. *darter*). Bret. *dars* ‘id.’ must surely be borrowed from Latin or Romance. The reading *incors* in the inscription from Larzac is uncertain (there is a space between the *r* and the *s*, so the two letters could belong to different words). The Cisalpine Celtic corpus has one example of *-rs-*, the patronymic *uarsileos* of unclear analysis. Finally, the names *Morsinus* and *Morsius* of unclear analysis are from Noricum and may not be Celtic at all.

Apart from cases that are recent, productively formed compounds, the precise nature of the clusters in the examples above is uncertain. Are they inherited instances of *-rs-* or did they also arise through secondary processes or through re-analysis? This question is all the more legitimate since Gaulish does provide examples that show the operation of the change *rs > *rr*. In order to reconcile the contradictory evidence, one can speculate that the change was just under way in Gaulish at the beginning of the historical period, and that some examples reflect the period before, and others the period after the change took place.

(1) PIE *$kr̥so$- ‘runner’ > *karso- > *karro- > Lat. *carrus* (from Gaulish), OIr. *carr*, W *car*, Bret. *karr* ‘cart’.

(2) PIE *$b^{h}r̥so$- ‘point’ > *barso- > *barro- > Gaul. *barro-*, *baro-* (?), OIr. *barr*, W *bar*, Bret. *barr* ‘top’.

In the Insular Celtic languages, the change has been carried through completely, as further examples demonstrate:

(3) PIE *$ĝ^{h}erso$- ‘small, short’ > *gerso- > *gerro- > OIr. *gerr* ‘short (cropped)’, beside *gair* ‘near’ < *$ĝ^{h}r̥i$-.

(4) PIE *h_1erseh_2- ‘hinder part, buttocks’ > *errā- > OIr. *err*. Perhaps also in *eirr* ‘chariot fighter’, if from *ers-sed- ‘he who sits at the back’, in contrast to *arae* ‘charioteer’ < *are-sed- ‘he who sits at the front’.

(5) It has been suggested to compare OIr. *currach*, *cuirrech* ‘swamp, bog’ with OE *hyrst* ‘hillock, wooded eminence’, OHG *horst* ‘wood’ < *$k^{u}r̥sti$- (cf. LEIA C-278; but see also 3.1. (9)), which requires an ablauting preform *$k^{u}orsi̯ko$- > *$k^{u}orri̯ko$-. It is not clear if the *o* in the sequence *orri would be raised to *urr in Old Irish.

(6) PIE *$u̯r̥seh_2$- ‘high point’ > *$u̯arsā$- > *$u̯arrā$- > OIr. *farr* ‘prop, post’, W *gwar* ‘nape of the neck’, OCorn. *guar* ‘neck’.

(7) PIE *$u̯erso$- ‘being on a high point’ > OIr. *ferr* ‘better’ This is the same etymon as the preceding item, but with a different ablaut grade. See also W *gwell* in 3.7. (2).

(8) qPIE *dor-so- > *dorro- > OIr. *dorr* ‘harsh, rough’, *doirr* ‘anger’. This example is uncertain, since the word is restricted to Irish and is late attested. If inherited, it may be built on the PIE root *der- ‘to tear’; for possible extra-Celtic cognates see LEIA (D-159) and Matasović (2009: 103). See also section 11.3. (2).

(9) PIE *$(s)tord$-s-V-, *$(s)tr̥d$-s-V- > *torsV-, *tarsV- > OIr. *tarr* ‘stomach’, *torrach* ‘pregnant’, OW *torr*, W, OCorn. *tor*, Bret. *tor*, *teur* ‘stomach’, W *torrog* ‘pregnant’. IEW 1024 and LEIA (T-33–34) derive this etymon from the extended Indo-European root *$(s)ter$-d- ‘to be stiff, sturdy’, with reference to an inflated belly. Matasović (2009: 385), on the other hand, doubts the Indo-European pedigree of the word.

(10) The merger of the preposition *for* ‘over, upon’ < *$u̯or$ ←*$u̯er$ < *$uper$ with enclitic pronouns starting with *s*- has resulted in inflected prepositions with *-rr-* in Old

Irish, namely 3sg. fem. *forrae*, *fuirre* < **u̯or-siu̯am*, and 3pl. *forru* < **u̯or-sūs*.

3.3. *ls > *ll

Like in the case of the preceding change, Celtiberian provides an example that shows the unassimilated cluster, namely VELSAM. Due to the lack of a clear analysis of the word, it is uncertain whether its internal cluster is the retention of an inherited sequence or is the result of a secondary development (see Jordán Cólera, 2019: 908 for proposals, none of which is compelling). Other names in Celtiberian sources with the sequence *-ls-* such as *belsa*, *belsu*, *kelse* are probably borrowings from Iberian. There is no secure example in Celtiberian of the change **ls* > **ll* having taken place. As for Gaulish, Delamarre (2007: 194) records the name (in the dative) *Velsounae Suiocae Vescleuesis f.* from Dalmatia. Its provenance, together with the un-Celtic name of the father (**eu̯* should have become **ou̯* in Celtic), renders it unlikely that *Velsouna* is Gaulish. The Gaulish inscription from Larzac (L-98 2b3) contains the form *(s)uolson*, which Lambert (cited in Delamarre, 2003: 327) tentatively compared with W *gwall* 'mistake, error' (see below).

In the Insular Celtic languages, at any rate, the change has been carried through:

(1) PIE **melso-*, **mḷso-* 'deceit, fault' > **melso-*, **malso-* > **mello-*, **mallo-* > OIr. *mell* 'destruction, confusion, error', W *mall* 'destruction, evil'; perhaps also the first element *mallo*, *mello-* in Gaulish personal names.

(2) PIE **pḷso-* 'rock' > **φalso-* > **φallo-* > OIr. *all* 'cliff'.

(3) From the Indo-European base **u̯el-s-* 'deceit (?)' (IEW 1140), a full complement of ablaut formations may be reflected in Celtic. The full grade PIE **u̯elso-* gives **u̯ello-* > OIr. *fell* 'deceit, treachery'. The zero grade PIE **u̯ḷseh₂-* > **u̯alsā-* > **u̯allā-* can underlie OIr. *fall* 'neglect, negligence', W *gwall* 'fault, error, failure', Bret. *gwall* 'bad'. Alternatively, W, Bret. *gwall* and Gaul. *uolson* (see above) can also continue the *o*-grade PIE **u̯olso-*.

(4) **kᵘelso-* 'distant' and **mḷso-* 'hesitant' are viable alternative reconstructions for the adjectives that have been discussed as **kᵘelno-* and **mḷno-* in section 2.1. above.

3.4. *sr

The case of **sr* is different. The alleged change **sr* > **rr* in OIr. *errach* 'spring' < **u̯esr-* (VKG i 82) is highly questionable. Not only is there no good parallel for this change (see the arguments below), but the loss of the initial **u̯* would also be unprecedented. *Errach* is better analysed as an adjectival derivate of *err* 'hinder part, extremity, tail' (see 3.2. (4)), *i.e.*, 'that which is at the end (of the winter)'. The spelling *dírruidiguth* 'derivation' beside *dírśuidigud* (both in the St Gall glosses, namely Sg. 53a11 and 188a8) for the verbal noun of *díśruthaigidir* 'to derive' does not prove the Proto-Celtic assimilation of **sr* > **rr* since the verb is an Old Irish calque on Latin *dērīuāre*, which shows the synchronic 'strengthening' or devoicing of *r* caused by lenited **ṡ*.

The best examples of inherited **sr* show a single *r* as the outcome in Irish and, to a lesser degree, in British, but the details of the intermediate stages are not entirely clear. It is possible that **s* became **ð* before **r*, perhaps passing through a stage **z* first.

(1) This is best illustrated by the feminine numeral **tisres* > Gaul. *tidres*, OIr. *teóir*, MW *teir* '3' (Kim, 2008: 160–161). On the basis of metrically disyllabic *teüir* in the two poems on the first page of the Milan manuscript of Old Irish glosses (MS Ambros. C301 inf.; *Thes.* ii 291–292), McCone (1996: 47) wants to derive this from **tēsūres*. However, the metrical evidence is not conclusive, since the poem is comparatively late, perhaps from c. 850, and the hiatus of *teüir* can be poetical licence.

The development **sr* > **ðr* > **r* suggested here is also compatible with the following examples:

(2) PIE **kēs-reh₂-* 'tool for combing' > **kīsrā-* > **kīðrā-* > **kīrā-* > OIr. *cír* 'comb'.

(3) PIE **u̯ōs-rV-* 'related to spring' > **u̯āsrV-* > **u̯āðrV-* > **u̯ārV-* > OIr. *fáir*, W *gwawr* 'sunrise, east'. In the case of W *gwawr* it has to be assumed that a segmental reflex of the **s* like in *teir* above was inhibited by the preceding long vowel.

(4) The case of **mēmsro-* (from **mēms-* 'piece of meat') > OIr. *mír* 'bit, piece of meal', although similar, is best kept apart, since a more complex consonant cluster is involved.

(5) Delamarre (2003: 128–129) explains Gaul. *craro-* as 'hornet' < PIE **ḱr̥h₂sro-* (cf. Lat. *crābrō*), but this is pure speculation since the meaning of this name element is unknown. If the etymology is correct, it would show the loss of **s*.

(6) PIE **su̯esōr*, *su̯esr-* 'sister' > **su̯esūr*, *su̯esr-* is continued as Gaul. dat.pl. *suiorebe*, OIr. *siür*, gen. *sethar*, nom. pl. *sethir*, W *chwaer*, OCorn. *huir*, LCorn. *hôr*, Corn. *hwoer*, Bret. *c'hoar*. Of these, the Gaulish and British forms are irrelevant here since they continue the full-grade allomorph of the stem. Only the allomorph *seth(V)r-* in Old Irish is potentially relevant for the present question. It has been suggested that **-θr-* in the oblique cases continue directly the zero grade **-sr-* of the suffix. However, this cannot be used as proof for the regular treatment of **sr* in Celtic since these forms may have been remodelled after the other kinship terms, which all have oblique stems in *-thr-* from **-tr* (McCone, 1994: 277–278, 283).

3.5. *sl > *ll

Only ambiguous evidence is found for the single relevant example in Ancient Celtic.

(1) PIE **koslo-* 'hazel' > **kollo-* > OIr., W, Bret. *coll*, OCorn. *colwiden*. Delamarre (2003: 127) refers various Gaulish personal and placenames to this etymon: the personal name *Collus* and the placename *Collex* < **Collācon* in Switzerland have geminate *ll*. However, the French placenames *Coole* < *Cosla*

(attested 983), *Coolus* < *Coslus* (attested 869), *Coulon* < ancient *Coslumnus* display unassimilated *sl* still in historical times, as does the divine name *Cuslanus* in Cisalpine Gaul. If the latter forms belong to this etymon, the change can only have happened subsequent to the Proto-Celtic stage.

In the Insular Celtic languages, at any rate, the change has been carried through.

(2) qPIE *d^hrus-li-* 'fragment' > *drusli- > *drulli- > W *dryll*, Bret. *druilh*, Corn. *dral*.

(3) qPIE *k^uei̯sleh$_2$-* 'faculty of seeing (?)' > *k^uei̯slā- > *k^uei̯llā- > OIr. *cíall*, W *pwyll*, Bret. *poell* 'sense, reason'. The underlying root is PIE *k^uei̯s-* 'to see, perceive'.

(4) qPIE *neu̯H-slo-* or *nou̯Hslo-* 'shout, cry' > *nou̯slo- > *nou̯llo- > OIr. *núall* 'cry, noise'.

(5) qPIE *(s)tug/k-slo-* 'treated with a hammer or chisel (?)' > *tuslo- > *tullo- > OIr. *toll* 'pierced', W *twll*, Corn. *toll*, Bret. *toull* 'hole, perforation'. There are personal and placenames with *tull-* in Gaulish, but their relationship with this etymon are unclear.

Geminate *-ll-* is also the outcome in Old Irish verbal compounds when root-initial **sl-* comes to stand between two vowels. For instance, the compound verb **to-sli-i̯e/o-* 'to earn' comes out as deuterotonic *do·slí*, but prototonic *·tuilli*. In British, in contrast, root-initial **sl-* in such contexts is treated like lenited *l*, *e.g.*, the root **slad-* 'to slay' occurs in the compounds W *ymladd* 'to kill, fight', Corn. *omladh* 'to fight', or OBret. *anlaedam* 'I attack' < **ambi-slad-*. In the case of W *dyrllid*, Bret. *dellit* 'to earn' < **to-ro-sli-i̯e/o-*, *ll* is not a trace of **sl*, but is due to the deleniting effect of *r* upon a following *l*.

3.6. **nl* > **ll*

Only examples from Irish come to mind for this assimilation. Two are compounds with the preverb **en-* + a second element starting in **l-*. It is not excluded that their behaviour is analogical after other preverbs and that the change occurred as late as the immediate prehistory of Irish.

(1) The prototonic stem of OIr. *in·loing* 'to join, unite' is *·ellaing* < PC **en-lunge-*, its verbal noun is *ellach* < **en-lou̯go-*.

(2) OIr. *ellam* 'ready, prompt' and 'bride-price' can be analysed as a compound of PC **en-* 'in' + **lāmā-* 'hand'.

(3) In Stifter (2011b: 558–559), I proposed to explain the legal term *noíll* 'oath' as a Proto-Goidelic compound **nou̯en-lug-* 'nine-oath'.

3.7. **rl* > **ll?*

I am aware of three candidates for the assimilation of **rl* > **ll*.

(1) Schrijver (1995: 421–422) derives OIr. *sell* 'iris', *sellaid* 'to see, perceive', MW *syllu* 'to gaze', Bret. *sellout* 'to watch' from PC **stillo-* < pre-Celtic **stīrlo-*, a derivative from PIE **$h_2$stēr-* 'star', *via* assimilation and an Osthoff-type shortening of the vowel. A semantic parallel for the metaphorical use of a heavenly body for 'eye' is the PIE dual form **suh$_2$lih$_1$* 'two suns', which underlies OIr. *súil* 'eye'. Pedersen (VKG i 78–79), on the other hand, compares Gr. *στιλπνός* 'brilliant' (2.5. (4)).

(2) Thurneysen (GOI 236) explains W, Bret. *gwell*, Corn. *gwell* 'better' from **u̯erlo-*, from the same PIE root **u̯er-* 'top' that gives OIr. *ferr* 'better' < **u̯erso-* (3.2. (7)), but with a different suffix. Alternatively, the word could continue **u̯ello-* < **u̯el-so-* or **u̯el-no-* (Matasović, 2009: 411) from the root **u̯elH-* 'to want, wish', with loss of the laryngeal.

(3) OW *guolleuin*, MW *gollewin* 'west' < **u̯or-lugu-īno-* (uel sim.) is a compound of the preverb **u̯or-* 'on, upon' + a derivative of the word seen in W *lleu* < PC **lugu-* 'light, brightness'. In ModW *gorllewin*, the preverb *gor-* < **u̯or-* has been re-introduced.

The chronological position of the change hinges on the first example. Because this etymon is continued in Goidelic and British, it would have to have been formed prior to the separation of those two branches. The other two words only give evidence for this change in British.

3.8. **ld* > **ll?*

None of the examples for the notion that **ld* became **ll* in Proto-Celtic are conclusive, and there is one strong argument against it. I will start with the alleged examples in favour of that change.

(1) The precursor of OIr. *caill* 'forest', W *celli*, OCorn. *kelli* 'grove, copse' has been reconstructed as PC **kaldi-* and compared with PGerm. **hulta-* 'wood' < **kl̥do-*. However, PC **kalnV-* would be equally possible, perhaps derived as 'place for making lumber' from the root **kelh$_2$-* 'to strike', with unclear loss of the laryngeal. See Zair (2012: 182–183) for further speculations.

(2) OIr., W *coll*, OCorn. *collet*, Bret. *koll* 'destruction, damage' has been compared with PGerm. **halta-* 'lame' < PIE **kol(h$_2$)do-*, from the same root **kelh$_2$-* 'to strike' as, perhaps, the preceding item. Because of the semantic distance between the languages, there is no compelling reason to operate with identical formations. Like before, the Celtic word could continue a formation with a nasal suffix, *i.e.*, **kolh$_2$no-* with loss of the laryngeal by Saussure's Law. Thurneysen (GOI 95) and Hamp (1974e: 196) compare Lat. *culpa*, but the notion of PC **ll* < **lp* was rejected in section 2.5. Hamp also sets up the alternative reconstruction **kol(d)no-* without further discussion.

(3) OIr. *sall* 'bacon' could be derived from the stem **saldo-* < PIE **sh$_2$l-dh$_3$-o-* (see NIL 586–587 for this reconstruction), but a formation with a nasal suffix, *i.e.*, **sh$_2$l-n-*, is equally possible and has been adopted in section 2.1. (9). Other evidence for **sald-* in Celtic, namely W *hallt* 'salt, salty', is ambiguous. British Celtic did undergo the relatively late sound

change **ld* > *llt*. The Latin loan W *swllt* 'shilling', OCorn. *sols*, Bret. *saout* 'money' < VLat. *soldus* < *solidus* furnishes the most solid instance for this change. A native item that possibly shows the change **ld* > *llt* is W *mellt* 'lightning', if it goes back to **meld^{h}V-* 'lightning, hammer of the god of thunder'. Other words are doubtful, as Anders Jørgensen (in peer-review) points out. W *melltith* 'curse' can hardly continue VLat. *maldictio* < *maledictiō* directly, but probably has secondary, learned *d*, devoiced by the preceding *ll*. For the form without *d*, compare MBret. *malloez* < *maledictiō*, *millizyen* < *maledictiōne*, Bret. *millig-*, MCorn. *mylyg-* < *maledīc-*. Secondary *d* also seems to be the case with MBret. *bennoez*, *binizyen* < *benedictiōne*, *binnig-* < *benedīc-* as opposed to W *bendith*, *bendig-* 'blessing'. That the medial cluster in *melltith* cannot be old also follows from the fact that Lat. *-ld-* gives Middle Welsh *-ll-* medially, cf. *caldāria* > **kaltɔr* > MW *callawr*, MBret. *cauter* 'cauldron, kettle'. While **saldo-* is therefore a conceivable preform of *hallt*, that word could also continue **sal-* + adjectival **-to-*, *i.e.*, 'provided with **sal-*'. Therefore, neither Irish nor Welsh provide unambiguous evidence for a Celtic preform **saldo-*. If *hallt* continued **sald-*, it would disprove the Proto-Celtic age of the assimilation **ld* > **ll*.

(4) In any case, PIE **meldo-* 'mild, soft' > OIr. *meld*, later *mell* 'pleasant' is decisive evidence that the cluster was retained unchanged up to Old Irish, and underwent assimilation to *ll* only during the medieval period.

The changes in section 3.1–section 3.8. were major additional sources of geminate resonants, especially in the Insular Celtic languages. Although they include no stop sounds, through their number they reinforced geminates as a phonological class in Celtic. Kuryłowicz (1957: 141–144) makes a similar point about the pivotal role that geminate resonants arising from soundchange played in establishing gemination as a marked phonological category, but his arguments for gemination as a morphological process as a whole are not convincing and rely only on a small number of examples.

3.9. *'tau Gallicum'*

In this section, a diversity of phonologically related clusters are treated together, although they belong to several different chronological layers. Clusters of dentals involving voiceless dentals, of dentals followed by *s*, and of *s* followed by *s* across a morphological boundary, are grouped together under a single heading, glossing over various issues that are not relevant to the present question. Already at the Proto-Indo-European stage, clusters of dentals developed an excrescent medial sibilant on an allophonic level (section 3.9.1–section 3.9.2.). These clusters merged with clusters of dentals with original **s* over the course of the ancient Celtic period (section 3.9.3–section 3.9.6.). In the ancient languages, their behaviour and orthographic representation is still differentiated, which indicates that they had not all fallen together indiscriminately yet. Depending on their origin, they may appear as biphonemic assibilated or as geminate sibilant sounds, traditionally referred to as *tau Gallicum*. The precise phonetic nature of this sound, which is here provisionally represented as **ts*, is unclear (see Eska, 1998 for more details). All these clusters end up as *s* in the Insular Celtic languages, *via* an intermediate stage of geminate or 'strong' **ss*.

The general outlines of these developments are well-known (*e.g.*, McCone, 1996: 48, 99; Stifter, 2017: 1192); the intricate minutiae do not need to detain us in this article, which focusses on the bigger, systemic picture of Celtic sound developments. Seven different types of input lead to the same outcome in the Insular Celtic languages. Completeness is not envisaged here, only a few notable examples will illustrate each type.

*3.9.1. *tt > *t^{s}t > *ts > ss*

(1) qPIE **ret$^{(s)}$-ti-*, **ret$^{(s)}$-tu-* 'running', an abstract noun of the verbal root **ret-* 'to run' > **retsi/u-* > **ressi/u-* > Gaul. *ressi-*, W PN *Rhys*, OIr. unstressed compound element *-ras/ -rus*.

(2) PIE **h$_2$lit$^{(s)}$-tu-* 'suffering' > **litsu-* > **lissu-* > OIr. *lius* 'disgust, loathing'.

(3) The PIE root **u̯et-* 'to turn to' formed the past participle **u̯et$^{(s)}$-to-*, which *via* **u̯etso-* > **u̯esso-* constitutes the root of the OIr. impersonal verbal form *do·cuäs* '(someone) has gone'.

*3.9.2. *dt > *d^{s}t > *ts > ss*

(1) PIE **med$^{(s)}$-tu-* 'mast, nourishment' > **metsu-* > **messu-* > OIr. *mess*, W *mes*, OCorn. *mesen*, Bret. *mez* 'acorn, mast'.

(2) PIE **med$^{(s)}$-tu-* 'judgement' > **metsu-* > **messu-* > OIr. *mess*, W *armes* 'prophecy'.

(3) PIE **ḱh$_2$d$^{(s)}$-ti-* 'strong emotion' > **katsi-* > **kassi-* > Gaul. *cassi-*, OIr. *cais* 'love, hate', W, Corn., Bret. *cas* 'hatred'.

(4) PIE **u̯id$^{(s)}$-to-* 'seen' > **u̯itso-* > **u̯isso-* > CisGaul. *-uiśeos*, Gaul. *-uistus*, OIr. *·fess*, W *gwys*, Bret. *gous* 'was/is known' illustrates a participal formation from a verbal root in *-d*.

*3.9.3. *Ds > *ts > ss*

Proto-Celtic must have possessed a large number of examples of this change as part of the morphologically regular formation of the *s*-subjunctive of strong verbs, and of the related formation of the future (continued in Old Irish) and the desiderative (attested in Gaulish). Here I will limit myself to the subjunctive, which was formed by adding the suffix **-se/o-* directly to the final consonant of the root (Schumacher, 2004: 49–57). Due to the phonotactic constraints of Celtic, this then underwent various types of changes. In the case of dentals, the result in Proto-Celtic was *tau Gallicum*. Because of the fragmentary or relatively late transmission of most Celtic languages, evidence for this formation is only plentifully attested in Old Irish, but sporadic evidence is found in the other languages as well. Two examples will suffice as an illustration for Irish:

(1) PIE **ret-se/o-* > **retse/o-* > OIr. *ress-* 'may run'.

(2) PIE **g$^{u̯h}$ed-/se/o-* > **g^{u}etse/o-* > OIr. *gess-* 'may ask, pray'.

British and Continental Celtic only have residual examples of this category.

(3) A possible case in Gaulish is *scrisumio* 'that I may spit' < **skritse/o-* < PC **skrit/d-se/o-* (determining the precise Indo-European root is more challenging, cf. Darling, 2019: 133–137).

(4) In Celtiberian, *robiseti* 'may cleave (?)' may continue **bitse/o-* < **bʰidʰ-se/o-* from the root **bʰeid̯ʰ-* 'to split, cleave' (Schumacher, 2004: 224–225).

It is conceivable that this category was still productive in the ancient Celtic languages and that new finds of texts will furnish more examples in the future.

(5) In British Celtic, the category had been largely abandoned by the period of medieval attestation. The only lexical fossiles are MW *gwares* 'he may succour' and *ryres* 'he may run', two compounds of **ret-se/o-*.

*3.9.4. *zds > *ts > ss*

The only candidate for this highly specific development is **nezd-is-m̥h₂o-* > **nezdʲsamo-* 'nearest, next' > **netsamo-* > Gaul. gen. pl. *neđđamon*, OIr. *nessam*, W *nesaf*, Corn. *nessa*, MBret. *nessaff* 'nearest, next' (Jasanoff, 1988–90: 185). In theory, PC **nezd-tamo-* with a different superlative suffix would also be conceivable.

*3.9.5. *st (> *ts) > ss*

In the Insular Celtic languages, the development may have been directly from **st* > *ss* without *tau Gallicum* as an intermediate stage. Only a small number of – sometimes ambiguous – Gaulish and Lepontic examples give evidence of a stage with *tau Gallicum*. It may accordingly have been specific to these languages. Whether this stage involved an actual metathesis of **st* > **ts* is equally unclear.

(1) PIE **gʰosti-* 'stranger, guest' > **gosti-* > Lep. *-kozis* 'guest'. If the Lepontic letter *zeta* actually represents the assibiliated dental sound [tˢ], this would be the best evidence for the metathesis of **st* > **ts*.

(2) The Gallo-Gr. name Ατεσθας is probably a compound of the preverb **ati-* 'back, again, re-' < PIE **áti-* and **stă̆t-* 'one who stands' < PIE **st(e)h₂-t-*. The spelling *-σθ-* indicates that the medial cluster **-st-* was associated with *tau Gallicum* in Gaulish.

(3) If the identification of Gaul. *tuðdos/tuθos/tuso* with Lat. *furnus* 'furnace' is correct (see Delamarre, 2003: 304), then the word could be derived from **tutso-* < **t(o)-us-to-*, from the root **h₁eu̯s-* 'to burn'.

(4) PIE **upo-sth₂o-* 'the one standing underneath' > **u̯osto-* > **u̯otso-* > **u̯osso-* > Gall.-Lat. *uassus*, OIr. *foss* 'servant', W *gwas* 'boy, lad, servant', OCorn. *guas*, Bret. *gwaz* 'young man'.

(5) PIE **h₂u̯osto-* 'staying' > **u̯osto-* > **u̯otso-* > OIr. *foss* 'rest, position', W *gwas* 'abode, dwelling'.

(6) PIE **kᵘis-to-* 'seen' > **kᵘitso-* > OIr. *ad·cess* 'was seen' illustrates a participal formation from a verbal root ending in *-s*.

The British languages sometimes have *st* where the other Celtic languages have *ss*. Schrijver (1995: 410–430; 2022) explains such cases as continuing **-sst-*, *i.e.*, clusters that arose when a suffix with *-stV-* was added to stems or roots ending in *s* or, occasionally, a dental. The evidence for this morphological structure is circular, since it is only recoverable from the fact that the British languages have *st*. Possible instances are:

(7) qPIE **h₂u̯es-steh₂-* 'spending the night' > **u̯esstā-* > OIr. *fess*, but W, Bret. *gwest* 'night's stay, feast'.

(8) PC **kisstā-* 'woven basket' > OIr. *cess* 'basket, wickerwork', Lat. *cissium* 'a light, two-wheeled vehicle, cabriolet', borrowed from Gaul. **cission*, but W *cest*, Bret. *kest* 'basket, box' from the root PC **keis-* 'to do, bring, plait' (Schumacher, 2004: 391–393; see 3.9.6. (4)). It is not excluded, however, that this noun was influenced by, or even borrowed from Lat. *cista* 'box'.

*3.9.6. *s-s > *ss*

Although inherited clusters of two consecutive *s* were simplified to a single *s* according to a well-known Indo-European sound rule, where such a formation was synchronically transparent, double *ss* could be reintroduced in the individual Celtic languages.

(1) This is directly seen in Gaul. *pissiumi* 'I shall see', which synchronically consists of the Gaulish verbal root *pis-* < PIE **kᵘeis-* 'to perceive' + the desiderative or future morpheme **sie/o-*.

The subjunctive and future stems of most roots where this rule could apply (PC **gus-* 'to choose' < **ĝeu̯s-*, **kᵘis-* 'to see' < **kᵘeis-*, **tau̯s-* 'to be silent' < **th₂eu̯s-*, **u̯os-* 'to spend the night'< **h₂u̯es-*, all after Schumacher, 2004) have been rebuilt in such a way that the phenomenon can no longer be observed in Irish.

(2) Only stray examples with *s* remain in the case of of *ad·cí* 'to see' < **kᵘis-*, *e.g.*, OIr. 3sg. *·accastar* 'is seen; may be seen', or future forms such as 1sg. *·accus* or 3pl. *at·chichset* < **kᵘikᵘis-se/o-*.

(3) Schumacher (2004: 525–527) sets up a root PC **φlus-* 'to drink', which he derives from **?pleu̯s-*, tentatively explained as an extended form of PIE **pleu̯-* 'to swim, float'. The subjunctive stem of this verb in Old Irish is *lós-* and *lús-*.

(4) Finally, Schumacher (2004: 391–393) sets up the root PC **keis-* 'to make, accomplish, create' (see also 3.9.5 (8)). No present stem of this verb is attested in Old Irish; the subjunctive stem *céiss-* 'may plait, make, bring' and the future stem *cichis-* appear to exhibit *-s-* originating from double **s-s*.

*3.9.7. 'Original' *ss*

The etymological origin of *tau Gallicum* is not always clear. Occasionally its origin may have been onomatopoetic. A case in point is Gaul. *bussu-* if it means 'kiss' or 'lip', which seems to have a parallel in Irish (*Bérla na Filed*) *bus*, *pus* 'lip' and in Southern German *Busserl* 'kiss'.

3.10. Stops plus **n*?

Sequences of stops + **n* have been claimed by some scholars to be a source for geminate stops ('Kluge's Law'; *e.g.*, Stokes, 1891; Stokes, 1891–3, Zupitza 100; VKG i 158–161; Lühr, 1985; Bammesberger, 1998; see the section on previous research in the introduction). It is not possible in the context of this study to discuss all the proposed items (for example, Lühr, 1985 discusses 72 relevant forms). Suffice it to say that the proposals are rarely convincing (cf. GOI 92–93), that they can be accounted for in alternative ways, or that they face counterexamples (see the discussions by Kuryłowicz (1957: 131–132) and Strachan (1891–4)). For almost every geminate stop, a combination of that same simple stop + *n* can be found that must be reconstructed as such for Proto-Celtic and which therefore runs counter to the predictions of Kluge's Law in Celtic.[2] I will illustrate these general objections with only a few examples.

(1) OIr. *ette* 'wing' has been frequently traced back to **petniio̯-*, but this is flatly contradicted by OIr. *én*, W *edn*, OCorn. *hethen*, MBret. *ezn*, ModBret. *evn* 'bird', which can only continue **φetno-* < **petno-*. OIr. *ette* may rather come from a formation with a different suffix, *e.g.*, **φetantio̯-* < **pet-n̥t-* 'flying'; cf. OIr. *ethait* 'bird' < **φetantī-* < **petn̥tih₂-*. OBret. *attanoc* 'winged, flying' is not an example of geminate *-tt-* (this would have given **th*!), but rather a variant spelling for *-d-*, cf. ModBret. *adaneg* 'winged' < **φatanāko-* < **pₔt-n̥-*.

(2) Gaulish attests to countless examples of the patronymic suffix *-ikno-* 'son/daughter of' < **i-kn-o-*, whose semantic core is the PIE root **ken-* 'to originate from'. The common ancient Celtic onomastic suffix **-kko-* has nothing to do with this, but is rather due to 'onomastic' gemination of the common adjectival suffix **-ko-*, in analogy to other onomastic suffixes with geminates, such as *-illo-* < **-ilno-* or *-ullo-* < **- ulno-* (see chapter 9.).

(3) On a related note, the complex suffix **-o-ĝn̥(h₁)-o-* 'born from' regularly develops to **-agno-* in Goidelic. This is attested in ogam inscriptions as *-AGNI*, and as the common suffix *-án* in Old Irish, not ***-aggo-* > OIr. ***-⟨ac⟩* = -/ag/, as would be predicted by Kluge's Law. I do not enter the vexed question whether **-agno-* is regularly reflected as *-an* in British Celtic, or whether that suffix has a different origin.

(4) Stüber (1998: 113) derives OIr. *derucc* 'acorn' from a preform **deru-kn-on-*, ultimately a compound of **deru-* 'oak' and the stem **knū-* 'nut'. However, **-kn-* > **-kk-* is not the only unexpected development required for this explanation to work: **-eru-* should result in **ir*, and the transformation of **knū-* into an *n*-stem **kn-on-* is to my knowledge unparalleled. In view of this cluster of necessary *ad-hoc* assumptions, this word cannot be used as an example to demonstrate a regular sound development.

4. Etymological gemination: developments across branches with divergent outcome

A small number of simplifications or assimilations of clusters led to rather different outcomes in the sub-branches of Celtic, or they occur so sporadically that they are better classified as isolated developments rather than proper sound changes.

4.1. **zd*

Proto-Celtic **zd* continues the rare PIE cluster **sd*. No evidence is known how this cluster was treated in Celtiberian. The ultimate outcome of **zd* in Old Irish is ⟨t⟩ = /d/, apparently *via* an intermediate stage **dd*. How early this assimilation occurred is unclear. In the case of **zg*, the spelling TASEGAGNI, probably for OIr. *Tadgán*, in the ogam inscription I-KIK-002 (= CIIC 28) seems to indicate the retention of the sibilant until very recently before Old Irish. The ultimate outcome of **zd* in the British Celtic languages is **θ* (W ⟨th⟩, Corn. ⟨th⟩, Bret. ⟨zh⟩), which is also the regular outcome of **tt*. A stage **dd* like in Irish is precluded since that yields *d* in the British Celtic languages. One way of explaining the British development is to assume that **zd* passed through the intermediate stage **tt*, even though the required change **zd* > **tt* is untrivial. It finds potential support in Gaulish where the stage **tt* appears to be attested in Gall-Lat. *petia*, *pettia* 'piece', which can hardly be separated from PC **kᵘezdi-* 'part, share', and in Gall-Lat. *bottia* 'blister, bump, boss' and **botto-* 'knob, button', which underlies various words in the Romance languages and which could be cognate with OIr. *bot* 'tail, penis'. Isaac (2004: 74–76) proposes a different pathway from PC **zd* to Brit. **θ* and the Gaulish reflex. Schrijver (1995: 376) is agnostic about the intermediate stages in the development. Martin Kümmel (in peer-review) suggests that the development **zd* > **dd* > **tt* in British and Gaulish was earlier than the rise of younger **dd*, which was not subsequently devoiced (see section 6.2.). He points out that devoicing of geminates is a rather natural change, and occasional examples of a voiceless outcome of originally voiced geminates (discussed in 6.2) could reflect that same development and might be older than those that are attested with voiced outcomes.

(1) PIE **ni-sd-o-* 'place to sit down' > **nizdo-* 'nest' > OIr. *net*, W *nyth*, OCorn. *neid*, Corn. *nyth*, OBret. *nith*, Bret. *neizh*.

(2) PIE **bʰrosdʰo-* > **brozdo-* > OIr. *brot* 'goad, spike', W *brath* 'biting, stinging; prick'. The vocalism of the Welsh word is unclear.

(3) PIE **gu̯osdo-* 'tail' > **gᵘosdo-* > **bozdo-* > Gall-Lat. *bottia* 'blister, bump, boss', OIr. *bot* 'tail, penis', W *both* 'nave,

[2] A noteworthy side aspect of Lühr's study (1985: 300–302) is that she sets up geminate **pp* for Proto-Celtic, which, unlike its simplex counterpart, she assumes to have been preserved. This differential treatment has an exact parallel in Japanese. A potential candidate is Gaul. **lapparo-*, which can be postulated as the source of Fr. *lapereau* 'bunny', against Fr. *lievre* 'hare' < Lat. *leporem*; another example is **kappilo-* (see section 13.1. (18)), to which Insular Celtic words for 'horse' go back.

hub, boss of a shield', Corn. **both* 'hillock (in placenames)'. However, the precise relationship of these semantically diverse words remains to be clarified.

(4) PIE **ĝʰasdʰo-* > **gazdo-* > OIr. *gat* 'withe, osier'.

(5) PC **kᵘezdi-* 'part' > Gall-Lat. *pettia* 'piece; share of land', OIr. *cuit* 'share, part', OW *ped*, W *peth*, MCorn. *peth*, *peyth*, *pyth*, Bret. *pezh* 'thing', Pict. *pet-* 'share of land'.

(6) qPIE **masdi̯o-* > **mazdi̯o-* > OIr. *maite* 'stick, staff'. This may be a loan from a pre-Indo-European substrate language.

(7) PIE **rasd-* 'to scratch' > **razd-* > W *rhathu*, Bret. *razhañ* 'to scratch, rub off'.

(8) PIE **su̯isde/o-* 'to whistle' > PC **su̯izde/o-* > W *chwythu* 'to blow, breathe', LCorn. *hụetha*, Corn. *hwytha*, Bret. *c'hwezhañ* 'to breathe'; and the nominal formation PC **su̯izdo-* > W *chwyth*, Corn. *wheth*, Bret. *c'hwezh* 'breath, breeze'. OIr. *fet* 'whistle' could belong here as well if the lenited anlaut *f-* < **su̯-* became generalised instead of expected *s-*. Alternatively, *fet* can come from **u̯into-* < **u̯ītnto-* < **u̯eh₁nto-* 'wind' (see section 7.2.).

(9) PC **trozdi-* 'starling' > OIr. *truit* 'starling'. The cognates W *trydw*, OCorn. *troet*, LCorn. *trodzhan*, OBret. *trot*, MBret. *tred* 'starling' are possibly loans from Irish (IEW 1096; *pace* Stifter, 2021: 174), unless the Insular Celtic languages continue **troddi-*, which is unsatisfactory with regard to the Germanic cognates ON *þrǫstr* < **þrastu-*, OEngl. *þræsce* 'thrush' < **þra(st)skōn-*, as well as cognates in other European languages. This item is suspect of being a prehistoric substratal loan into several European languages.

4.2. **χt*

A handful examples appear to attest to the irregular assimilation of PC **χt* (of diverse origins) > **tt*, instead of retaining them as two separate segments as is the norm. Since examples for this assimilation are isolated across the family and within the languages, it is best to view them as sporadic, perhaps dialectal, simplifications of the clusters, the motivation for which remains obscure.

(1) The name of the Ibero-Celtic people *Vettones* in Central Spain can be understood as **u̯eχt-on-*, a form with the individualising suffix *-on-* built on the stem seen in Gaul. *Vectirix*, *Vecturius*, OIr. *fecht*, MW *gweith* 'journey, time', Corn. *gwyth*, Bret. *gwezh* 'time' < **u̯eĝʰ-to-* (Prósper, 2005: 305–309).

O'Brien (1954; 1956) has drawn attention to three Irish words that appear to exhibit a similar behaviour:

(2) OIr. *utlach* 'lapful' stands beside *uchtlach*, synchronically derived from *ucht* 'lap' < PIE **pektu-* 'breast'. It is conceivable that the internal *-ch-* was lost through dissimilation against the final *-ch*.

(3) OIr. *littiu* 'gruel, porridge' stands beside *lichtiu* and W *llith* 'food, sustenance' < **liχti̯on-* (of uncertain origin). Matasović (2009: 135) regards the form with *-tt-* as primary and reconstructs PC **φlittV-*, which he compares with Lat. *puls*, *pultis* 'porridge' < **polt-/ pl̥t-*.

(4) OIr. *aittenn* 'furze, gorse' can be reconstructed as **attīno-*, but its cognates W *eithin*, OCorn. *eythinen*, OBret. *ethin* continue **aχtīno-* < PC **ak-tīno-*, perhaps 'sharp tree', from the root **h₂eḱ-* 'sharp, pointed'. McCone (2005: 409) compares this with Basque *ote* of the same meaning and suspects a borrowing from an unknown source for all languages.

5. Etymological gemination: morphological gemination in derivation and inflection

Many instances of gemination treated above (*e.g.*, **ln* > **ll*, **sm* > **mm*, **sn* > **nn*) occur across morpheme boundaries and are in fact indirectly the result of processes of word formation. Evidence for a few more such changes arising through derivation or inflection is usually restricted to a single branch within Celtic. Therefore, it will be apposite to discuss them separately.

5.1. **n-n* > **nn*

A group of words with *nn* in the Celtic languages are best analysed as derivates with a nasal suffix from roots ending synchronically in *n*. Although the phenomenon appears to be pan-Celtic, the evidence for it is often confined to a single language or two languages at best. It is noteworthy that the resulting sequence *-nn-* was not simplified, in contrast to what may have been the case with early instances of **-mm-* > **-m-* (see 6.3. (2)). Perhaps these words were formed at a time when geminates were already established as a phonological class in the language.

(1) The double *nn* in the Gaulish personal names *Adgennos*, *Congennolitanos*, *Adgonnetios* has in the past either been ignored or explained as a sporadic gemination of **-geno-* < **ĝenh₁o-* 'born' in personal names. However, onomastic gemination is typically coupled with shortening, but none of these names show shortening. Furthermore, all examples of *-nn-* are found in compounds with preverbs as first members, whereas compounds with other elements, *e.g.*, Cib. *mezukenos*, Gaul. *Medugenos* 'mead-born', *Rextugenos* 'law-born', OIr. *Muirgen* 'sea-born', *etc.*, never show gemination. The two phenomena seem to be causally linked. I propose to explain *-genno-* as a thematic derivative of the Celtic neuter *n*-stem **genen-* 'birth, origin' (cf. OIr. *gein*), namely **-gen-n-o-* 'having birth, origin'. In this interpretation, *Adgennos* can be understood as meaning 'having birth into (a specific family), having been born to (a family)', and *Congennolitanos* as 'being wide-ranging with regard to relations'. PC **genen-* itself continues earlier **genmen-* with loss of *m* in heavy clusters such as **genmn-*; **genmen-* in turn is simplified from PIE **ĝenh₁men-* (NIL 153).

(2) Gaul. *linna* 'coat', OIr. *lenn*, W *llen*, OCorn. *len* 'coat', OBret. *lenn* 'piece of cloth' all go back to PC **linnā-*. The proposed etymology as **φlitnā-* 'spread out (cloth)' (Delamarre, 2003: 203) is unsatisfactory since **tn* does not otherwise assimilate to **nn*. Perhaps it is to be analysed as **lin-neh₂-* 'made of linen' (cf. Gr. *λίνον* 'linen', OCS *lьnъ* 'flax') or as

**liHn-neh₂-* with Osthoff shortening before the geminate (cf. Lat. *līnum* 'flax, linen' and OIr. *lín*, unless the latter is borrowed from Latin).

(3) The Gaulish name element *rinno-* of unknown meaning, the rare Irish adjective *renn* 'swift, hasty', and W *rhyn* 'rigid, stiff, brave, rough, cold' may form an equation, although the divergence in meaning renders this less than obvious. If the Irish adjective represents the original meaning, one can think of a formation from the root **h₃reiH-* 'to whirl', perhaps from a hypothetical nasal-infixed verb **rinati* 'to run', comparable with Goth *rinnan* < **h₃ri-n-H-*, with the addition of the adjectival suffix **-no-* to the present stem. Alternatively, the formation could derive from the noun **reino-* 'great flowing mass of water' < **h₃rei(H)-no-*, with ablaut and with the addition of a second nasal suffix. However, both alternatives are morphologically highly speculative.

(4) The Celtiberian name *Stennoco* (in Latin script) is a derivative of the shape **sten-n-o-* from the *on*-stem name *stenū* < **stenōn* with a single *n* (Jordán Cólera, 2019: 613–614). The name ⟨stena⟩ in Celtiberian script could likewise stand for **stennā*, since the vernacular script does not graphically mark geminates. This development is parallel to that of the genitive *abulos* < **abul-nos* in section 2.1.

(5) The first element of Gaul. *sonnocingos* has been thought to contain the oblique stem of the Indo-European word for 'sun'. If this doubtful proposal is right, the structure of the word must be something like **su(h₂)n-no-*, *i.e.*, a derivative with a nasal suffix. However, this will not explain the vowel *o*.

(6) Old Irish has two distinct verbs that influenced each other formally: in PIE **sn̥-n(e)-h₂-* 'to obtain', a nasal-infix present of a root containing **n* led to geminate **nn* in **sanna-* → **sanne/o-* > OIr. *seinnid*. Schumacher (2004: 558) defines its meaning as 'to attain, achieve, procure'; in eDIL (https://dil.ie/36945), the meaning is given as 'to strike', with many examples that are classified as "obscure". Although entirely unrelated, PIE **su̯n̥h₂-* 'to sound' > **su̯ana-* acquired geminate **nn* probably after its model, *i.e.*, **su̯anne/o-* > OIr. *seinnid* 'to sound, make music'; cf. MW *honni* 'to assert, proclaim, make known' from the iterative formation **su̯onn-ī-* (Schumacher, 2004: 608).

(7) In British, the same development underlies the verb PIE **tn̥-nu-* > **tannu-* > W *tannu* 'to extend', Corn. *tan* 'take!' (Schumacher, 2004: 618–619).

5.2. Other examples

It is conceivable that (very rare) combinations of nominal stems ending in **-b* + the athematic dative and instrumental plural endings **-bis* and **-bos* may have led to geminate **-bb-* within nominal paradigms, but no such examples are attested from ancient Celtic. Probably there were other, rare contexts of a similar character in which geminates arose through regular processes of inflectional and derivational morphology.

6. Etymological gemination: morphological gemination in composition

Compounding, that is the univerbation of two or more lexical items to form a new lexical item that is more than just the sum of its constituent parts, is an important process of word formation of ancient and medieval Celtic languages. When in the course of compounding two consonants come into contact with each other, new geminates can emerge, either because the consonants were identical or similar already before, or because assimilation takes place.

6.1. *dzd* >*dd*?

In Stifter (1998: 212–218), I proposed to explain PC **ruddo-* 'rust', reflected in W *rhwd*, Corn. *-roade*, *-rode* in placenames, OBret. *rod* 'rust; mud, dirt', either as a **h₁rudʰ-dʰh₁ó-* 'that which makes redness' or as **h₁rudʰs-dʰh₁ó-* > **rudzdo-* > **ruddo-*, where the first element is an *s*-stem with zero-grade in root and suffix and with subsequent simplification of the cluster **dzd*. The alternative proposal by Hill (2003: 196–202) < **h₁rudʰ-sd-ó-* with the root **sed-* 'to sit' is phonologically equivalent to the second option. If one of these proposals is correct, note the different treatment of **dzd* from PC **tst* > **ts* (see section 3.9.1.–section 3.9.2.), which could have to do with the difference in voice, or is due – in this isolated example – to some kind of analogical influence (Martin Kümmel in peer-review). Schaffner (2016–7: 114–5), on the other hand, reconstructs **h₂ru-ti-* 'redness', which does not involve a geminate.

6.2. Assimilations across morpheme boundaries

A special source for geminates are clusters where the first stop assimilates totally to the articulation of the second. Such assimilations can only be found across morpheme boundaries. Most of the following examples relate to phonotactic phenomena between lexically distinct items (preverbs and nouns) where it can be argued that external sandhi applies (Martin Kümmel in peer-review). It remains to be investigated in future research if there is a difference in outcome between clusters that result from composition and those that arise from derivational suffixation or whether it is possible to assign conflicting results to full vs. incomplete univerbation.

The simplest examples of assimilation are those where the involved sounds are already identical from the outset. Relevant examples are Gaul. *readdas* 'has given', perhaps from **φro-ad-dast*, and the very similar OIr. *do·rat* 'has given' < **to-φro-ad-dast*, both ultimately containing a sequence of the preverb **ad-* and the PIE root **deh₃-* 'to give'. Welsh provides the examples *adynu* 'to suck' < **ad-dina-* (Schumacher, 2004: 274), *adygaf* 'to take, seize' < **ad-duke/o-* (Schumacher, 2004: 287), and perhaps the Welsh adjective *edifar* 'regretted, deplored' < **ad-dī-maro-* (cf. Jackson, 1953: 427). In these cases the synchronic outcome of the geminate in Welsh is an unlenited voiced stop. The late by-form *addygaf* /aðəgav/ of *adygaf* /adəgav/ owes its lenited *d* to the influence of other compounds of **duk-* after leniting preverbs, *e.g.*, MW *dyđwc* 'to bring, carry' < **to-duk-*, *ymđuc* 'to conceive, bear' < **ambi-duk-*, *kyfyrđuc* 'to remove' <

**kom-ro-duk- etc*. (Schumacher, 2004: 287). Such reanalyses are found elsewhere as well: the compound *atygaf* 'to bring back' (*e.g.*, MW *attwc*) < **ad-ddug-* < **ati-duk-*, with the different preverb *ad-* < **ati-*, received the reanalysed by-form *ad-ddygaf*. W *addysg* 'education, learning' is either a formation within Welsh or, if it ultimately continues Lat. *addiscō* 'to learn in addition', it may have undergone the same analogical lenition as *addygaf*.

In the majority of instances, however, the consonants across morpheme boundaries are different and some sort of assimilation takes place, either assimilation in voice or in articulation, or in both. This is readily observable in verbal and nominal compounds where the first element is a preverb ending in a stop, in particular **ad-* and **eχs-*. The latter occurs in the *s*-less allomorph **ek-* in at least a subgroup of such formations (see the discussion in Russell (1988: 118–121) as to how and why **ek-* and other preverbs lost their **-s*). Examples come from all branches of Celtic. However, the way how cross-morphemic clusters are treated differs in complex ways, not only between the branches, but sometimes even within a single language. This difficult matter deserves a detailed discussion.

It seems that already in Proto-Celtic there was full assimilation to the second stop when it was voiceless. An equation involving three branches of Celtic is PC **ad-trebā-* > **attrebā-* 'abode, dwelling', reflected in OIr. *attrab*, W *athref*, and in the Gaulish ethnonym *Atrebates*; cf. OIr. *ad·treba* 'to inhabit' where the two elements are separated by a strong morphemic boundary. This contrasts with the adverb W *adref* 'homewards' that belongs to a younger stratum. *Pace* GPC, which calls its initial *a-* an "elfen anhysbys" (obscure element), it is evidently the old preposition **a* < **ad* 'to' (identical with the sentential connector a^2) + **trebā-* 'home', corresponding to Bret. *a-dreñv* 'behind'. Even though *adref* contains the same etymological elements as *athref*, it must have been univerbated from a prepositional phrase when **ad* had lost its final consonant and consequently had acquired a leniting effect. An equation encompassing Gaulish, Galatian and Irish is CisGaul. *Ateporix*, Galatian *Ατεπορειγος* and OIr. *attach* 'refuge' < **ad-tekᵘo-* 'running to, refuge'. Several other examples demonstrate that full assimilation and, most likely, gemination were in fact the rule in Gaulish in such contexts, namely *appisetu* 'let see' < **ad-kᵘise-*, the names *Aclutius* < **ad-klut-i̯o-* 'very famous' and CisGaul. *akluśamoualos* < **ad-klut-(i)samo-u̯alo-* 'most famous ruler', and Lat. *attegia* 'cabin, hut', which was probably borrowed from Gaul. **attegi̯ā-* < **ad-teg-i̯ā-*. Old Irish furnishes numerous additional instances of voiceless geminates resulting from assimilation to a second element with a voiceless stop, *e.g.*, *frettech* 'fore-swearing' < pre-OIr. **frith-tech-*, verbal noun of *fris·toing*, *accobar* 'desire, wish' < **ad-kuφro-*, verbal noun of *ad·cobra*, and many others. Examples in British are not quite as numerous, but still unambiguous, *e.g.*, W *achas*, Corn. *ahas* 'keen, severe, cruel', which correspond to OIr. *accuis* 'hate, enmity' < **akkassi-* < **ad-kad-ti-*; and W *achan* 'whisper, murmur' < **ad-kanV-*, W *athwll* 'perforated' < **ad-tullo-*.

The behaviour of the preverb **ek(s)-* is special. The expected development in front of a voiceless consonant is that the first **k*, *via* the intermediate stage **χ*, was either lost or assimilated to the **s*. This is shown by the agreement of Gaul. *escingo-* 'warrior, infantrist' < **eχs-keng-* 'striding out' (but cf. also *extincon* of unknown meaning) and OIr. *escarae* 'enemy' < **eχs-karant-* (however, its lack of syncope points to a recent date of formation); cf. also OIr. *sesca* 'sixty' < **su̯eks-kont-*. Celtiberian provides the comparable example *eskenim*, perhaps 'foreigner', but in this case in front of a voiced stop, if it goes back to **eχs-gen-i-*. For the position before **t*, a phonological parallel is provided by **trek-stu-* > **treχstu-* > **trestu-* > OIr. *tress* 'contention, fight', W *tres* 'battle'. However, in young, as it were *nachgrundsprachlich*, compounds, it seems that the consonantal part of the preverb assimilated totally to the following consonant, perhaps in parallel with the treatment before voiced consonants (see the following paragraph). This may have occurred separately in the already differentiated Celtic branches after the break-up of Proto-Celtic. Old Irish has only three relevant examples: *etaim* 'chance, opportunity (?)' < **eχs-tud-sman-* or **eχs-dī-tud-sman-* 'act of falling out', *ettech* 'refusal' < **eχs-teg-o-*, verbal noun of *as·toing* 'to refuse', and *etal*, *etail* 'pure, sinless' < **eχs-tol-o/i-* 'being outside of desire'. The latter has a parallel in W *ethol* 'chosen' < **ettol* < **ek-tol-*. This treatment of **k-t* > **tt* across compositional boundaries contrasts with the regular development of pre-Celtic **Kt* > PC **χt* within simplex lexemes, and must be due to morphophonological analogy after the model of voiced sounds.

When the second element was a voiced stop, the outcome differs decisively between the languages and no comprehensive picture emerges. Old Irish shows full assimilation in such cases. Verbal and nominal compounds provide countless examples; it will suffice to mention a few representatives, *e.g.*, OIr. *·eipir* 'says' < **eχs-ber-* or *·acair* 'sues, accuses' and *acrae* 'act of suing, bringing an action' < **ad-gar-*. In order to explain the difference from how **eχs-* is treated before voiceless consonants in old formations seen above, the development before voiced consonants may have been the following: **eχs-b°* > **eγzb°* > **eγb°* with loss of 'sandwiched' **z* and with subsequent assimilation to **ebb-*. As OIr. *naidm* 'act of binding, bond' < **nad-man-* and *maidm* 'act of breaking' < **mad-man-* demonstrate, **d* did not assimilate regularly to a following **m*. However, in analogy to the verbal compounds cited above, **dm* did become **mm* in verbal compounds with *ad-* as first preverb, *e.g.*, the verb *ad·midethar* 'to aim at, evaluate' has the prototonic stem *·aimdethar* and the verbal noun *ammus* < **ad-med-*. In the verb *fo·ammámaigedar* 'to subjugate', which is restricted to the Würzburg and Milan Glosses, the productive character of the secondary assimilation **dm* > *mm* is synchronically observable. The verb has been created from the elements *fo-* 'under', *ad-* 'at' + *mám* 'yoke' in order to calque Lat. *subiugāre*. Its verbal noun *foammámugud* underlines the non-inherited nature of the formation: contrary to what would regularly be the case, the vowels *o* and *a* have not been elided or merged in it, but sit uneasily side by side of each other. In Stifter (2017b: 221) I tentatively suggested that, in contrast to **dm*, **tm* may have regularly given **mm* in Celtic, in order to explain OIr. *amm* 'time' < **amman-* < PIE **$h_2$et-men-* 'the act of going around', but alternative explanations are available for this word, for which see the cited article. The case of the compound **eχs-med-* is ambivalent. From

its prototonic stem a neo-simplex verb *éimdid* ‘to refuse, reject’ was created, which is frequently written with *mh*, indicating lenition of the *m*. On the other hand, the further-derived compound *fo·émid*, *for·émid* ‘to be unable, fail; refuse’ never shows such a spelling and is in fact once transmitted with a double *mm*. One of the two stems must have undergone analogical change.

The evidence is not so resounding in sheer numbers in British, and it seems to differ between the classes of sounds involved and according to the chronological layers of word formation. Clear examples of full assimilation are found for the preverb **ad-* followed by *b-*. W, OCorn., Bret. *aber* ‘river-mouth’ < **abber-* < **ad-ber-* is reflected the same in all three languages and must be an old formation; W *aberth* ‘sacrifice’ < **ad-ber-tā-* and W *abwyd* ‘bait, lure’ < **ad-* + *bwyd* ‘food’ show the same treatment. However, the evidence for *-d* followed by a *g-* is ambivalent. Some examples with **ad-* in intensifying and preverbal function display the plain simplification of the geminate: W *agos*, OBret. *ocos*, Bret. *hogos*, OCorn. *ogos* ‘near’ < **aggossu-* < **ad-gostu-* ‘at hand’ (Hamp, 1981), W *agarw* ‘rough, stern, bitter’ < **ad-garu̯o-*. That these are in all likelihood old formations is underlined by their precise cognates in Irish, namely OIr. *acus* ‘near’ and *acarb* ‘very rough’.

Other examples are restricted to Welsh: W *agwrdd* ‘strong, mighty’ (related to Hisp.-Lat. *gurdus* ‘dolt’?); W *agwedd* ‘manner, fashion’ (beside *gwedd* ‘appearance’ < **u̯idā-* ‘look’) can only have been formed after initial **u̯-* had developed into a stop. On the other hand, a handful of examples appear to illustrate first assimilation of **dg* to **gg* and then fricativisation to **χ*, effectively making it fall together with the outcome of **kk*. These are the hapax W *achwir* ‘true, genuine’, ostensibly a compound of **ad-* + *gwir* ‘true’, W *achlan* ‘all, total’, if it is from **ad-* + *glan* ‘clean’, and W *achwre*, *achre* ‘part of a roof or fence; covering’, perhaps from **ad-* + **u̯regi-*, related with OIr. *fraig* ‘wall’. Again several of these words feature *g-* that arose word-initially rather late before PC **u̯-*. Compounds of the root PC **gab-* ‘to take’ do not provide conclusive evidence. Schumacher (2004: 321) argues that the inherited stem was replaced by **kab-* in British, *e.g.*, W *dyrchafael*, Corn. *drehevel*, *derevel* ‘rising, ascending’ continue **to-ro-ud-kab-aglā-* rather than **-ud-gab-*.

The outcomes are also rather diverse for the preverb **eχs-*. It is evident that some of the developments must be due to analogy, but it is difficult to determine with certainty what the regular inherited treatment was. Some examples suggest that in such cases the outcome was a fricative, *e.g.*, W *differ*, MCorn. *difres* ‘to defend, protect’ < **dī-eχs-ber-*, and W *dichlyn* ‘to choose, pick’, Bret. *dilenn* ‘to choose, select’ < **dī-eχs-glenn-* (cf. OIr. *as·gleinn*, *·eclainn* ‘to examine’ < **eχs-glenn-*; see Hamp, 1974c and Hamp, 1974d). Fleuriot (1964: 138) cites OBret. *diclinatuiu*, gl. *legendae* ‘to be chosen’, without indicating the source. The OBret. spelling ⟨cl⟩ is ambiguous and could stand for /χl/ or /gl/. I regard this verbal adjective as the Old Breton representative of ModBret. *dilenn*. Fleuriot, however, explains it as the equivalent of OIr. *do·gleinn* ‘to select, collect, gather’ < **dī-glenn-*. This matter remains ambiguous.

Even though I have been writing these reconstructions with a medial *-s-*, chiefly for reasons of etymological transparency, it is not certain that the *-s-* was retained or, if it was, if it played a role in the sound changes. Schrijver (1995: 376; 1999: 2), on the other hand, suspects that the development in such contexts was “**-ksb-* > **-xsb-* > **-sb-* > **-hb-* > **-p-* > **-f-*” (and analogically for **-ksg-*), *i.e.*, that *-s-* played a crucial role. However, in view of W *achlan* < **agglan* < **ad-glano-* *etc*. mentioned above (if that is the right explanation), I consider the possibility that these British forms involved the asigmatic allomorph **ek-* of the preverb (see the following paragraph), with full assimilation to the initial of the following element, *i.e.*, **dī-ek-glenn-* > **dīgglenn-* > **dīχlenn-* *etc*. In view of *aber* < **abber-* < **ad-ber-*, it must then be assumed that the case of the labial in W *differ* < **dībber-* < **dī-ek-ber-* is analogical to the velar. Another piece of evidence for the development of **eχ(s)-g°*, unfortunately not shedding light on the crucial intermediate stages, is MBret. *elas* ‘gizzard, liver’, corresponding to OIr. *eclas* ‘gizzard, oesophagus, stomach’ < **eχ(s)-glasso/ā-* (Hayden & Stifter, 2022). Because of the regular disappearance of the velar before *l* in Breton, it permits no inference about the precise treatment of the cluster.

That there existed an asigmatic allomorph **ek-* in British Celtic, or at least in the stage immediately preceding Welsh, is evident from other formations. This allomorph is most conspicuously visible in W *eglwg* ‘conspicuous, visible’ < **ek-luko-* from the root **leu̯k-* ‘to become bright’. Moreover, a series of words starting with *e-*, all exclusive to Welsh, can conceivably be analysed as compounds with **ek-*, giving evidence of a sequence of reanalyses of such formations. They were therefore probably only formed productively within the Welsh language during the historical period. They include *eglan* ‘sea-shore’ < **egglannā-* < **ek-* + **glannā-* (*glan* ‘shore’), *egwan* ‘very weak’ < **ek-* + **u̯anno-* (*gwan* ‘weak’), as well as the evidently late neologisms *egwyl* ‘respite’ (16th–17th century, beside *gŵyl* ‘holiday’) and *egwal* ‘cabin’ (18th-century, beside *gwâl* ‘lair, den’); furthermore *eban* ‘feeble’ beside *ban* ‘top’ and *edif* ‘greedy’ beside *difiog* ‘voracious’ (Russell, 1988: 120). At first glance, the first of these appear to contradict the claim above that **ek-g°* became **egg°* and ultimately **eχ°*. However, since these formations are late and secondary, various analogical steps are involved. *Eglwg* ‘conspicuous’ was synchronically analysable as consisting of *eg-* + a lenited allomorph of **llwg* ‘light, visibility’. *Eglan* *etc*. can therefore have been formed as compounds of the same *eg-* + *-lan*, the lenited allomorph of *glan*, *etc*. Since, on the other hand, the relationship of the simplex *glan* to the compound *eglan* could also be analysed as that of a prefix **e-* + unmutated base *glan*, the way was then free to form *e-ban* from *ban* and *e-dif* from *dif°*.

The situation is very different in Gaulish. In the certain instances of the preverb **ad-* followed by **b-* or **g-*, assimilation is typically lacking, *e.g.*, *adgarios*, *adgariontas*, *adgarie* < **ad-gar-* (contrast OIr. *acrae* ‘act of suing, prosecution’, *·acair* ‘to sue, prosecute; bewitch’ < **aggar-* of the same structure),

Adbugiounus < **ad-bug-*, *Adbogius* < **ad-bog-* (contrast OIr. *apach* 'corpse, remains' < **abbo̯u̯g-*, possibly of the same structure). Cisalpine Gaulish shows the same treatment if *aśkoneti(o)* represents the name written *Adgonnetius* < **ad-gonn-et-* in the Latin script. Assimilation is likewise missing between *d* and *m*, as the name *Admina* (*aśmina* in the Lepontic script) attests. The name *Annamat(i)us* and the placename *Annamatia* are undoubtedly to be connected with the name that is more commonly written *Adnamat°* 'against the enemies', but note that attestations for the assimilated variants do not come from the Gaulish heartland, but from 'marginal' areas such as Noricum or Pannonia where influence from other languages is conceivable. De Bernardo Stempel (2010: 69) cites the divine epithet *Agganaicus* < **ad-gen-ak̯o-* of Jupiter (Pavia, 2nd c. A.D.) as an example for the assimilation of *dg* > *gg*. Perhaps this reflects a late development in Cisalpine Gaulish, but the number of additional changes required for the etymology casts doubt on the relevance of this example. The preponderant lack of assimilation in most of the Gaulish evidence means either that assimilation of voiced clusters across the morpheme boundary was not a pan-Celtic phenomenon, or that these compounds were morphophonologically so transparent that the assimilation could be easily undone by reanalysis.

Its ambiguous writing system renders the pertinent Celtiberian evidence meagre and difficult to interpret. The gentilic name *abo[..]kum*, which has been emended to *aboiokum* on the basis of *Abboiocum* in a Latin inscription, has been explained as **ad-bog-i̯o-ko-* by Prósper (2005: 252–254). The verbal form *usabituz*, perhaps 'let him cut out', may continue **uts-ad-bi-tūd* (Schumacher, 2004: 226–231). While the deficiency of the Celtiberian script with regard to writing obstruent clusters leaves it undecidable if **d* has been assimilated to **b* in these cases or if it is just not graphemically expressed, the spelling *Abboiocum* in the Latin script indicates a genuine geminate resulting from assimilation, provided the etymology is correct.

The examples above are all taken from compounds where the identity of the involved elements is well established and would also have been easily recoverable for native speakers. This transparency of the elements may have entailed a special, analogical treatment across the morpheme boundary. It is therefore possible that geminates that are not transparently analysable as resulting from assimilation may show different outcomes. This is conceivably the case in the British Celtic languages, as Anders Jørgensen (personal communication) reminds me. As argued in the sections above, sometimes what must have been voiced geminates in prehistory are reflected by unlenited sounds, sometimes the outcome appears to be fricative sounds, at least in the case of **gg*, as if an intermediate stage had consisted of a voiceless geminate (cf. Russell, 1988: 115–125). W *achlan* and *eglan* show contradictory behaviour within a single language. The conditions for the divergent treatments are not always obvious. When the outcome of the geminate in Welsh is a single voiced stop, but lacks parallels in Breton and Cornish, it may have been formed more recently than words with parallels, where the outcome is a voiceless fricative.

At this point it is apposite to look at the treatment of other possible examples of voiced geminates, especially of **gg*, in the British languages, when they do not occur across transparent morpheme boundaries. They allow placing the ambiguity of the treatment of **gg* in a bigger picture.

(1) The change **gg* > **χ* may have a parallel in PC **biggo-* 'small' > W *bychan*, OCorn. *boghan*, Bret. *bihan*; the reconstruction with **gg* is necessitated by OIr. *bec*. The unsatisfactory alternative is to assume two different formations **biggo-* and **bikk(an)o-* for the two branches.

(2) The case of the word for 'bell' is uncertain. OIr. *cloc* presupposes **kloggo-*. W, OCorn. *cloch*, Bret. *kloc'h* could either continue that same preform, or they could be borrowed from MLat. *clocca*, which has a geminate **kk* (cf. also Fr. *cloch*, Germ. *Glocke*). The origin of this word is clearly sound-symbolic (see section 11.2.).

(3) Another item is equally ambiguous: W *gwraig*, OCorn. *grueg*, Bret. *gwreg* 'woman' speak in favour of a preform **u̯rakī* or **u̯rakū* with *k*. On the other hand, the spelling variants of the rare Irish word *frac*, *fracc*, *frag* 'woman' are most straightforwardly interpreted to stand for /g/ < **u̯raggā-* '(old) woman'. This interpretation finds support in Scottish Gaelic *fràg* 'a kind woman'. While the spellings *frac* and *fracc* could conceivably stand for /frak/ < **u̯rakkā-*, the spelling *frag* and ScGael. *fràg* would remain isolated in that case. A uniform explanation that accounts for all Gaelic forms is preferable. In addition, the British Celtic languages possess the by-form W *gwrach*, OCorn. *gruah*, Bret. *gwrac'h* 'old woman'. This could either reflect **u̯rakkā* with 'expressive' gemination of the **k* of **u̯rakī/ū* or, if my analysis above of *bychan etc*. as **biggo-* is accepted, it could reflect **u̯raggā-* and correspond directly to the Irish form. Phonaesthetically, *-ch* /x/ acquired a negative connotation especially in Welsh (Rodway, 2019; Wmffre, 2007: 59; Zimmer, 2000: 278). This may have been prompted by loans from Irish with their frequent suffix *-ach* that conferred a particularly alien and in consequence pejorative feeling on such words (Sims-Williams, 2011: 183–184). This has been suggested to explain the choice of the 'phonemaestheme' *-ch* and the semantics of *gwrach*. However, while this phonetic attitude is specifically Welsh, the negative connotations of *gwrach* appear to go back already to Proto-British, since it is equally attested in all three British languages. This detracts from the phonaesthetic explanation of *gwrach*.

(4) A few other examples are even less certain. The Welsh hapax sg. *gwre* 'worm, insect, mite' < PC **u̯rigā* < **u̯rg$^{(h)}$eh$_2$* stands in a suppletive relationship with pl. *gwraint* 'mites, worms', cf. OIr. *frige*, pl. *frigit* 'vermin, flesh-worm', Gallo-Lat. *brigantes* 'worms in the eyelid' < PC **u̯rigant-*. Breton has a related word ending in *-c'h*, namely sg. *gwrec'h*, pl. *gwrec'hent* 'mite'. The relationship between the Welsh and Breton words is reminiscent of that of W *lle*, Corn. *le*, OBret. *le* 'place' (< **legā*? cf. Schrijver, 1995: 308) and OBret. *legh*, Bret. *lec'h* 'place'. Several explanations are possible. The forms with the fricative could continue a sporadically geminated

*-*gg*-; or they could continue by-forms with suffixal *-*s*- of unclear function added to the stem, *e.g.*, *$*u̯ri̯χsā$ and *$*leχsā$. Finally, the *-c'h* of Breton could be due to a phonetic strengthening in sandhi of final *-γ > *-χ in Late Proto-British *$g^{u}re\gamma$ and *$le\gamma$, comparable to the strengthening of *-*h* > *-χ seen in Bret. *dec'h* 'yesterday' < *deh < *$\gamma des(i)$, in contrast to W *doe* (Schrijver, 1995: 390). The latter appears to be the simplest explanation. It only entails the extra assumption that the stem form of the singular was then also extended to the plural. In consequence, these words are not relevant for the treatment of geminates.

Several items seem to exhibit the simplification of *dd > *d*, like across the morpheme boundary seen above.

(5) The best evidence is W *credu*, MCorn. *crysy*, *cresy* (with secondary assibiliation), MBret. *cridiff*, Bret. *krediñ* 'to believe', cf. OIr. *creitem* 'belief' < PC *$kred$-*dī-mā* 'belief' < PIE *$\hat{k}red\ d^{h}eh_1$- 'to put one's heart (*uel sim.*)', also supported by Lat. *credere*, Vedic *śrad dhā*-, Avestan *zrazdā*- 'to believe' (see now Weiss, 2020). In this word, the geminate does not result from composition with a preverb, but the compositional joint between the constituents may arguably have remained transparent until late in prehistory.

(6) This is more difficult to argue for a number of other words in which *dd can be set up for etymological or comparative reasons. For W *rhwd*, OBret. *rod* 'rust; 'mud, filth', and the Cornish placename *Polroad*, *Polrode*, possibly from *$ruddo$- < *$rudzdo$- see section 6.1. above.

(7) Bret. *red* 'bog myrtle' is surely cognate with OIr. *rait* 'id.', in which case the preform must be *$ro/addi$- or perhaps *$ruddi$- from the same root *$h_1 reu̯d^{h}$ 'to become red' as the previous item (Stifter, 1998: 216–217).

(8) Another possible instance is the equation OIr. *gat* 'theft' and Bret. *gad* 'hare', which leads to the reconstruction of the common preform *$gaddo/ā$- (in Stifter, 2021, I linked the two semantically distant words through the popular belief that hares steal milk).

(9) The Latin loan *abbatem* 'abbot' is reflected as W, Bret. *abad*, OCorn. *abat*.

While in Irish it is the rule that geminates that resulted from the assimilation of two stops were subsequently simplified to a single stop, the foregoing discussion makes it inevitable to conclude that geminate clusters were treated differentially in British. *bb and *dd seem to have been reduced to single voiced stops, but *gg may have become the voiceless fricative *χ. Where, on the other hand, single *g results in Welsh, this may be due to analogy. However, Martin Kümmel and Anders Jørgensen (in peer-review) make the valid point that if we allow for an analogical treatment of clusters across morpheme boundaries, *zd, which does not occur across synchronically analysable boundaries, might show the regular development of *dd. This is to say that *zd could have passed through the stage *dd, only to undergo, together with other tautosyllabic instances of *dd, devoicing to *tt and subsequent spirantisation, thus being part of a general devoicing of voiced geminate clusters (thus Jørgensen, 2022: 145–146). Morphemically analysable clusters of two *dd, on the other hand, could have analogically preserved their voicing and subsequently have fed the later British Celtic rule *dd > *d. This scenario entails that Bret. *gad* 'hare' and *red* 'bog myrtle' must be borrowings (presumably from Irish, even though the examples seem semantically random). See also section 4.1. It is evident that more research needs to be done on the treatment of geminates, and it cannot be guaranteed that all relevant examples were taken into consideration in this study, since no exhaustive search was carried out.

6.3. *RR

This paragraph brings together a handful of diverse instances in individual languages where geminate resonants do – or do not – arise as a consequence of identical sounds coming into contact across the composition boundary; phenomena which do not fit precisely into any of the preceding sections.

(1) A special, language-internal case of identical sounds across the morpheme boundary is Gaul. *petorritum* 'four-wheeled wagon'. The geminate *rr* must have arisen from metathesis of earlier *$petru$-*rito*-, which consists of the composition form *$petru$- < *$k^{u}etru$- of the numeral '4' (itself metathesised from earlier *$k^{u}etur$-), and a nominal formation of the root *ret- 'to run'. Alternatively *$k^{u}etur$- could be an archaism preserved in this compound, or *petor*- is simplified from the younger Gaulish form *$petu̯or$- with influence from the cardinal *$k^{u}etu̯ores$ '4'.

(2) A word that is best reconstructed as a compound of the preverb *kom- and the root *men- 'to think' shows the very exceptional behaviour that, instead of expected *$kommen$- (a compound with a root noun?), the lexical stem is reflected as *$komen$- with a simple *m*, namely OIr. *cuman* 'reminiscent', W *cof*, Corn. *cof*, MBret. *couff*, NBret. *koun*, *kouñ* 'memory'. The behaviour of *$komen$- finds no parallel in other compounds of *kom- with a root beginning with *m*-. Instead, gemination is the rule, *e.g.*, OIr. *commailt* 'consumption' < *kom-*melti*-, *commus* 'competence' < *kom-*med-tu*-, or W *cymaint* 'as large as' < *kom-*mantī*-, *etc*. The difference is perhaps one of time-depth (VKG i 171; Jackson, 1953: 481). The simplification of *mm > *m* could even be an echo of the Proto-Indo-European constraint on geminates. Alternatively, the allomorph *ko- of *kom- could have been extrapolated from adjectives such as PC *ko-*u̯ar-i̯o*- < *kom-*u̯ar-i̯o*- 'proper' (OIr. *coäir*, MW *cyweir*) or *ko-*u̯īr-V*- < *kom-*u̯īr-V*- 'correct' (W *cywir*) where *m was regularly lost before *$u̯$. To make things even more complicated, OBret. *commin*, glossing *annalibus* 'annals', must surely reflect the etymon of OIr. *cuman etc*., too. However, it is written with double *mm*, which otherwise represents unlenited *m*, namely in *camm* 'oblique', *gueimmonou* 'seaweed', *lammam* 'I jump', *pimmont*, *pimmunt* '50'.

(3) A potential parallel is OIr. *neim* 'poison', which inflects as a neuter *n*-stem in the singular. The underlying formation would be expected to be **nem-men-* (root **nem-* 'to apportion' + suffix **-men-*), but since the Old Irish word descriptively continues **nem-en-*, it is attractive to operate with the same early simplification of **mm* here. A different possibility for both **kommen-* and **nemmen-* is to assume the loss of the medial consonants in cases where the second syllable came to stand in the zero grade, *i.e.*, **kommn-* > **komn-*, with subsequent generalisation of the simplified stem **komen-* (personal communication Michael Weiss). Such a strategy has been employed to explain OIr. *gein* 'birth' < **genen-* ← **ĝenh₁men-* (see 5.1. (1)).

At the end of the foregoing developments, the phonological system of Celtic had morphed into the one in Table 2.

There was a contrast between *s* and a strong sibilant (probably an affricate, represented by *ts* in Table 2), although that contrast does not mirror that of simple *vs*. geminate consonants elsewhere. It is probable that single voiced stops already had lenited allophones in intervocalic position, but this is not indicated in the table. Geminate *p:* has been tentatively included in this table as a marginal phoneme to allow for the theoretical possibility that **pp* had been preserved or acquired in contrast to simple Indo-European **p*. On the whole, geminated voiced stops were rare except across the morpheme boundary (an observation already made by Martinet, 1952: 198). They have been tentatively included as phonemes in Table 2, but Gianguido Manzelli and Elisa Roma (in peer-review) make the point that, in view of their rarity, **dd* etc. could be considered as consonant clusters rather than geminates at this stage of the language.

7. Etymological gemination: Proto-Goidelic **NT* > **DD*

Notwithstanding the many diverse developments presented so far, one of the most common sources for geminate voiced stops in Irish are Proto-Celtic clusters of nasal + voiceless stop that developed into the corresponding voiced geminates after the separation of Goidelic from the rest of the Celtic languages, *e.g.*, Proto-Goidelic **gg* < PC **nk*. At some point in the prehistory of Irish, the voiceless stops in such clusters assimilated in voice to the preceding elements, while the nasals assimilated in the mode of articulation to the following stops (GOI 126–127; McCone, 1996: 106–109; Schrijver, 1993: 35–39). The two met, as it were, in the middle and a geminate voiced stop resulted. While geminate voiced stops are a rarity in Proto-Celtic and in other ancient European languages (Martinet, 1952: 198), thanks to this change they are very common in Primitive Irish.

In ogam inscriptions of the classical period (5th–7th century A.D.), the resulting voiced stops are written with the letters for D and G, *e.g.*, DECCEDDA < **dekantos* or TOGITTACC < **tonketākko-* (the double spelling DD or of the other consonants has nothing to do with geminate sounds, but is an orthographic convention in ogam that is independent of the phonological nature of the consonant). In Old Irish orthography, they are usually expressed by ⟨c⟩ and ⟨t⟩ (and by ⟨p⟩ for /b/ arising in other contexts; B is used in such cases in ogam inscriptions, e.g., TEBICATOS < **t(o)-ebbīkatu-*). Preceding Proto-Celtic short **a* and **e* (the latter of which had been raised allophonically to mid-high **ɪ before the tautosyllabic nasal) were lengthened in the process and fell together as ⟨é⟩ = /eː/. They are retained as such in accented syllables, but are shortened along with other long vowels in unaccented syllables. As a consequence, words with initial *éc-* = /eːg/ and *ét-* = /eːd/ very commonly go back to compounds consisting of the negative prefix **an-* 'un-' + etyma in *c-* and *t-*. There are countless, uncontroversial examples for this chain of sound changes. To name just a few representative examples:

(1) PIE **h₁dn̥t-* > **dant-* > OIr. *dét*, W, Bret. *dant*, OCorn. *dans* 'tooth'.

(2) PC **sentu-* > OIr. *sét*, Gaul. *sentu-*, W *hynt*, OBret. *hint*, Bret. *hent*, Corn. *hens* 'path'.

The unstressed development of the vowel is exemplified by

(3) PIE **h₂r̥ĝn̥to-* 'shining one' > **arganto-* > OIr. *argat*, Gaul. *arganto-*, Cib. *argato-*, W *arian*, *ariant*, OCorn. *argans*, MCorn. *arhans*, *arans*, OBret. *argant*, Bret. *arc'hant* 'silver, money'.

(4) and by the Old Irish preverbs *ceta·* < **kanta* < **km̥th₂* 'together with' (cf. Gaul. *canti-*, OW *cant* 'with') and *ceta·* < **kentu-* 'first'

Table 2. the early Common Celtic sound system.

	stop	nasal	fricative	glide	liquid
bilabial	(pː) b (bː)	m mː	(φ)	w	
dental	t (tː) d (dː)	n nː			l lː r rː
alveolar			s (z) ts		
palatal				j	
velar	k (kː) g (gː)		(x ɣ)		
labiovelar	kʷ (kʷː?) gʷ (gʷː?)			w	

(both also written *cita·*, the latter also *cíata·*), and by the preposition/preverb *etar* < **enter* 'between'.

Preceding **o* and **u* were not lengthened, *e.g.*, **tonketo-* > ogam TOGITTACC, OIr. *tocad* 'fortune', **kon-toletu-* > OIr. *cotlud* 'act of sleeping', or **slunke-* > OIr. *slucaid/sluicid* 'to swallow' (Schumacher, 2004: 593). The evidence for **i* – in contrast to mid-high **ɪ*, the allophone of **e* before a tautosyllabic nasal – has been ambivalently assessed, and it warrants more detailed attention. The following paragraphs are therefore a digression from the overall topic of geminates.

McCone (1996: 107–108) cites *ro·ic* 'to reach' < **inket* ← **īnkti* < **h₂ēnḱti* and *fet* 'whistle', with regular lowering from **u̯into-* 'wind' < **u̯īnto-* < **h₂u̯eh₁n̥to-* 'blowing' as relevant evidence for the question of the treatment of **iNC*. Rix makes the tentative alternative, but phonologically ultimately equivalent suggestion of analysing *ro·ic* as a continuation of **īnk-* from a reduplicated present **h₂i-h₂n̥ḱ-* (LIV 282–283, fn. 7). This position implies that **ink* became intermediate **igg* with a short vowel.

This approach is gainsaid by Schrijver (1993: 39–42) and Schumacher (2004: 200–204) who explain *ro·ic* as **ɪnket* < **ænnket* < **annketi* ← **h₂n̥nkti*; *fet* can be explained semantically more easily as **u̯izdo-* from the root **(s)u̯eiz-* 'to whistle'. On the other hand, Schrijver and Schumacher derive *léicid* straightforwardly from **link^{u}i̯eti* ← **li-n-k^{u}-*, while McCone (1998: 474–475) has to argue for analogical influence from a long-vowel form of the subjunctive stem at a pre-Irish stage.

Other forms have received less attention in this dispute, even though they are equally relevant for determining the outcome of PC **inC*. McCone (1996: 107–108) mentions a couple of "problematic instances with unlengthened stressed vowel", namely *ecor* 'arrangement' (the verbal noun of *in·cuirethar*), *tecosc* 'instruction', *do·ecmaing* 'to befall', *do·ecmalla* 'to collect', and conjugated and therefore stressed forms of the preposition *etar* 'between', such as *etruinn* < **enter-snī* (*uel sim.*). Relevant forms are also discussed by Armstrong (1976: esp. 64–66). McCone wonders if "the following *o* (plus *r* or *m/ṽ*) played a role in the loss of the nasal", but he arrives at no other conclusion than that "the precise conditioning remains unclear". He refers to GOI 518–519 in this context, but does not actually quote or discuss Thurneysen's alternative solution, even though it merits a closer look.

Thurneysen proposes that "these examples can best be explained by assuming that in them the preposition had at one time the form *in-*". The Celtic preposition with the meaning 'in' had been inherited from Indo-European as **eni* (Dunkel, 2014: 224–225). It is attested as a plain preposition in the most archaic form in Celtiberian *eni*, and in a few nominal compounds in other Celtic languages, *e.g.*, Gaul. *Enignus*, OIr. *ingen* 'daughter', ogam INIGENA < **enigenā*, OIr. *inis*, W *ynys*, Bret. *enez* 'island' < **eni-sth₂-ih₂-*. However, in the Gaulish preposition *in* it appears with loss of final *-i* and with unexpected raising. This same form **in* underlies also the ordinary Old Irish preposition *i*^{N} and Welsh *yn*. That the Irish preposition continues the high vowel **i* follows from the fact that **en* with mere apocope of final **-i* would have resulted in ***a*^{N} in Old Irish, like the masculine infixed pronoun *-a*^{N} < **-en* did. Dunkel (2014: 223) suggests that the vowel of Gaul. *in* reflects the raised allophone **ɪ* of **e* before a nasal in tautosyllabic position, *i.e.*, before consonant.[3] In Irish, **ɪn* is usually kept distinct from original **in*, but in this case it has to be assumed that mid-high **ɪ* was further raised to **i*.

In Gaulish, this allomorph also intruded into compounds, *e.g.*, *Indercillus* < **en(i)-derk-* 'having an eye inside (?)' or *in-dutio-* 'being endowed with (?)' in the name *Indutiomarus*, perhaps from **en(i)-dʰoh₁t-* (Delamarre, 2003: 163). In Irish, **in* was semantically and morphologically conflated with Proto-Celtic **andi/ande* < **n̥dʰi*.[4] This is most readily seen in the preposition itself: while the plain preposition is *i*^{N} < **in*, the inflected forms such as 1sg. *indium*, 3pl. *intiu etc*. are built on the stem *ind-* < **ande*. In the 3sg. masc. *and*, yet another etymon **andom* < **n̥-dom* 'at home, inside', was drawn into the paradigm (Dunkel, 2014: 159, 230).

In verbal composition, the situation is even more complex. The allomorphs of the preverbs have been conflated to such a degree that a clear distinction is not always possible. This is not aided by the fact that the presentation and morphological analysis of verbal forms in eDIL and other handbooks is often imprecise or incorrect. The archaic allomorph **eni-* shines through OIr. *do·infet* 'to inspire' and its 3sg. present subjunctive *·tinib* < **to-eni-su̯izd-*, but otherwise evidence for it is hard to come by and hard to distinguish from the more common allomorphs **ande-* and **in-* (reconstructed preverbs are written with a final hyphen in order to distinguish them formally from prepositions and adverbs). The 2sg. imperative of *in·cosaig/in·coisig* 'to signify, indicate' is *inchoisc*. As is evident from the lenited *-ch-*, this cannot be from **in/en-kom-sech-* but must either go back to **enikom-sk-* or **ande-kom-sk-*. eDIL quotes no attestations with *-d-*, *e.g.*, **·indchoisc*. The lack of such forms in a verb that is well attested in early sources can be taken as indirect evidence that it involves **eni-*.

The most common allomorph in verbal compounds appears to be **ande-*. It seems that in the earliest period it occurred

[3] There are other seemingly sporadic examples of *i* where *e* is expected, but none is parallel to **in* in distribution. The 3rd singular present indicative of the copula OIr. *is*, W *ys* must come from **isti* < **esti* < **h₁esti*, but these seem to be independent developments in Goidelic and British Celtic. In Old Irish, the *i* is the result of raising of unaccented vowels before palatalised **s* (Griffith, 2016: 48–51), in Welsh it is an instance of *i*-affection of **e* (Hamp, 1974b: 33; Schrijver, 1995: 265–268). In Gaulish, PC **esti* may in fact be attested with unraised *e* as *esi* (L-98 1a9) and as *esti* in the defixio found 2022 in Orléans. On the other hand, the 1st singular of the copula appears as *imi* (L-120), *immi* (G-13) in Gaulish, but since the corresponding Old Irish form *am* continues **emmi* < **h₁esmi* with unchanged vocalism, this raising must be specific to Gaulish.

[4] *Pace* Dunkel (2014: 224) who derives it from **en-dʰe*. The comparison of OIr. *indel* 'preparation, machinery' with Welsh *annel* 'trap, deception' < **ande-lo-*, and Gaul. *ande-* leave no alternative than to reconstruct PC **andi/ande* < **n̥dʰi*.

as *ind·* in pretonic position and as *·in(d)-* in stressed position, as opposed to **in-* that appeared as *in·* and *·in-* respectively. Ultimately both merged in *in·* in pretonic position and cannot always be kept apart in other positions either. A precise description of all contexts is not intended here. Finally, and to complicate matters even further, it will become clear from the following that beside **eni-*, **ande-* and **in-*, there was also a fourth allomorph **en-* with a much more restricted domain. Leaving aside **eni-*, whose presence in a handful of Irish compounds was demonstrated above, it can be shown by minimal pairs of phonological microcontexts that a three-way contrast between **ande-*, **in-* and **en-* needs to be made. The distinction between **ande-* and **in-* is required to account, for instance, for the different behaviour of *in(d)·lá* 'to enter into, arrange, *etc*.' and its prototonic stem, represented by the 3sg. subj. pres. *· indell* < **ande-la-* (not ***·ell* < ***·in-la- uel sim.*), versus *in·loing* 'to join, bring together, put upon, *etc*.' and its prototonic form *·ellaing* < **in-long-* (not ***·in(d)laing*); or *in·samlathar*, *·intamlathar* 'to imitate' < **ande-saml-* (not ***essamlathar*) versus *in·snaid* 'to insert', verbal noun *essnaid* < **in-snad-* (not ***intnaid*).

The reason why **in-* instead of **en-* is set up here for the preforms, at least for deuterotonic verbal forms when the preverb stands before the accent, is that **en-* in unstressed position would have become ***an·*.[5] On the other hand, the presence of the allomorph **en-* (or potentially even **an-*!) is required in other contexts, for instance by verbs such as *con·éitet* 'to accompany' < **kom-en-teiĝ-*; or *in·túaisi* 'to listen to' whose prototonic stem must be from **en-touṣtī-*, *e.g.*, 3pl. *·éitset*; or by the augmented subjunctive stem of *in(d)·fét* 'to tell, relate', *e.g.*, 1sg. pres. subj. *·écius* < **en-kom-uei̯d-s-*. All of these verbal forms show the regular change of **en-* > *é-* before a Proto-Celtic voiceless stop described at the beginning of this chapter. In the case of *·ellaing*, no decision can be made between **in-* and **en-*, since the outcome would be the same. Ultimately, in this way a fourfold allomorphy of **eni-*, **en-*, **in-* and **ande-* can be demonstrated in verbal morphology, with a number of distributional restrictions on their occurrence, as illustrated in Table 3.

While all the phonological developments invoked so far are trivial in the diachrony of Old Irish, this is not the case for the words quoted by Thurneysen (GOI 518–519) and McCone (1996: 108–109), namely *ecor* 'arrangement', *tecosc* 'instruction', *do·ecmaing* 'to befall', *do·ecmalla* 'to collect' and *etruinn*. These cannot go back to **en-koro- etc*. since it was just seen that this would have regularly resulted in ***écor etc*. The logical conclusion, after the alternatives have been excluded, is that they must continue **in-koro-*,[6] **to-in-kom-skᵘo-*, **to-in-kom-ink-*, **to-in-kom-la-*, and **inter-snī* respectively, with the development of **in-k-* > **igg-* in a first step, and then with regular lowering of **igg-* before a non-high non-front vowel > **egg-*. The same solution applies to *itge*, *itche* 'request, petition' < **in-tech-i̯o-* < **in-tekᵘ-i̯o-*, only with the regular absence of lowering.[7]

What about the distribution of **in- vs. *en-*? Perhaps its obscured rationale is that **in-* was originally at home in pretonic position, that is, as a plain preposition and before the accented part of the verbal complex, while **en-* occupied the stressed part of the verbal complex, as reflected in *con·éitet*, *·éitset* or *·écius*, and was used in most verbal nouns. In a subsequent development, the unstressed variant intruded by analogy into the stressed portion of the verb, which explains cases such as *do·ecmaing* or *ecor*. The replacement of **en-* by **in-* in this environment is not a regular, but a sporadic, analogical process. Operating with a sporadic replacement is not completely arbitrary or random. That there could be a formal distinction between

[5] The pretonic form *in·* could also be accounted for on the basis of **en-* if influence of the – contentious – main-clause particle **et* is invoked, namely **enet· > in·*.

[6] This has been compared with Gaulish *incoro* in the inscription from Châteaubleau (L-93), but in view of the latter's uncertain interpretation this comparison has to remain speculative.

[7] The word seems to have /t/, not the /d/ expected from the proposed etymology, cf. ModIr. *itghe*, *itche*. Perhaps *itge* was influenced by the semantically related *attach* 'invoking, refuge' or devoicing occurred before *ch*.

Table 3. the four allomorphs of preverbal 'in' in Primitive Irish.

	reconstruction	preverb in pretonic position	preverb in tonic position
1.	**to-eni-sụi̯zd-*	-	*do·infet, ·tinib*
	**eni-kom-sech-*	*in·cosaig, in·coisig*	*inchoisc*
2.	**ande-la-*	*in(d)·lá*	*·indell*
	**ande-samlā-*	*in·samlathar*	*·intamlathar*
3.	**in/en-long-*	*in·loing (in-)*	*·ellaing (in/en-)*
	**in/en-snad-*	*in·snaid (in-)*	*VN essnaid (in-)*
	**in/en-touṣtī-*	*in·túaisi (in-)*	*·éitset (en-)*
4.	**kom-en-teiĝ-*	-	*con·éitig*
	**en-kom-uei̯d-s-*	-	*·écius*

pretonic and tonic allomorphs of preverbs is commonplace elsewhere in Old Irish, for instance in the alternation *do· vs ·to-* or *a^{H}/as· vs. e(s)-*. It is equally well known that the two allomorphs could influence each other. In the case of *a^{H}/as·/e(s)-*, the spread of the pretonic allomorph into tonic position is observable in such verbs as *as·beir*, *· epir*, where the imperative is found as *apair* already in Old Irish, or in pronominal forms such as 1sg. *asum* or 3sg. masc./neut. *as* instead of archaic *es-*. In the case of *do·/to-*, the tonic allomorph was frequently used in pretonic position, in archaic and in archaising orthography, for example archaising *to·beir* for *do·beir*.

The presence of two allomorphs side by side with each other in one paradigm is not isolated in the wider Celtic perspective either (see also Schumacher, 2022: 196–199). It is securely attested for the prehistory of the Old Irish preposition *do* 'to, for' and for the preposition/preverb *de* 'from', the distribution of neither of which follows predictable rules throughout their respective paradigms. In the case of *do*, the allomorph PC **dū* < PIE **doh$_1$* (Dunkel, 2014: 148–149) underlies the plain preposition and conjugated forms such as *dúinn* 'to us' < **dū-snī(s)* or *dúib* 'to ye' < **dū-su̯ī(s)*, as well as British Celtic **dī* and perhaps Gaulish *duci*, whereas **do* < PIE **do* underlies conjugated forms such as *duit* 'to you' << **do-tī* < **do-to̯i̯* or *dó* 'to him/it' < adverbial **do*. In the case of *de*, its allomorphs PC **dī* < pre-Celtic **dē* and **de* < **de* occur without apparent distributional rationale in several compound verbs, *e.g.*, *·díltai* 'to deny' < **dī-slondī- vs. dermat* 'forgetting' < **de-ro-mento-* (but see the critical discussion in Dunkel, 2014: 148–156).

A side remark: The preposition/preverb PC **enter* is always continued with a short vowel in OIr. *etar* (cf. Lash, 2017 for other Old Irish allomorphs of it). This is expected for the preposition in pretonic position, but the short vowel is not expected in inflected forms of the preposition and when the preverb occurs in nominal compounds. For instance, the 1pl. is attested as *etruinn*, not as expected ***étruinn*, or the word for 'boundary-ditch, fence' is *etarbae*, not expected ***étarbae*. It might look attractive to derive those forms from an innovatory allomorph ***inter*, in which the more basic local preposition **in* had analogically caused the replacement of inherited **en* by **in*. However, such a scenario is excluded. There is no way how the hypothetical preform ***inter* would have led to **ed'er* with lowered **e-* in the first syllable. It is therefore more economical to assume that the preform is indeed the traditional **enter* > **ēd'er*,[8] which *via* vowel shortening and depalatalisation in pretonic position resulted in **eder*. In a further step, this shortened pretonic variant replaced the original tonic variant **éter* everywhere. The complete replacement across the board of all tonic allomorphs by the pretonic ones in this scenario nicely illustrates the randomness of analogical change. In the case of several other prepositions, the language tolerated the coexistence of dual stems in unstressed and stressed prepositions: *co vs. cuc-*, *amal vs. saml*, *dar vs. tor-* or, indeed, as demonstrated above, *i vs. ind-/ and*.

[8] Celtiberian *entor* and *entara*, and the Gaulish placename *Entarabo* ostensibly retain the old vocalism of **entVr*; Late Gaulish *inter ambes* in Endlicher's Glossary could be due to influence from Latin. Welsh *ythr* is ambiguous since preconsonantal **en* was regularly raised to **ın* in British.

8. Etymological gemination: Insular Celtic mutations

When we broaden the perspective from the level of isolated lexemes to that of accentual units and constituent phrases, we can see that some of the sandhi-effects across words that ultimately became the initial mutations of the Insular Celtic languages, resemble the types of geminations that arose word-internally. These sandhi-related processes, namely the interaction of word-final sounds with word-initial sounds, thus extended the positions where geminate sounds were phonotactically permissible from word-internal to word-initial position. A comprehensive diachronic description of the emergence of mutation in Insular Celtic is outside the aims of this study. The present section only aims at highlighting the fundamental parallels of certain types of mutations with word-internal gemination. Only sandhi involving two interacting consonants is relevant for the question of gemination. Leniting contexts, where an initial consonant originally followed a word ending in a vowel, are therefore ignored here.

8.1. Irish

In the following, the examples will be drawn mainly from contexts that underlie mutations in Irish. Traditional grammars of Irish mention a mutation which they call 'gemination' (*e.g.*, GOI 150–153). This is a misnomer, both synchronically and diachronically. Diachronically, other sandhi contexts also led to geminate sounds in Primitive Irish in the appropriate contexts. Synchronically, the occasional geminate spellings of certain sounds are best regarded as markers of non-lenition (Greene, 1956). Instead, the main effect of this mutation, which finds no graphic expression in genuinely Old Irish sources, but which is directly observable in Middle and Modern Irish written sources and in Modern Irish pronunciation, is the prefixation of *h-* to a following word if it starts with a vowel. Therefore this mutation is called 'aspiration' here (cf. Stifter, 2009: 65).

The sandhi contexts that are relevant for the diachronic study of gemination are therefore nasalisation, aspiration, and non-mutation, *i.e.*, the appearance of the unmutated radical sound. It is essential to acknowledge that in addition to the morphosyntactically recognised mutations, the absence of an overt change, *i.e.*, 'non-mutation', also is a mutation in its own right. From the point of view of diachronic phonology, aspiration and non-mutation are for the most part just two sides of the same coin. In the vast majority of cases, both continue contexts where the mutated word was preceded by a word ending in *-s*.[9] At the time when final syllables had not been apocopated yet, the effects of non-mutation must have been identical to those of aspiration. The split between aspiration and non-mutation depends on the context: either intervocalic or adjacent to a consonant. Aspiration occurs when the final **-s* > **-h* combined with the initial vowel of the following word in Primitive Irish after the shift of the syllable boundary. When a consonant followed, the *-h* merged with that consonant. In most cases,

[9] Non-mutation furthermore occurs after words ending in *-r*, but this is marginal in comparison to the cases with final **-s*. Aspiration is also found after **-th* < **-t*, if the Old Irish prepositions *la* and *fri* continue **let* and **u̯rit*. It is a moot point if the conjunct ending of third-person-singular verbs, which likewise ended in *-t* in the prehistory of Irish, would have caused aspiration. Since no evidence for it is known, this possibility will not be further pursued here.

this must have led to a phonetically slightly longer pronunciation that, at least in the case of the resonants, but perhaps also of stops, meant phonetic similarity to or identity with word-internal geminates. However, the effect was different if the following word started with *$\underset{\frown}{u}$-. In that case the chain of events was *-*s* $\underset{\frown}{u}$- > *-*h* $\underset{\frown}{u}$- > *# *h*$\underset{\frown}{u}$- > OIr. -*Ø f*-, where the outcome *f* (phonologically equivalent to a geminated *$\underset{\frown}{u}$!) is identical to that of the internal group *-*s*$\underset{\frown}{u}$-, but is not phonetically similar to single *$\underset{\frown}{u}$, which rather gets lost intervocalically. On the level of surface phonetics this means that when *-*h* stood immediately before a consonant, it was absorbed by it and non-mutation ensued.

Synchronically, aspiration is only caused by proclitic mono- or disyllabic particles that end in a vowel. It is conceivable that aspiration was originally also caused by Old Irish inflectional forms that ended in a vowel, for example the accusative plural of masculine and feminine nouns, but there is no written evidence for this. This effect can only be conjectured. Non-mutation, in any case, occurs after all classes of words that ended in a consonant and in phrase-initial position.

Through the outlined processes, non-lenited initials and non-nasalised initial voiced stops came to be phonetically identifiable with internal geminates. The phonetic salience of geminates word-internally, and the rise of geminates word-initially, must have mutually reinforced each other. Sandhi-generated geminates will have considerably increased the overall token frequency of gemination in speech and strengthened its phonological status in the system. In contrast, gemination was rare in absolute word-final position.

In the case of the nasal mutation, several contexts need to be distinguished. A probably weakly articulated nasal sound in final position was attached and merged with initial consonants. It assimilated fully to initial resonants. That this sound must have been phonologically *-*n* can be gleaned from the different treatments of inherited initial and internal clusters with **m*, which do not turn into geminates, *e.g.*, **mrogi*- > OIr. *mruig* ‘land’, **mligeti* > *mligid* ‘to milk’, or **kom-rigo*- > OIr. *cuimrech* ‘binding, bond’. Furthermore, this -*n* is directly visible in the nasal mutation *n*- on initial vowels. It is therefore evident that, like in Gaulish and British Celtic, final PC *-*m* in the ending of accusative singulars, genitive plurals, and neuter nominative singulars, had become *-*n* in the prehistory of Irish. There are not many inherited word-internal clusters of **n* + **r* or **l* that illustrate their identical treatment to that across the word boundary. The best examples are furnished by compounds with the preverb **en*- ‘in’, where it hadn’t been replaced by other allomorphs. Practically speaking, the only good examples are those with **l*, which indeed show the expected assimilation to *-*ll*-, namely OIr. *ellam* and *ellach* (see section 3.6.). For *-*nr*-, the only potential example that I am aware of is OIr. *eirr* ‘chariot-fighter’. This has been suggested to continue **en-ret*- ‘he who runs into (the battle)’ (see NIL 577 n. 3), but the analysis **ers-sed*- ‘he who sits at the back’ is preferable because of its parallel to *arae* ‘chariot-driver’ < **are-sed*- ‘he who sits at the front’ (see 3.2. (4)). Nasalisation of voiceless initial stops has exactly the same outcome as in word-internal position, *i.e.*, a voiced single stop results synchronically. It can be surmised that the sound was a geminate at an intermediate stage (cf. chapter 7.). Nasalisation of voiced initial stops does not lead to geminates, but to prenasalised stops.

8.2. British and Gaulish

In contrast to the situation outlined for Irish, there is a broad consensus that gemination of initial sounds played no role in the emergence of British Celtic lenition (Harvey, 1984; Russell, 1985; Schrijver, 1999; Sims-Williams, 1990; Sims-Williams, 2008; Thomas, 1990); as for its role in British spirantisation, see section 14.2. below

Very occasionally, external sandhi phenomena that look similar in structure to Insular Celtic mutations can be found written in Gaulish inscriptions, *e.g.*, *reguccambion* < *$reg\bar{u}$-$k^{(u)}$’ *kambion* ‘and I straighten the crooked’ in the inscription from Chamalières (L-100). The final *c* of *reguc* is probably enclitic *-$k^{u}e$ for ‘and’ that merged with the initial *c*- of *cambion* ‘crooked’ (differently De Bernardo Stempel, 2010: 69).

9. Non-etymological gemination: onomastic gemination

9.1. Vocative morphophonology

While phonological developments represent the largest source for geminate sounds, other factors may also have contributed to their increase in the language. One such factor is the pragmatics of onomastic morphology. It is a general observation that, especially in the ancient Celtic data, gemination is especially frequent in anthroponomastics, but not among the ‘long’ dithematic compound names of the Indo-European type, which often have martial and heroic connotations, but rather among more colloquial short or hypocoristic names (Stüber *et al.*, 2009: 37–38). This morphophonological behaviour is not limited to Celtic, but is a feature of many ancient Indo-European naming systems (cf. Schmitt, 1995: 425, 618, 620; and Ellis Evans, 1967: 296–297, 376 with earlier literature). Masson (1986: 220) stresses the central pragmatic importance of the vocative as the context in which gemination of stem-final consonants could arise, in personal names but also in expressions that belong to colloquial registers.

The origin of this process of onomastic gemination may lie in ‘vocative reduction’ or ‘vocative truncation’, one method of the formal marking of vocatives (see Daniel & Spencer, 2009: 628–629 for this and other types of vocative marking). Descriptively speaking, vocatives can be shorter than the forms of names that fill the regular argument slots in a sentence (cf. the pertinent examples cited in Daniel & Spencer, 2009: 629; Janson, 2013: 224–231; Schmitt, 1995: 419–425, 618). Indo-European languages bear this out in many ways. The shortening can take the form of the reduction of the number of segments or of the reduction of moras, for example in the Proto-Indo-European vocative of *o*-stems such as *$\underset{\frown}{u}$*iHros* → *$\underset{\frown}{u}$*iHre* ‘man’; or the vocative of -eh_2-stems, caused by laryngeal loss in pausa, *e.g.*, qPIE *$g^{u}enah_2$ → *$g^{u}ena$ ‘woman’, *e.g.*, in OCS voc. *ženo*; for consonant stems, cf. Greek nom. *Σωκράτης* (*Sōkrátēs*) → voc. *Σώκρατες* (*Sṓkrates*), which shows both stress shift and mora reduction.

In other languages, whole syllables are dropped at the end. A contemporary illustration is the neo-vocative of Russian whereby male names in *-a* lose the final vowel, *e.g.*, nom. *Saša* → voc. *Saš*. This habit has recently been borrowed into Georgian (Amiridze, 2022: 1–2). At the same time, this pragmatic shortening of names in situations of address – no doubt connected with the emotional urgency typical of conversations – can be counteracted on other levels of the speech act. The reduction or loss of segments, especially at the very end, is sometimes balanced by increasing the moraic count further to the front of the word. This is illustrated by the Georgian neo-vocative, where the loss of the final vowel is accompanied by the lengthening of the root vowel, *e.g.*, *Šota* → voc. *Šoot*, or *Gvanca* → voc. *Gvaanc* (Amiridze, 2022: 2), so that the overall moraic count of the name remains the same. Vowel lengthening is a typologically common process of vocative marking (Daniel & Spencer, 2009: 629). I believe that consonant gemination is a comparable phenomenon that is equally linked to vocative truncation, with the difference that, instead of lengthening a vowel as in Georgian, it is accompanied by the lengthening of the final consonant. Gemination can thus be regarded as a process that compensates for phonological or morphological loss, perhaps as a way of making up towards the end of the utterance for the overall reduction in the moraic structure. Due to the absence of suitable textual genres in the surviving documentation of the ancient Celtic languages, the connection between gemination and vocatives cannot be demonstrated with actual examples, but the relatively high number of geminates in personal names is indirect evidence of this tendency.

9.2. Gemination in Ancient Celtic names

Once gemination had been established as a morphological feature of personal names in a highly specific context, namely in the address of persons in intimate or informal speech acts, it could then be transferred also to contexts outside of vocatival function (for the generalisation of vocatives to other contexts more broadly, see Stifter, 2013). Examples of shortened names with gemination in core syntactical functions are well attested in ancient Celtic, *e.g.*, Gaul. *Eppo*, reflecting a compound name with **epo-* 'horse' as first member, *Blattia* ← **blātu-* 'flower', or *Sammus* ← **samo-* 'summer'. Commonly the 'root' (or rather onomastic basis) of such names is monosyllabic. Sometimes the etymology of Gaulish names is not entirely certain. For instance, the Gaulish name element *poppo-* has been interpreted as 'cook' < **kᵘokᵘo-* < PIE **pokᵘo-* (Delamarre, 2003: 252), and the names *Peccia*, *Peccio* have been compared with ogam Ir. QECIA, QECEA < PIE **kᵘekᵘ̯̊io-* 'courageous, strong' (Delamarre, 2003: 247), related with W *pybyr* 'eager, vigorous, brave' < **kᵘekᵘro-*. Onomastic gemination is the best explanation in such cases.

Conceptually related is gemination in kinship terms and in generic nouns for persons that are intimately known, *e.g.*, OIr. *macc* 'son, boy' < **makᵘkᵘo- vs.* ungeminated **makᵘo-* in Gaul. *mapon*, W, OCorn., Bret. *mab*; or W *geneth* 'girl' < **genettā- vs.* Gaul. *geneta*, OIr. *geined* < **genetā-* (more on this in section 11.1. on symbolic gemination). Gaulish has the personal name *Matta*, perhaps created from **mati/u-* 'good', but this word also appears in Raeto-Romance as a generic term for 'girl'. Maybe it had a similar generic function already in spoken Gaulish, from where it was retained as a substratum loan in Romance.

But also names with a larger phonetic body show gemination. A consonant, usually the final consonant before the ending, can be geminated without concomitant shortening of the name. Suffixes with voiceless velars are common in all Celtic languages. They continue earlier **-ko-* or **-k̄o-* added to vocalic stems, whence more complex suffixes such as **-iko-*, **-uko-*, **-īko-*, or the very productive **-āko-* arose in the Celtic languages. In Gallo-Greek inscriptions, *-Vkko-* is occasionally found instead of expected *-Vko-*, especially when the vowel is *-i-*: *Δονικκα*, *Ουαλικκο(νε)*, *Ουηβρουκκου*, and the fragmentary *Βιλλικκ*], *Ερικκ*[,]*ουικκος*,]*ουλικκι*.[10] Sometimes there is a geographical bias to such formations. For example, four of the five examples of the name *Maricca*, most likely derived from the adjective **māro-* 'big, great', are found in Noricum and Pannonia. Since there is no phonological factor discernible that could have triggered the gemination in such formations, it is natural to assume that it has to do with the onomastic character of the data.

The comparatively compact Gallo-Greek corpus in *Recueil informatisé des inscriptions gauloises* (RIIG) is an ideal field to study the distribution of geminate sounds in names. RIIG currently consists of 426 inscriptions. Of these, 234 are too short or too fragmentary for a meaningful analysis. The remaining 192 inscriptions contain 65 examples of geminate sounds. Under close inspection, their distribution turns out to be heavily skewed. The following results are based on a small sample and are therefore only preliminary.

As Table 4 shows, some geminates have a clear predilection for suffixes, while others occur only in the root or in the semantically salient portion of the names. *-ll-* (whose possible origin as a suffix was discussed in section 2.1.) and *-kk-* are by far the most common geminated stem-final consonants. Of the twelve examples of *-kk-*, nine (75%) occur in suffixes, only three in roots. A similar situation is observable among the seventeen examples of *-ll-*: eleven (65%) are in suffixes, six in roots. The ratio is reversed in the case of the other sounds. All eleven examples of *-nn-* in Greek-letter inscriptions are in roots. Suffixal *-nn-* is only found twice, both in inscriptions in the Latin script. This fact is significant and may have something to do with the rendering of Gaulish suprasegmental phonology in Latin (see section 10.2.). Of the nine examples of *-ss-*, six are in

[10] As a side remark, it can be speculated if the Welsh suffix *-ach*, which occurs in the names of rivers such as *Afallach* (based on *afall* 'apple-tree'), persons and peoples, may have a similar explanation, namely an onomastic gemination of an older suffix **-āko-* > **-akko-* with concomitant shortening of the vowel. Another possible example is *Berneich*, the Old Welsh name of the Old English kingdom of Bernicia, which presupposes a preform **bernakkī-* < **bern-ākī-* 'those at the gap (?)' (cf. Jackson, 1953: 701–705).

Table 4. geminates in the Gallo-Greek inscriptions.

	total	stem	suffix	unclear
ll	17	6	**11**	
kk	12	3	**9**	
nn	13	**11**	2	
ss	9	**6**	3	
tt	8	**7**	1	
rr	3	**3**		
pp	2			2
mm	1	**1**		
bb	—			
gg	—			
dd	—			

roots and three in suffixes. In the case of **-tt-* (lumped together with *-θθ-*), there are seven examples in roots and only one in a suffix. Of those seven, five belong to the onomastic stem **att-*, apparently from the etymon PIE **atto-* 'daddy' (see 1. (1)). Other geminates are too rare to draw clear conclusions. *-mm-* occurs only once in a root, and *-rr-* is found two or three times, apparently always in roots. The two examples of *-pp-* are too damaged to draw any conclusions. The most striking observation, however, is that voiced stops do not occur as geminates at all in this corpus. This links in with the overall impression that emerges from this study, namely that geminate voiced stops arose only very late through processes in the individual languages or that they are chiefly found in lexemes that are suspect of borrowing.

On a more general note, mainly based on anecdotal observations, geminates seem to be more common in Gaulish short names or in names formed with suffixes, as opposed to compound names. This may be a special characteristic of Gaulish and seems to be less common in Irish. The distribution of geminate sounds in ancient and medieval Celtic names deserves research on a much wider and more diverse material basis, but this goes beyond the aims of this article. There are some practical methodological limitations to identifying relevant examples in the written record. The Lepontic script of the Cisalpine Celtic languages and the Celtiberian script do not distinguish graphically between single and geminate consonants. Geminates in these languages can only be identified when vernacular names are also transmitted in the Latin or Greek alphabets, either in epigraphy or in manuscript texts. For instance, the Celtiberian name ⟨lubos⟩ in the Celtiberian script has a counterpart in the genitive *Lubbi* in the Latin alphabet, but, to complicate matters, *Lubus* is also found in epigraphy. Cisalpine Celtic *aśkoneti(o)* in vernacular writing has a parallel in *Adgonnetius*, *kasilus* corresponds to *Cassillus*, and *esanekoti* possibly contains the adjective **kotto-* 'old'. In other cases, such as *koimila* and *anteśilu*, the lack of a Latin parallel does not permit to decide if the *l* is single or geminate. Furthermore, the frequency of gemination in the attested *written* corpus of ancient and medieval Celtic languages may give a distorted picture. Dithematic, i.e. compound, names without onomastic gemination, which may be typical of the small aristocratic elite, may be overrepresented in our available sources. It is conceivable that the frequency of short or hypocoristic names with onomastic gemination was higher among the non-aristocratic population and accordingly in everyday spoken language.

While the examples in this section are mostly taken from Gaulish, the formal processes of vocative truncation counteracted by compensatory gemination are valid for all Celtic languages. These processes established gemination as a morphophonological feature in the naming system. Through them, names, especially high-frequency variants such as hypocoristics, came to contain a relatively high proportion of geminate stops, especially geminate voiceless stops, in contrast to the rest of the lexicon, where gemination was most prominent among resonants. This relative increase in geminate stops will in turn have led to an overall reinforcement of the status of geminates in the phonological system of Celtic languages.

10. Non-etymological gemination: accent-related or graphic gemination

De Bernardo Stempel (2010: 71–79) has observed that many words in ancient Celtic, especially in Gaulish, display etymologically unexplained gemination, in particular in post-tonic position – assuming that her hypotheses about the placement of the accent on the penultimate syllable in Gaulish are accurate. It is not possible here to repeat all of the sub-types she discusses, but a few select examples shall suffice.[11] The putative position of the Gaulish accent will be indicated by an acute accent: *simiuisonna* 'the name of a month' < **sēmi-u̯ēs-ón-ā* 'half spring (?)', *ogronno-* 'the name of a month' < **ou̯gró-no-* 'cold one (?)'; *uxello-* 'high' < **uχsélo-* < **upselo-*, cf. OIr. *úasal*, W *uchel*, Bret. *uhel* < **ou̯χselo-*, apparently with the same suffix, but with a different ablaut grade in the root.

De Bernardo Stempel thinks of genuine phonological gemination that is dependent on the stress. While this is possible, I do not want to exclude the alternative possibility that, at least in a subset of the material, such spellings could be a merely graphic convention in order to replicate Gaulish accentuation on the penultimate syllable within the framework of Latin suprasegmental phonotactic rules, which only permit stress on the penultimate syllable when it is positionally long.

[11] The assessment of De Bernardo Stempel's examples and of the various formal subgroups depends crucially on accepting her proposed etymologies, not all of which are uncontroversial. De Bernardo Stempel's 2023 monograph on Celtic accentuation appeared too late to be taken account in this article.

The reverse of De Bernardo Stempel's rule or tendency is the simplification or degemination of geminates in pretonic position (De Bernardo Stempel, 2010: 79). A good candidate for this is the group of Gaulish names around *Biracos* (stress presumably on ā) beside unsuffixed *Birros*, the latter being a likely cognate of OIr. *berr*, W *byr*, OCorn. *ber*, Bret. *berr* 'short'.

11. Non-etymological gemination: 'symbolic' gemination

11.1. 'Expressiveness'

Another factor – psychological and therefore outside the domain of regular phonetic change – played a role in the increase of geminate sounds. A number of instances of geminates are found in words that have, or are believed to have, an affective or emotive quality. This category of gemination is commonly referred to as 'expressive gemination' (*e.g.*, De Bernardo Stempel, 2010: 80; De Bernardo Stempel, 1999: 508–521; Lühr, 1985: 275–276; Stüber *et al.*, 2009: 260; Kuryłowicz, 1957: 132, 138, 142–144; these references are not meant to be exhaustive). In some kinship terms an emotional involvement, perhaps connected with vocatival gemination, is evident, especially where the creation of a geminate occurs only in a single language, *e.g.*, OIr. *macc* 'son, boy' < **makᵘkᵘo-*, which contrasts with the older, ungeminated **makᵘo-* that is continued in Gaul. *mapon*, W, OCorn., Bret. *mab*; or W *geneth* 'girl' < **genettā- vs.* Gaul. *geneta*, OIr. *geined* < **genetā-*. Perhaps **u̯raggā-* or **u̯rakkā-* 'old woman' (see 6.2. (3)) and Gaul. *ninno-*, OIr. *nen* 'handmaid', OW latinised personal name *Nennius*, Bret. *nen(n)* < **ninno/ā-* (?) 'servant (?)' can be included in this category as well. In these cases, there is no discernible phonological reason for the divergent development in just a single language. It is undeniable that a rule of arbitrary gemination, whether we call it expressive or hypocoristic, must have been synchronically operative at least in kinship terms. This was already evident in the words **atta* 'dad' and **mamma* 'mum', which can perhaps be reconstructed with geminates even for Proto-Indo-European, a language that otherwise actively avoided geminate sounds. Gemination in personal names may be of a similar nature.

Interjections are a class of words that combine expressiveness with a performative aspect, which can find formal expression in gemination.

(1) A relevant example in Old Irish is *nacc/naicc* 'no', probably from **nakᵘkᵘe*. This contrasts with the related **nakᵘe* < prohibitive **na* + the sentential connector **kᵘe*, which underlies W *na(c)*, Corn., Bret. *na(g)* 'not'.

(2) Once-attested OIr. *upp*, W *hwff*, Bret. *ouf* 'ooff' could in theory be derived from a common pre-form **upp*, beside **up* > W *wb* 'woe, alas', but they can also be viewed as spontaneous utterances.

However, expressiveness as an explanatory strategy is often extended beyond these narrowly circumscribed cases to words where emotional involvement of the speaker is not easily discernible. For example, we may ask ourselves why so many Celtic adjectives for basic concepts such as 'small', 'old', 'weak' display geminates. Most of them cannot be captured by an emotion-based concept of expressiveness. Instead I propose to use the more neutral, descriptive term 'symbolic gemination'. It takes its motivation from the concept of sound symbolism that refers to a perceived resemblance between the phonological structure of a word and its meaning. Sound symbolism can have a strong element of iconicity and onomatopoeia, but, like almost anything in language, a lot of the symbolism is arbitrary and can be triggered by culturally and grammatically specific conditions that, in the case of prehistoric languages, must remain obscure to us.

The etymologies discussed in this study so far comprise phonologically or morphophonologically motivated instances of gemination in inherited words (if we consider onomastic gemination as a morphophonological process). 'Symbolic' or 'iconic' gemination, on the other hand, is not triggered by phonetic cues, but is motivated by pragmatic considerations such as the semantics of words, or it has no discernible motivation at all – or at least none that is recoverable to us. It translates a special relationship, which speaker and hearer have to a concept, into linguistic markedness of words.

I am aware of the danger that any non-native classification of words as 'symbolic' may be arbitrary. I have therefore decided to restrict the category of symbolic gemination in Celtic to two types of words: onomatopoetic nouns (in fact a single example) and adjectives. Other scholars may apply different criteria.

11.2. Nouns

In one instance a reasonably good case can be made for sound symbolism, *i.e.*, onomatopoeia, namely OIr. *cloc* 'bell' < **kloggo-*. Medieval languages outside of Celtic presuppose the variant stem **klokkā-*, namely MLat. *clocca*, Fr. *cloch* and Germ. *Glocke*. W, OCorn. *cloch* (fem.), Bret. *kloc'h* (masc.) can either be compared directly with the Irish word, if the change **gg* > **χ* (see section 6.2.) is accepted, or they continue **klokko/ā-*, either as a genuine variant or as a borrowing from Middle Latin.

11.3. Adjectives

One semantic class of words in which phonologically obscure gemination is noticeably frequent, are comparable adjectives, very often those for basic concepts such as physical qualities. My speculative approach to explaining this descriptive fact is that the iconic component of gemination allowed speakers to give expression to uncertainty about the precise magnitude of those qualities, *i.e.*, it was an iconic way of 'hedging one's bets' (in the same way as speakers of German tend to qualify adjectives by *irgendwie* 'somehow', in order not to commit themselves too strongly to a specific value). In addition to their symbolic or iconic component, most of the adjectives in this category are also notable for being etymologically opaque. Many of them are not only found in Insular Celtic, but have parallels in Gaulish. Even though the contexts in which they appear in Gaulish, namely preponderantly in personal names, do not allow us to determine their semantic content beyond all doubt, the formal correspondence with words known from Insular Celtic renders their interpretation as adjectives fairly certain. For practical purposes, the material will be arranged

in groups that reflect their distribution. Some adjectives are found in several languages, and some are exclusive to a single sub-branch. In those cases that are found across several branches, borrowing between the branches can never be excluded.

11.3.1. Inherited formations

A subset of adjectives with geminates is of more or less good Indo-European inheritance. Even though gemination in these cases is phonologically regular and not due to sound symbolism, and they have already been discussed in previous sections, they are being repeated in this panorama for the sake of the overall argument. As will be argued at the end of this section, their presence in the language may have provided one morphosemantic trigger for the spread of gemination among adjectives.

(1) PC **ad-gostu-* 'at hand' > **aggossu-* > OIr. *acus*, W *agos*, OBret. *ocos*, Bret. *hogos*, OCorn. *ogos* 'near'.

(2) PC **dallo-* 'blind' > Gaul. *dallo-*, OIr., W, Corn., Bret. *dall*. As discussed in section 2.1. (5), it is attractive to explain this as PIE **dl̥̊no-*, but the lack of preserved **l̥̊* in the potential Gaulish examples is worrying. If the word is of Indo-European origin, its geminate can be explained through regular sound change from **ln*.

(3) qPC **dorro-* > Ir. *dorr* 'harsh, rough' and the abstract *doirr* 'anger'. Both words are only attested late. An Indo-European etymology **dor-so-* from the root **der-* 'to tear' has been suggested, but such an explanation is by no means unavoidable.

(4) PIE **gʰerso-* 'small, short' > **gerso-* > **gerro-* > OIr. *gerr* 'short (cropped); castrated'. If the ogam name GIRAGNI belongs here and equates with the Old Irish name *Gerrán*, the proposed Indo-European etymology becomes invalid because it cannot explain the **i*.

(5) PIE **kerso-* 'cropped' > **kerro-* > OIr. *cerr* 'crooked, maimed' could belong to the root **kers-* 'to cut'.

(6) PIE **korso-* 'cropped' > **korro-* > Gaul. *corro-*, *coro-* in names, W, OCorn. *cor*, Bret. *corr* 'dwarf'. This appears to be an ablaut doublet of the preceding item.

(7) qPIE **kurso-* > **kurro-* > OIr. *corr* 'peaked, pointed; projecting part', W *cwr* 'corner, pointy edge'. This could be related to Lat. *curuus* 'bent', Gr. *κυρτός* 'bulging, swelling, convex', but the underlying root **kur-* does not look Indo-European. De Vaan's (2008: 158) alternative suggestions for Lat. *curuus* eliminate a connection with the Celtic adjective. The Irish examples cannot always be distinguished from the preceding item.

(8) qPIE **kᵘelno-* or **kᵘelso-* > PC **kᵘello-* > W, Corn., Bret. *pell* 'far'.

(9) PC **mallo-* > Gaul. *mallo-*, OIr. *mall* 'slow', W *mall* 'evil, rotten, bad' can come from PIE **mḷ̊no-* or **mḷ̊so-* 'hesitant' (?).

(10) qPIE **polno-* 'full' > **φolno-* or **(h₂ˀ)olno-* > **ollo-* > Gaul. *ollon* 'big?; whole?', OIr. *oll* 'great, ample', W *oll*, *holl*, Corn., Bret. *oll* 'all'.

(11) PIE **(s)lh₂go-* 'slack' >> PC **laggo-* > OIr. *lac* 'weak'. If this word is of Indo-European origin, its geminate is unmotivated. Matasović (2009: 232) instead suggests PIE **lh₂g-ko-*, but I am unaware of evidence for the postulated development **gk* > **gg*. See also Zair (2012: 59). The gemination may perhaps have arisen secondarily and be due to systemic pressure from the class of primary adjectives discussed below.

(12) PIE **truCsmo-* > **trummo-* 'heavy, pressing' > OIr. *tromm*, W *trwm*, OBret. *trom* 'heavy', Corn. *trom* 'sudden, immediate'. A connection with PIE **treû̯d-* 'to thrust, press' or PC **truk-* 'unfortunate, sad' has been suggested (see Matasović, 2009: 391).

(13) qPIE **(s)tug/k-slo-* 'treated with a hammer or chisel?' > **tuslo-* > **tullo-* > OIr. *toll* 'pierced', W *twll*, Corn. *toll*, Bret. *toull* 'hole, perforation'.

One adjective has parallels in other Indo-European branches, but, because of its unusual structure, is believed to be of substratal origin:

(14) PC **menekki-* 'frequent' > OIr. *menicc*, W *mynych*, Corn. *menough* has correspondences in Germanic **mana/iga-* 'many' < **mono/igʰo-* and Old Church Slavonic *мъnogъ* 'much, many' < **munaga-* < **mn̥̊gʰo-* and perhaps even in Finno-Permic, *e.g.*, Finnish *moni* 'many', Votic *mõni*, Northern Sami *moanak*, Udmurt *мында* 'as much as'. This item is atypical for adjectives with gemination because of its disyllabic base.

11.3.2. Common Celtic formations

The following adjectives have no convincing extra-Celtic cognates:

(1) qPC **biggo-* > OIr. *bec* 'small, little'; OW *bichan*, W *bychan*, OCorn. *boghan*, Corn. *byghan*, Bret. *bihan* 'small, little' continue either **biggano-* or **bikkano-*, either with diminutive *-an* or with the suffix copied from PC **φlitano-* 'wide, broad' < PIE **pl̥̊th₂no-*.

(2) PC **birro-* > Gaul. *Birrus*, OIr. *berr*, W *byr*, OCorn. *ber*, Bret. *berr* 'short'. The Gaulish adjective seems to have been borrowed as a technical term into Lat. *birrus* 'a cloak to keep rain off', Gr. βίρρος 'cloak', perhaps originally a 'short cloak'.

(3) PC **brikko-* > Gaul. *Briccus*, OIr. *brecc*, W *brych*, Corn. *brygh* 'speckled', Bret. *brec'h* 'variola'. A connection with

PIE **perḱ-* 'speckled' has been suggested (LEIA B-82; Matasović, 2009: 78), but would be phonologically irregular. The early attestation of OIr. *brecc* with initial *br-* suggests that it is not etymologically connected with *mrecht* 'variegated' < **mrek-to-*.

(4) qPC **buggo-* 'soft' > OIr. *boc*, OBret. *buc*, *boc*, Bret. *bouk* (*boug*). However, Jørgensen (2022: 145) suggests that the actual reflex of **buggo-* in Breton is *bouc'h* 'blunt' and that *bouk* is a borrowing from Irish.

(5) qPC **burro-* 'swollen' > OIr. *borr* 'swollen', W *bwr*, OCorn. *bor* 'fat, strong, big'. LEIA (B-73) derives this from PIE **bʰorso-*, *i.e.*, from the same root as 3.2. (2) **barro-* < **bʰr̥so-* 'point', but this does not explain the vowel of the Welsh word.

(6) qPC **gollo-* > OIr. *goll* 'blind of one eye'. IEW 545 connects this *via* a 'secondary voicing' with "air. (acymr.?) *coll* 'luscum, einäugig'", a hapax in the Munich Sortilegia (*Thes.* ii 236.3). I suspect the latter to be a misunderstanding of *coll* 'loss, destruction'.

(7) qPC **glikki-* > OIr. *glicc* 'clever, ingenious, skilled'. Could there be a connection with *glacc* 'hand, grasp, clutch' < **glakkā-*? For the connection of 'cognition' and 'hands' cf. Germ. *begreifen* and Lat. *comprehendere* 'to touch with the hand = to understand'.

(8) qPC **gozdo-*, **gonto-* or **goddo-* > MIr. *got* 'stammering, lisping'.

(9) qPC **kokko-* 'red, scarlet' > Gaul. *cocco-*, W *coch*, Corn. *cough*. Probably an ancient *Wanderwort*, originally referring to kermes, the dye produced by the insect cochineal cf. Delamarre (2003: 120–21).

(10) qPC **kotto-* 'old' > Gaul. *Cottos*, OCorn. *coth*, Bret. *kozh*, and perhaps in the Old British ethnonym *Atecotti*.

(11) PC **metsamo-* (?) > **messamo-* 'worst' > Gaul. *messamobi*, OIr. *messam*. Originally a superlative with the suffix PC **-samo-*, it is morphologically isolated and no longer synchronically analysable in the attested languages and therefore warrants its inclusion in this list.

(12) qPC **kittu-* (?) > Early ModIr. *cittach* 'lefthanded, clumsy'. Unlike most of the other adjectives in this group this requires a stem in **-u-*, if it is an old formation at all. See Schrijver (2003: 4–5) for the problematic comparison of this item with W *chwith* 'left, awkward, unlucky' and related words.

(13) If the Gaulish names *Suallius/a*, *Sualius* are cognate with ogam Ir. SUVALLOS and OIr. *súaill*, *súail* 'small, trifle', they probably have to be analysed as compounds of **su-* 'good' + an element **u̯al(l)-i-* (in the case of OIr. *súail(l)* hardly identical with the root **u̯al-* 'ruler, to rule'). It is uncertain if the geminate *ll* is original or caused by individual secondary factors within each language.

(14) qPC **tretti-* > OIr. *treitt* 'quick, swift'; the more common form of the adjective, OIr. *traitt*, is perhaps influenced by the structurally and semantically similar OIr. *prapp*. W *trythyll*, *drythyll* 'lively, high-spirited; lascivious' and OIr. *treittell*, *dreittell* 'pet, favourite; warrior' could be derivatives from this. Alternatively *traitt* can be reconstructed as **tratti-*, but it would need to be separated etymologically from *treittell etc.* in that case.

11.3.3. Exclusively Irish examples

Several adjectives only found in Irish are ambiguous as to the precise reconstruction of the sound that synchronically in Old Irish looks like the descendant from a geminate:

(1) qPC **u̯ozdo-* > OIr. *fotae* 'long'. Together with Lat. *uastus* 'vaste, immense', IEW 1113 derives this from **u̯asdʰo-* or **u̯osdʰo-*. However, for De Vaan (2008: 655–656) Lat. *uastus* cannot be separated from *uāstus* 'desolate, waste', which comes from **h₁u̯(e)h₂sto-* 'empty' (= OIr. *fás* 'empty, void'). The latter is not a viable preform for OIr. *fotae*. In view of PC **oNT* > PrimIr. **oDD* (see section 7.1.), the reconstruction qPC **u̯onto-* is also possible, as is an original **u̯oddo-*.

(2) A mechanically reconstructed immediate preform of MIr. *láitir* 'strong, powerful' would be something like PrimIr. **lāddiri-*. Perhaps the conspicuous fact that the degrees of comparison, *e.g.*, comparative *laiteri*, and the derived abstracts *láitire* and *láitirecht* show no syncope, is an indication that the preform was more complex, *e.g.*, **lāde/idVri-*. No analysis suggests itself for any of those reconstructions.

(3) qPC **menno-* or **minno-* > OIr. *menn* 'clear, visible'. The homonym OIr. *menn* 'stammering, inarticulate, mute' could have an identical preform. Since the word is not attested in contemporary Old Irish sources, it could conceivably also have been **mend* < **mendo-* or **mindo-* originally.

(4) MIr. *prapp* 'sudden, swift' could be a language-internal neologism that makes iconic use both of the phonologically highly marked sound *p* and of gemination, or rather its reflex, *i.e.*, a voiceless unlenited intervocalic stop. Alternatively, it may reflect a symbolically reduplicated Latin loan, namely *rap-rapidus* (personal communication, Paulus van Sluis).

(5) qPC **trikki-* > OIr. *tricc* 'swift, active, sudden, urgent'.

Where do the adjectives without Indo-European pedigree come from? What is notable about this collection of adjectives, which strongly relies on Irish, is that, apart from being formally united by gemination, they form several semantically coherent subgroups. They relate to basic physical concepts; to physical defects (which could involve taboo formations); and they are symbolic of swiftness and agitation. Gemination was phonologically original in the subset that continues Indo-European formations. From there, it can have spread as a

semantic marker within the semantic field. External influence need not be invoked. It is evident that words such as **biggo-* 'small', **buggo-* 'soft', **laggo-* 'weak' constitute a 'phono-semantic' class whose structure was self-replicating. That such a process was productive, is evident from words such as *prapp* 'rapid, quick', which, if it isn't a substratal loan, can have been created sound-symbolically at a late date.

Another reason for ascribing these words to *Urschöpfungen* within Celtic, and not to loans from unknown languages, has to do with the typological tendency of adjectives to be less amenable to borrowing than nouns (cf. Matras, 2007; Matras & Adamou, in press). The fact that such a number of conceptually central adjectives display gemination points rather to sound-symbolism as the driving force. Once gemination had been established as a phono-semantic marker for basic adjectives, there must have been enough internal momentum to create new material through this morphophonological process, without taking recourse to borrowings from a substrate.

12. Unclear gemination

A small number of nouns and verbs of certain or likely Indo-European origin possess geminates that cannot be explained straightforwardly from their putative preforms by phonological developments. Neither do these words manifest an affective character. For some of them, I will make tentative morphological proposals, usually involving some sort of derivational process that can account for the gemination. However, these attempts are merely speculative. In some words, the gemination has to remain unaccounted for for the time being.

12.1. nouns

(1) PIE **h₁eh₂no-* 'ring' > **ānо-* → **ānniio-* > OIr. *áinne*. This item is related to Lat. *ānus*, Arm. *anowr* 'ring'. Maybe the derivation involves the addition of the suffix *-no-*, *i.e.*, **ān-no-*, to explain the geminate, or it is due to the syncope of a preform such as **ān-in-i̯o-*. The lack of Osthoff shortening and the palatalisation speak in favour of the latter option.

(2) PIE **pl̥keh₂-* 'flat stone' > **φlikā-* → **likkā-* > OIr. *lecc*, W *llech*, Bret. *lec'h*, Corn. *leghen*. The Gaulish placename *Arelica* appears to have a single **k*; various Gaulish names starting with *licc-* could contain the etymon with a geminate. Perhaps the geminate is the result of a 'precocious' syncope between identical sounds in an adjectival formation of the type **φlikо-kā-* 'stony', which became substantivised and ousted the original simple noun **φlikā-* 'stone', or the suffix was added directly to the root, *i.e.*, **φlik-kā-*.

(3) The relationship among OIr. *ícc* 'healing' < **īkkā-* and W, OCorn. *iach*, Bret. *yac'h* 'healthy' < **i̯akko-* or **i̯ekko-* is notoriously difficult, as is its extra-Celtic relationship, if there is any, with Myc. *a₂-ke-te-re*, *ja-ke-te-re*, Gr. ἄκος 'cure, medicine'. Matasović (2009: 171) and Zair (2012: 68) tentatively and fully conscious of the formal problems think of **i̯h₂ko-* as a starting point, but neither addresses the question of the geminate. Schrijver (1995: 103–104) dismisses the link with Greek and instead suggests, also tentatively, a formation **i̯et-ko/ā-* from the root **i̯et-* 'to position onself'. Although it cannot account for the vocalism of the Irish word, it explains the geminate. Finally, I want to join in the tentative chorus myself. A possible source for the geminate could be a Proto-Celtic adjectival formation **i̯ek-ko-* from the root **i̯ek-* 'to speak'. The semantic connection would be the premodern notion of healing through words of power. But I need to stress that aside from explaining the geminate, this etymology creates other issues with the initial vowel.

(4) qPC **kruttā-*, **krutto-* 'bulging object, belly (?)' > OIr. *crott* 'harp', W *croth* 'womb, uterus, belly', *crwth* 'fiddle, violin; hump, hunch; anything of bulging shape', Corn. *crothak* 'big-bellied', Bret. *kourzh* 'vulva'. MLat. *crotta* 'harp' looks like a loan from Irish. Lith. *krūtìs* 'breast' and Latv. *krùts* 'hill, breast' look related, but have a long vowel. A common preform could be set up as **kruHto-* with laryngeal loss in Celtic. Perhaps it contains the PIE root **kreu̯H-* 'to pile up, cover'.

(5) qPC **makko-* > OIr. *macc*, W *mach*, OBret. *meic* 'bond, enforcing surety'. Perhaps **magʰ-ko-* from the PIE root **magʰ-* 'to be able, powerful'?

(6) PIE **pr̥keh₂-* 'furrow' > PC **φrikā-* > Gall.-Lat. *rica* (whence Fr. *raie* 'ray'), Bret. *regenn*. OIr. *etarche*, *eitrige* continues a compound with *etar-* 'between'. However, W *rhych* 'furrow' and Bret. *rec'h* 'grief' (perhaps metonymically from 'worry-lines on the forehead'?) presuppose PC **φrikko-* with a seemingly unmotivated geminate. The difference in stem-class could indicate that this is in fact an adjectival derivative **φrik-ko-*. Alternatively, those forms go back to a preform with a sigmatic suffix, *i.e.* **φrik-so-*. Formal influence from **φrikkā-* < **pr̥d-keh₂-* 'fart' is unlikely. Stefan Höfler (personal communication) suggests a connection with Lat. *rixa* 'noisy quarrel, brawl' < PIE **h₁riḱ-s-eh₂-* instead.

(7) qPC **truddo-* from the root PIE **treu̯d-* 'to push' is set up in IEW 1095 as the preform of OIr. *troit* 'fight, battle', but neither word-formation, stem-class nor vocalism are satisfactorily explained by this. If an Indo-European connection is abandoned, the word can also be set up as **troddV-* or **trontV-*.

(8) PIE **u̯ei̯h₁tV-* 'winding (plant/thing)' > **u̯ei̯ttV-*. OIr. 1 *féith* 'kidney, fibre, sinew, vein' and 2 *féith* 'some kind of twining plant' have been derived from this; if and how 3 *féith* 'a swamp, marsh' belongs here, is unclear (Irslinger, 2000: 202–204; Matasović, 2009: 419; Zair, 2012: 230). W *gŵyth*, Corn. *gwythyen* 'vein, sinew', OBret. *goed*, Bret. *gwazh* 'valley, stream' are evidently related, but continue **u̯ei̯ttV-* with an unmotivated geminate. Bauer (2015: 66–67) tentatively suggests a borrowing from Irish into British.

12.2. verbs

Only a select number of Old Irish verbs will be included here.

(1) qPC **u̯ekkā-* > OIr. *feccaid* 'to bend, stoop; turn towards'. IEW 1135 derives this from the root **u̯ek-* 'to bend' "mit expressivem *-kk-*", but that root is now reconstructed as **u̯enk-* with a nasal (LIV 683). Unclear.

(2) On the evidence of ModIr. *dímhigean* 'contempt', OIr. *do·meicethar* 'to despise, condemn' has a medial /g/. This

precludes the reconstruction *?*mik$^{(\prime)}$né/n-h$_2$-*, tentatively proposed in LIV 429. Operating with the formation **mi-n-k(h$_2$)-e/o-*, where the position of the infixed *n* would have to be secondary, results in PrimIr. **migge/o-*. This explains the medial *g* of OIr., but leaves the stressed vowel *e* unaccounted for. W *edmygaf* 'to admire, honour' and Bret. *dismeg* 'opprobrious' < **-mik-* seem to continue a formation without a nasal infix, unless all go back to a preform **migg-*.

(3) Matasović (2009: 344) derives MIr. *scrípaid* 'to scratch' from PC **skrībbā-* from the root **skri(H)b^{h}-* 'to scratch' (LIV 562: **sk$^{(\prime)}$reiḇ^{h}-*). This is unnecessary, the Middle Irish verb may rather be a learned reborrowing of Lat. *scrībere* with unlenited *b*.

13. Gemination in potential loanwords

13.1. Possible loanwords

After having scrutinised the possible sources and the types of geminate consonants in the lexicon of the Celtic languages, the question to what extent geminate sounds can be relied on as indicators for the substratal origins of words can finally be broached. This applies of course only to words for which no etymological explanation within the Indo-European paradigm can been found, and to words for which no case – how ever flimsical – for a sound-symbolic neologism can be made. Some words with geminates are evidently foreign, a case in point being **katto-* 'cat', which is a trans-European *Wanderwort*, perhaps ultimately originating in a Bronze-Age Afro-Asiatic language such as Nubian. At the same time, this word is atypical for the substratum loans that we are interested in here since it refers to a foreign concept brought in from outside, not to one that was encountered *in situ* by incomers.

After having dealt in the preceding sections with words for which some sort of language-internal explanation can be found, we are left with a list of items that have no good etymology and no obvious source. This makes them suspect of being loans. The collection below expands on Matasović's (2009: 441–443) list of suspected non-Indo-European loanwords in Celtic, enlarged by additional examples that have been found in non-systematic searches. Some of the words have manifest parallels in languages outside of Celtic, but formal irregularities prohibit their reconstruction for the Indo-European protolanguage or even for a common Western Indo-European subnode:

(1) qPC **ātti-* > OIr. *áitt* 'place'. A variety of Indo-European etymologies have been proposed for this word, all requiring the 'Kluge's Law' treatment **-tn-* > **-tt-* (Pedersen, VKG i 161: **pōthni-*; Klingenschmitt in Lühr, 1985: 303: **ō-i̯et-nā-*; Bammesberger, 1998: **pōtni-*). Since the validity of Kluge's Law for Celtic is not accepted in this article, the word is only mechanically projected back to a possible Proto-Celtic pre-form, without further analysis.

(2) qPC **bakko-* 'angle, bend, hook, crook' > OIr. *bacc*, W *bach*, Corn. *bagh*, Bret. *bac'h*. This has cognates in Lat. *baculum* 'stick', Germanic **pag-* 'peg', which presuppose ungeminated sounds.

(3) qPC **bekko-* > Gaul. *beccus* > Fr. *bec* 'spout, muzzle'. Instead of continuing a variant **beko-* without gemination, Bret. *beg* 'spout, mouth' is best understood as a borrowing from Gallo-Romance (Piette, 1973: 80–81).

(4) qPC **braddo-* or **brazdo-* > OIr. *brat* 'spoil, plunder, robbery'.

(5) qPC **braddo-* or **brazdo-* → OIr. *bratán* 'salmon'. This etymon is perhaps identical with the previous, with a special semantic development from 'a good catch' > 'salmon'.

(6) qPC **bratto-* > OIr. *bratt* 'mantle, cloak', W *brethyn* 'cloth'; the vocalism of Bret. *brozh* 'skirt' is irregular.

(7) PC **brokko-* > Gaul. *Broccus* > Fr. *broche*, ogam BROCI, OIr. *brocc*, W, OCorn. *broch*, Bret. *broc'h* 'badger'. The early Irish forms beginning with *br-* exclude a connection of this word with PC **mrogi-* 'land'.

(8) PC **bukko-* > Gaul. *Buccus*, OIr. *bocc*, W *bwch*, OCorn. *boch*, Bret. *bouc'h* 'he-goat'. Perhaps a loan from Germanic **bukka-*.

(9) qPC **bunno-* 'owl, bittern' > OIr. *bonnán*, W *bwn* 'bittern', Bret. *bonn* 'crane'.

(10) qPC **donno-* > Gaul. *Donnos* '?', OIr. *donn* 'chief, noble'. This is an uncertain item. The meaning of the Gaulish word is unknown and the Irish word could be the substantivised adjective *donn* 'brown' (see 3.1. (4)).

(11) qPC **ētt^{s}(V)lūm(b)-* 'bat' > OIr. *íatlu*, W *ystlum*. See Van Sluis (forthc.) for the unusual reconstruction.

(12) qPC **gaddā-* 'theft' > OIr. *gat* 'theft', Bret. *gad* 'hare'. See Stifter (2021) for the semantic relationship between the two words.

(13) qPC **garrV-* 'calf of the leg, shank' > OIr. *gairr*, MW, MBret. *garr*, W, Corn. Bret. *gar*. Matasović (2009: 152) suggests a connection with PIE **ĝ^{h}esr-* 'hand', but this is semantically and formally not convincing since **sr* doesn't give **rr* in Celtic (see section 3.4.). Alternatively he considers the possibility of a connection with PIE **ĝ^{h}erso-* 'small, short', namely a zero-grade formation **ĝ^{h}r̥so-* > **garso-* > **garro-* 'short one', with the semantic evolution to 'limb > shank'.

(14) qPC **glakkā-* > OIr. *glacc* 'hand, grasp, clutch'.

(15) PC **gobann-* 'smith' > Gaul. *gobann-*, OIr. *gobae*, gen. *gobann*, W, OCorn. *gof*, MW pl. *gofein*, MBret. *goff*, NBret. *gov*. Beside the stem *gobann-*, which is attested in personal and divine names, Gaulish also has the stem *gobet-*. See Blažek (2008) for a formally not satisfying Indo-European derivation of this word.

(16) PC **gobbo-* 'beak, gob' > OIr. *gop*; Gaul. **gobbo-* is presupposed by Fr. *gober* 'to slurp', *gobelet* 'goblet'.

(17) The nominal stem *illo-* or *illio-* is frequent in Gaulish onomastics, but neither its meaning nor its analysis are known. It can be formally compared with OIr. *ell* 'flush, sudden emotion', or *ell* 'opportunity, advantage'.

(18) qPC **kappilo-* 'horse' > OIr. *cappall*, dat.pl. *caiplib*, W *ceffyl*, OB *chefel*; Bret. *kefeleg* 'woodcock' is a derivative. Related to Lat. *caballus* 'pack-horse', but not a borrowing.

(19) qPC **karreko/ā-* > W *carreg*, Corn. *karrek*, Bret. *karreg* 'rock'; borrowed into OIr. *carrac*.

(20) PC **kau̯anno-* 'owl' > Gaul. *cauannos*, Fr. *chouan*, *chouette*, W *cuan*, OBret. *cou(h)ann*, Bret. *kaouenn* (cf. Jørgensen forthc.).

(21) qPC **kīmmukko-* 'lobster' > W *cimwch*, *ceimwch*, *gimwch*. It is most economical to assume that ModIr. *gliomach*, *giomach* was borrowed from W *gimwch* at a comparatively late date.

(22) qPC **kladdo-* 'stone (?)' > ModIr. *cladach* '(rocky) shore', *cladar* 'heap of stones', *cladán* 'fence-like pile of stones', all with [d] (see Stifter, 2023a: 179 for a very speculative connection with **kaleto-* 'hard' *via* a loan from British). Since the word is restricted to Goidelic, a reconstruction with **zd* would also be possible.

(23) qPC **knukko-* > OIr. *cnocc* 'protuberance, hill', W *cnwch* 'protuberance', OBret. *cnoch*, MBret. *cnech*, NBret. *krec'h* 'hill'.

(24) OIr. *corr* 'heron' < **korrā-* or **kurrā-* < **korsā-* or **kursā-* could simply be a substantivisation of the adjective 'hooked, pointed, humped' (see 11.3.2. (6); thus Schrijver, 1997: 297) or it could continue a more complex preform qPC **korχsā-*. The reconstruction with the complex cluster **rχs* is prompted by the comparison with W *crychydd*, Bret. *kerc'heiz etc*. (Schrijver *ibid*.), which would require the presence of a velar sound in the coda of the root. Proto-Germanic **hraigran-/*higra(n)-* has a related shape, but ultimately the starting point for this word is probably onomatopoetic.

(25) qPC **krettā-* > OIr. *crett* 'framework', W *creth* 'nature, appearance, form'. The Old Irish oblique cases *creitte*, *creitt* point to the Proto-Celtic preform **krettā-*, but **krittā-* is not entirely excluded.

(26) qPC **krokkenno-* 'hide, skin' > OIr. *croiccenn*, MBret. *crochenn*, Bret. *kroc'hen*. The palatalised *cc* of Old Irish must have spread from syncopated forms such as **croiccnea*. Gallo-Lat. *crocina* 'mastruca, a garment made of skin' has a different suffix; the precise formal relationship of W *croen*, OCorn. *croin*, Corn. *croghen* 'skin' is unclear.

(27) PC **kᵘesno-* (?) > **kᵘenno-* 'head' > Gaul. *penno-*, OIr. *cenn*, W, OCorn. *pen*, Bret. *penn*.

(28) qPC **luttā-* or **lottā-* > OIr. *lott* 'ruin, damage, injury'.

(29) qPC **mekkVno-* > OIr. *meccon* 'an edible root'. The word is occasionally written with double *nn*, *e.g.*, dat. *meccunn*. If this is old, the suffix could be similar to that of the 'avian' suffix **-anno-* discussed by Jørgensen (forthc.). For speculations about this etymologically obscure word see Lühr (1985: 293–295).

(30) qPC **met(t)o-* (?) > OIr. *meth* 'decay', W *meth* 'failure, error', Corn. *meth*, Bret. *mezh* 'shame'. The irregular correspondence between the Irish and the British words suggests that one branch must have borrowed the word from the other. In the absence of an extra-Celtic cognate, establishing even the Celtic preform is impossible. If it was **meto-*, Irish *meth* is regular and the word was borrowed into British (thus Bauer, 2015: 71–73). If it was **metto-* with a geminate, it is the other way round. Because of its isolation, the word may be a substratal loanword, in which case the probability of **metto-* increases.

(31) qPC **nūsso-* > OIr. *nús* 'beastings, milk of a newly calfed cow'. Perhaps borrowed into W *nus*, Bret. *luzenn*, unless the British word continue **n/lou̯sso-* with a different ablaut grade.

(32) W *pwll*, OCorn. *pol*, Bret. *poull* 'hole, pit; pool, pond'. OIr. *poll* 'hole' is evidently a borrowing; usually its source is believed to be Welsh. A mechanical back-projection results in qPC **kᵘullo-*, which is phonologically implausible, since labiovelars were delabialised before **u* in Proto-Celtic. A borrowing from OE *pól* 'pool' is phonologically and genetically unlikely: it cannot explain the vocalism of the Welsh and Breton words and a loan from Old English would not be expected to be present in all three British languages. Alternatively, it could be a substratal borrowing from the so-called '*p*-language', a pre-Celtic language postulated to have been present in Ireland as late as the first half of the 1st millennium A.D. by Schrijver (2000; 2005; see also Stifter, forthc.).

(33) qPIE **rou̯kko-* > ModIr. *ruacán*, *rócán* 'cockle', W *rhuch* 'bran, husks; garment'. The final *-t* of MIr. *rucht* 'cloak' is unexplained (Stifter, 2023a: 186). A connection with the root **$h_1$reu̯d^h-* 'red' is unlikely; Germanic **hrukka-* 'cloak' cannot be related because of the different initial.

(34) qPC **skaddo-* 'herring' > OIr. *scatán*, W *sgadan*, *ysgadan*. This substratal word has also been borrowed into Germanic languages, *e.g.* OE *sceadd*, ON *skaddr* (Schrijver, 2005: 141–143; Stifter, 2023a: 184).

(35) Pre-Irish **sladdā-* 'plundering, robbery' > OIr. *slat*. If this noun is related with the verb *slaidid* 'to strike, slay; plunder' from the PC root **slad-*, then the geminate **dd* of **sladdā-* may be due to influence from the synonyms OIr. *brat* 'plunder' and *gat* 'theft'.

(36) qPC **slattā-* 'rod, staff' > OIr. *slatt*, W *llath*, Corn. *lath*, Bret. *lazh*. This has cognates in Proto-Germanic **laþ(þ)a/ōn-*, **latta(n)-* 'lathe'.

(37) Pre-Irish **sliggi̯o-* or **sleggi̯o-* > OIr. *slice* '(oyster)shell'. Alternatively, **slinki̯o-* is also conceivable (Stifter, 2023a: 186).

(38) qPC **tanno-* 'green oak' > Gaul. *tanno-*, Bret. *glastannen*. OIr. *tinne*, which is glossed as 'holly, elder' in some glossaries, should be kept separate. The word is perhaps only an artificial construct to press the ogam letter into an arboreal scheme. Formally and semantically, Germanic Old Saxon *danna* 'pine' and OHG *tanna* 'fir wood' have to be kept apart.

(39) qPC **u̯annellā-* 'swallow' > OIr. *fannall*, W *gwennol*, Bret. *gwennili*, OCorn. *guennol*. Gall.-Lat. *uannellus* 'lapwing' may represent the same etymon (see Stifter, 2010: 149–150). On the other hand Matasović, for whom the Celtic words for 'swallow' are inherited, reconstructs **u̯ennālā-* < **u̯esn-*, derived from the Indo-European word for 'spring'. However, this reconstruction cannot explain the vocalism of the Irish word, nor of the Gaulish, if it belongs here. The different reconstructions at least agree in setting up a geminate **-nn-* for the preform.

13.2. Assessment

We can now proceed to a provisional assessment of this collection. Most of these words are isolated and lack parallels outside Celtic. Where parallels exist, they are often in Germanic (but this may be partly a consequence of how the data was compiled). It is therefore conceivable that these words may have been borrowed from unknown languages in the prehistory of Celtic, in Western Europe or on the Western Archipelago. Unlike the adjectives in section 11.3., the borrowing of nouns is typologically a trivial phenomenon. The case for borrowing can be made more rigorously if it can be shown that the candidate words share specific phonological or morphological features apart from gemination. A number of relatively unspecific observations can indeed be made (I will occasionally also include words from previous sections in this discussion):

Many of these words are monothematic and monosyllabic (if we disregard the Proto-Celtic ending), where gemination is found at the end of the stem or root syllable. From a Proto-Celtic perspective, the geminate sound occurs in the onset of the second syllable that also contains the ending. Typical examples are **bratto-* or **gobbo-*. Examples with more than one syllable are **ētt^{s}(V)lūm(b)-*, **gobann-*, **kappilo-*, **karrekā-*, **kīmmukko-*, **krokkeno-*, **mekkVno-*, **u̯annellā-*; **kau̯anno-* takes a special place since **-anno-* can be isolated as a suffix (see Jørgensen forthc.).

Another feature is the preponderance of short vowels, or rather the absence of long vowels, in the root syllables (cf. Van Sluis forthc.). Including disyllables, the most common vowels are *a* (11) and *o* (6 or 7); *e* (5) and *u* (4 or 5) lag somewhat behind. Intriguingly, *i* is absent from the collection, unless the obscure Gaul. *ill(i)o-* is a substratal borrowing. This is in contrast to the adjectives in section 11.3., where *i* is not uncommon. Exceptions with long vowels are **ātti-*, **ētt^{s}(V)lūm(b)-*, **kīmmukko-*, **nūsso-*, items that are all only attested in the Insular Celtic languages. Maybe they reflect borrowings from a local prehistoric language in these islands. Perhaps items such as **rou̯kko-* and **u̯ei̯ttā-*, that have been mechanically reconstructed with a Proto-Celtic diphthong, can be referred to the group with long vowels as well. Among the words with geminate stops, the majority have voiceless sounds (the comparative dearth of geminate voiced stops was already highlighted by Martinet (1952: 198)). Exceptions are **gaddo/ā-*, **gobbo-*, **skaddo-* and **sle/iggi̯o-*. **Braddo-*, **kladdo-* and **tru/oddo-* could be further instances, but since they are restricted to Irish, their **dd* could in theory continue earlier **zd*. **Kloggo-* is a special case since it may be onomatopoetic. Among geminated resonants, *nn* and *rr* dominate.

The words have been preponderantly adopted as *o-* or *ā*-stems. Exceptions are **ātti-*, **ētt^{s}(V)lūm(b)-* and the consonantal stem **gobann-*. The complex morphology of the words for 'smith' (*e.g.*, Gaul. **gobet-*) raises the question of a language-internal origin.

From the words included in the group with possibly Indo-European origin (section 12.1.), **krutto/ā-*, **makko-* and **tru/oddo-* would also fit formally into the group of potential loanwords. On the other hand, **φlikkā-* and **φrikko-* stand apart with their radical *i*.

One must never lose sight of the fact that the borrowings need not come from a single source. The formal difference between some of the words can be an indication that more than one donor language is involved. The phonetic shape of the extant words owes as much to the target phonology of the Celtic languages to which it had to be adapted, as it does to the donor languages. Depending on when the borrowing occurred, the donor language need not have possessed geminates at all: if, at the time of borrowing, the receiving language already had allophonic lenition of inherited intervocalic stops, any foreign unlenited intervocalic stop could have been perceived as a geminate.

14. The insular outlook

At the end of the developments described in the foregoing chapters, the phonological system of Celtic had morphed into the one in Table 5.

Automatic, positional allophones are in brackets. This system reflects, in a somewhat idealised fashion, the state of affairs in the ancient Celtic languages and in the prehistoric Insular Celtic languages before the emergence of phonological lenition. For the descriptively *p*-Celtic languages, **p* has to be substituted for **k^{w}*. The relative frequencies of geminate stops, especially of the voiced stops, may have differed between the individual Celtic languages. Geminated labiovelar stops like in **mak^{w}k^{w}o-*, the precursor of OIr. *macc* 'son', probably were not very numerous. The notation **k^{w}k^{w}* = [k^{w}ː] is meant as purely arithmetical. For the purposes of this paper, it makes no practical difference if a geminate **k^{w}* is phonetically [kk^{w}] rather than [k^{w}k^{w}], contradictory positions advocated by two reviewers of this paper.

Table 5. the pre-Insular-Celtic sound system.

	stop	nasal	fricative	glide	liquid
bilabial	(p:) b b:	m m:	(β)		
dental	t t: d d:	n n:	(ð)		l l: r r:
alveolar			s (z) ts		
palatal				j	
velar	k k: g g:		(x ɣ)		
labiovelar	k^w k^w: g^w g^w:			w	

Table 6. the Primitive Irish sound system.

	stop	nasal	fricative	glide	liquid
bilabial	(p:) b(:)	m(:)	β β̃		
dental	t(:) d(:)	n: n	θ ð		l: l r: r
alveolar			s		
palatal				j	
velar	k(:) g(:)		x ɣ		
labiovelar	k^w(:) g^w(:)		x^w $ɣ^w$	w	
glottal			h		

14.1. Irish

After it had taken such a long time to build up a phonological system that had an opposition of length for almost every consonant, it is surprising to see how quickly this system disintegrated even before the full attestation of the medieval Celtic languages set in. Chiefly responsible for the quick abandonment of gemination, at least on the phonological level, was the rise of the opposition between unlenited *vs*. lenited sounds in the Insular Celtic languages. The emergence of lenition as an allophonic feature of single stops in intervocalic position may well at first have been a side effect of the strong opposition between single *vs*. geminate consonants, caused by a maximum phonetic polarisation of the manner of articulation (see also Martinet (1952: 198–203, 215–216) for a more phonetically oriented account of the loss of gemination and for the systemic implications of gemination as a prominent phonetic class). However, after the contrast lenition : non-lenition had emerged as the central phonological opposition in Insular Celtic, the gemination of stops lost its phonological significance and turned into a phonetic feature that was concomitant with non-lenition, indicated by the bracketed length marks in Table 6 below. The old opposition geminated : ungeminated was re-phonologised as the new phonemic opposition unlenited : lenited only among resonants. The resulting phonemic system of Primitive Irish is shown in Table 6.

For the ancient, medieval and early modern Gaelic languages, information about the phonetic status of the sounds can only be inferred indirectly from the written sources. The fact that single and double resonants were distinguished fairly consistently in manuscripts up to the modern period is an indication that the graphic distinction corresponded to a real-life phonetic distinction. At no point in the history of Irish can a comparable consistency be observed for the writing of unlenited stops. This is not the place to discuss the intricacies of Old, Middle and Early Modern Irish spelling conventions. Suffice it to say that, for example in Old Irish, the representation of word-internal unlenited stops by single or double letters appears to have been chiefly a matter of personal discretion of each scribe. This is exacerbated in later copies of Old Irish texts by the interference of a host of alternative spelling conventions for Modern Irish.

For the modern languages, the earliest recordings of traditional speakers allow some insight into the phonetic realisation of the sounds. For Irish, these are the recordings on wax cylinders by Rudolf Trebitsch in 1907 (kept at the Phonogrammarchiv of the Austrian Academy of Sciences; Lechleitner & Remmer, 2004) and the recordings by Wilhelm Doegen on shellacs, carried out between 1928–1931 (Royal Irish Academy; Conroy *et al*., 2009). Recent phonetic studies of the recordings of Donegal Irish in the Doegen collection shed some light on the phonetic length of unlenited sounds (Wheatley & Iosad, 2021). It appears that in the early 20th century, a contrast in length was still perceptible for some sounds, best preserved among resonants where the researchers noted perceptible length distinctions. In contrast, there was no significant length distinction in stops. These findings agree with the orthographic observations made above.

14.2. British

The details of the prehistoric developments of geminated sounds are rather different in the British Celtic languages and can only be sketched here. A striking difference to Irish is the sound **p* as the regular outcome of the labiovelar **k^w* in the phonological system of British Celtic. It seems that the voiced labiovelar **g^w* was reinterpreted as an allophone of **w* and that the two occurred in complementary distribution. When the merger of **w* and **g^w* happened is unclear.

The transformation of the Proto-British system (Table 7) into that of the individual neo-Celtic languages Welsh, Cornish and Breton has been the subject of a long controversy (see the section on previous research in the introduction). From a structural point of view, the operation of intervocalic lenition on single stops removed single voiceless stops from intervocalic position for a short period. In response to this, a chain-shift was set in action in which gemination lost its phonological significance and became merely allophonic with non-lenition, or, in other words, geminate voiceless stops refilled the now empty phonotactical slot of simple voiceless stops. Geminated voiced **b:* and **d:* fell together with the new simple **b* and **d*, which themselves were the results of the earlier lenition of simple **p* and **t*.[12] The case of **g:* is less clear. It may also have

Table 7. the Proto-British sound system.

	stop	nasal	fricative	glide	liquid
bilabial	p: p b: b	m:	β β̃		
dental	t: t d: d	n: n	ð		l: l r: r
alveolar			s		
palatal				j	
velar	k: k g: g		(x) ɣ		
labiovelar	g^w g^w:?			w	
glottal			h		

developed into **g*, thereby falling together with the product of lenited simple **k*, but in some cases it seems to have merged with **k:* in yielding **x* (see the discussion in section 6.2.). Unlike in Irish, there was probably never a stage where geminates occurred word-initially as allophones of unlenited sounds. See Van Sluis (2019: 30–35) for an account that differs in the details with regard to initial consonants.

The operation of syncope in Late Proto-British brought previously intervocalic stops and fricatives into contact with each other and thus created new complex consonantal clusters, which were prone to assimilation. The outcomes of this process called 'provection' in British historical phonology are voiceless stops, *e.g.*, the Welsh river name *Calettwr* < **kaled †đuβr* 'hard-water' < **kaleto-dubro-*, Bret. *klopenn* 'skull' < **klog †benn* 'stone-head' < **kloko-k^uenno-* (Harvey, 1984: 98–99). By virtue of two segments coalescing into one, one could expect that the assimilation product was a long sound at first, at least phonetically and at least for a short period of time. Van Sluis (2019: 69–74) argues that word-external provection (*i.e.* in sandhi) also produced long consonants up to Early Middle Welsh.

However, the ultimate result word-internally is a phonologically single, voiceless stop (Harvey, 1984: 99). Provection triggered yet another chain-shift by which existing single voiceless stops, *i.e.*, the old voiceless geminates, underwent a new round of phonetic weakening when intervocalic inside a word, after resonants, or within accentual units. This weakening, which has the appearance of a 'secondary lenition' (Greene, 1956: 289), but which is traditionally called 'spirantisation', resulted in the introduction of voiceless fricatives into the system, *i.e.*, old **pp* > **p* > **f etc*. In order to grasp the bewildering aspects of the diachronic treatment of geminates in British Celtic it is quintessential to rigorously differentiate between phonology and phonetics. Different classes of sounds (voiced stops, voiceless stops, resonants) were degeminated first phonologically, then phonetically, at different times (Isaac, 2004: 70).[13] Like in the Gaelic languages, geminate resonants remained longest as a phonological class and their simplification only occurred during the historical period.

The system immediately before the emergence of the individual British languages was therefore the one in Table 8. (*:*) indicates that length may have remained for some time as a phonetically concomitant feature, even though it played no role in phonology (cf. also Schrijver, 2011: 30–33).

15. Conclusions

Geminate consonants were a distinct and prominent phonological class in the history of the Celtic languages, namely in the attested ancient Celtic languages of the Continent, as well as in the reconstructable stages immediately before the attestation of the medieval Celtic languages in the insular world. Geminates also played an important role on the interface between phonology and morphophonology. The phonology of the historically attested Insular Celtic languages, on the other hand, has evolved into different systems where gemination as a phonetic feature is concomitant to other oppositions at best.

While sound change in inherited lexical items can explain the regular emergence of the core of this Common Celtic phonological class, it cannot account for all the types of gemination and for the entirety of the examples. By and large, a dichotomy is discernible in the Celtic lexicon between, on the one hand, words with geminate resonants and a 'strong' sibilant, and, on the other hand, words with geminate stops. Very crudely put, geminates of the first group tend to have good Indo-European

Table 8. the late Proto-British sound system.

	stop	nasal	fricative	glide	liquid
bilabial	p b(:)	m(:)	f β β̃		
dental	t d(:)	n: n	θ ð		l: l r: r
alveolar			s		
palatal				j	
velar	k g(:)	ŋ	x ɣ		
labiovelar	g^w(:)		h^w?	w	
glottal			h		

[12] In some positions, geminated voiced stops may have been retained as allophonic variants of simple unlenited stops. Especially in the case of Breton, phonetic, but non-phonological gemination was maintained and was still observed by researchers in Breton of the 20th-century (Falc'hun, 1951; Greene, 1956: 285; Jackson, 1953: 545–548; Jackson, 1967: 317). Van Sluis (2019: 73) argues that such a situation obtained also in Welsh up to Early Middle Welsh.

[13] Martin Kümmel (in peer-review) points out that "the relative length of voiceless obstruents may be argued to have remained relevant until the split of vowels in (Modern) Welsh into a long/tense variant in originally open syllables (preceding simple voiced consonants and some fricative-stop clusters) and a short/lax variant in originally closed syllables (preceding clusters, original geminates, voiceless consonants and *m*, *ng*)".

etymologies, whereas those of the second group are less amenable to explanations along traditional Indo-European principles, unless they occur across a morphemic boundary especially in compound formations. In the latter, assimilations across the morpheme boundary contributed to the rise of geminates, but to different degrees in the different Celtic languages. The evidence especially of ancient Celtic personal names indicates that the various sounds showed a skewed distribution between the root and the suffixal part of words. This observation illustrates that gemination had acquired a morphological function in addition to its purely phonological status. Voiceless geminate stops were fairly common in Proto-Celtic, but voiced geminate stops are virtually absent from Proto-Celtic. Words in which the latter did not arise across morpheme boundaries, and in Goidelic as the product of the change *NT > *DD, are mostly either strongly suspect of being sound-symbolic neologisms or borrowings from unknown sources.

Although there is no clear-cut demarcation line between the groups, it is useful for analytic purposes to divide the material into one with phonologically explicable gemination and one with gemination that has no known diachronic source. In many of the steps outlined in this survey, gemination can be seen as a phonetic strategy to reduce the number of phonotactically admissable consonant clusters while retaining the moraic structure of the words. Etymological gemination is thus a compensatory reaction on the prosodic level to the loss of phonological segments. In the case of non-etymological gemination, two possible sources are conceivable: either they are loans from prehistoric substratal languages, or they are *Urschöpfungen*, *i.e.*, neologisms created within Proto-Celtic or in the individual languages. For typological reasons it has been suggested above that nouns rather belong to the first category, whereas adjectives belong to the second.

There was no single pathway to the rise of gemination in Celtic. Over the long period of at least one and a half thousand years, a variety of factors conspired and various processes reinforced each other to slowly increase the number of types, and the number of tokens, of geminate consonants in the Celtic languages. Some of the earliest steps in this direction constitute 'uneven' phonological change, when, for instance, **sm* and **ln* became **mm* and **ll* already in Proto-Celtic, whereas **sn* and, perhaps, **sl* did not immediately follow suit. The material assembled in this investigation also demonstrates that the emergence and treatment of geminated voiced stops went rather diverse paths in the Celtic languages.

In addition to sound changes that provide purely phonological factors for the rise of gemination, there are also psychological drivers, namely borrowing, symbolism and pragmatics. In previous scholarship, expressiveness has been claimed to be an important factor, but this has not proved to be a useful concept on any grand scale. While the emotional attitude of speakers clearly plays a role in kinship terms, that segment of the lexicon accounts only for a tiny portion of instances of gemination. After the systemic establishment of gemination as a phonological category, it may rather have been the concept of sound symbolism that played a central role in the further spread of gemination, in particular in the creation of adjectives describing physical reality. It is very rare that we find, as it were, 'spontaneous' gemination in words of Indo-European inheritance, *i.e.*, gemination that can neither be explained by phonological nor by psychological factors.

Because of the imbalance in documentation between Continental and Insular Celtic languages, it is difficult to assess if the frequency of gemination, especially of stops, is higher in the latter than in the former, that is to say, if their frequency increased over time. However, it is noteworthy that geminated voiced stops are surprisingly rare, albeit not absent, in Gaulish. They may have been reinforced as a numerically substantial class only in Insular Celtic. Throughout the observable history and the reconstructable prehistory of Celtic, it is always the resonants that have been particularly prone to the creation of new instances of gemination (or rather 'fortification' during the younger stages of the Insular Celtic languages, when one can no longer speak of true geminates). However, when 'full saturation', as it were, of gemination was achieved in the Insular Celtic languages, it was quickly abandoned as a phonologically distinct class.

Of the data analysed in this study (excluding personal names), over 50 lack an obvious inherited, *i.e.*, Indo-European explanation. We may allow for the possibility that a subgroup of words with geminates was created language-internally through the operation of analogy and sound symbolism. This view is taken here especially for adjectives. However, the overall number of unetymologised words is too large to ascribe them entirely to those factors. It is suggestive that at least a subgroup of them, in particular nouns, were borrowed from prehistoric local precursor languages in Western Europe and in the Western Archipelago. Geminates are only one sign for substratal loans, and ideally they should go hand in hand with other markers for non-inherited origin, such as other unusual phonological and morphological features. The question to what extent the number of unetymologised words with geminates in Celtic is different (larger, equal, or lower) than in other languages of Europe, especially the Germanic (cf. Kuiper's (1995: 68–72) source 'A2' for loans into that language branch) and Italic languages and Greek, goes beyond the present article and needs to be investigated in a broader, comparative context.

In any case, from the foregoing it emerges that Celtic is an Indo-European branch that seems to contain a substantial body of words with geminate sounds borrowed from unknown substratal languages. The question imposes itself if this observation may be linked with the fact that in the Celtic languages intervocalic single stops underwent weakening through lenition, that is to say that *unlenited* single intervocalic stops of the donor language were *perceived* by speakers of Celtic as their only available unlenited intervocalic stops, namely geminated. The evidence for such a scenario is not clear-cut: Single intervocalic voiced stops were probably already

allophonically lenited in Proto-Celtic. This would render foreign unlenited voiced stops particularly prone to be replaced by geminates in the borrowing process, but in fact voiced geminate stops constitute only a small minority among the words with unknown provenance in the Celtic languages. It is therefore best to assume that the donor language or languages did indeed have geminates.

Ethics and consent

Ethical approval and consent were not required.

Data availability

All data underlying the results are available as part of the article and no additional source data are required.

References

Amiridze N: **Final-Vowel Truncation in the Forms of Address in Modern Spoken Georgian**. In: Aybüke Özgün, Yulia Zinova (eds.), *Language, Logic, and Computation. 13th International Tbilisi Symposium, TbiLLC 2019*. Batumi, Georgia, September 16–20, 2019. Revised Selected Papers. [= Lecture Notes in Computer Science]. Cham: Springer, 2022; **13206**: 3–25.
Publisher Full Text

Armstrong J: **Phonological irregularity in compound verb forms in the Würzburg Glosses.** *Ériu.* 1976; **27**: 46–72.
Reference Source

Bammesberger A: **The etymology of Irish *áit.* 'place'.** *Ériu.* 1998; **49**: 41–43.

Bauer B: **Intra-Celtic loanwords**. PhD-thesis, University of Vienna, 2015.

Blažek V: **Celtic 'Smith' and his Colleagues**. In: Alexander Lubotsky, Jos Schaeken, Jeroen Wiedenhof (eds.), *Evidence and Counter-Evidence. Essays in Honour of Frederik Kortlandt. Volume I: Balto-Slavic and Indo-European Linguistics*. [= Studies in Slavic and General Linguistics]. Leiden: Brill, 2008; **32**: 67–85.
Reference Source

Conroy J, *et al.*: **The Doegen Records Web Project. Irish Dialect Sound Recordings 1928-31**. Dublin: Royal Irish Academy, 2009.

Daniel M, Spencer A: **The vocative - an outlier case**. In: Andrej Malchukov and Andrew Spencer (eds.), *The Oxford Handbook of Case*. Oxford University Press, 2009; 626–634.

Darling M: **The Subjunctive in Celtic: Studies in Historical Phonology and Morphology**. PhD-thesis, University of Cambridge, 2019.
Reference Source

De Bernardo Stempel P: **Nominale Wortbildung des älteren Irischen. Stammbildung und Derivation.** [= Buchreihe der Zeitschrift für celtische Philologie 15]. Tübingen: Niemeyer, 1999.
Reference Source

De Bernardo Stempel P: **Die Geminaten des Festlandkeltischen**. In: Karin Stüber, Thomas Zehnder, Dieter Bachmann (eds.), *Akten des 5. Deutschsprachigen Keltologensymposiums. Zürich, 7.-10. September 2009*. Wien: Praesens, 2010; 65–88.

De Bernardo SP: **The Accents of Celtic. New Light on the Older and Oldest Stages.** Heidelberg: Universitätsverlag Winter, 2023.
Reference Source

Delamarre X: **Dictionnaire de la langue gauloise: Une approche linguistique du vieux-celtique continental**. Paris: Éditions Errance, 2003.
Reference Source

Delamarre X: **Noms de personnes celtiques dans l'épigraphie classique. Nomina Celtic Aniqua Selecta Inscriptionum**. Paris: Éditions Errance, 2007.
Reference Source

De Vaan M: **Etymological Dictionary of Latin and the other Italic Languages.** [= Leiden Indo-European Etymological Dictionary Series 7]. Leiden – Boston: Brill, 2008.
Reference Source

Dunkel GE: **Lexikon der indogermanischen Partikeln und Pronominalstämme. Band 1. Einleitung, Terminologie, Lautgesetze, Adverbialendungen, Nominalsuffixe, Anhänge und Indices. Band 2: Lexikon**. Heidelberg: Universitätsverlag Winter, 2014.

Egurtzegi A, Ariztimuño B: **Remarks on the Etymology of the Basque Word for 'Swallow' and Its Potential Relation to Celtic.** *Ériu.* 2013; **63**: 79–90.
Reference Source

Ellis Evans D: **Gaulish Personal Names: A Study of some Continental Celtic Formations**. Oxford: Clarendon Press, 1967.
Reference Source

Eska JF: **Tau Gallicum.** *Studia Celtica.* 1998; **32**: 115–127.

Eska JF: **Laryngeal realism and the prehistory of Celtic.** *Trans Philol Soc.* 2018; **116**(3): 320–331.
Publisher Full Text

Falc'hun F: **Le système consonantique du breton arec une étude comparative de phonétique expérimentale**. Rennes: Imprimeries Réunis. 1951.
Reference Source

Favereau F: **Geriadur Bras. Brezhoneg-galleg & français-breton**. Montroulez: Skol Vreizh, 2016.
Reference Source

Feuth E: **Germination: An Old Irish Mutation Rule?** *Ériu.* 1983; **34**: 143–156.
Reference Source

Fleuriot L: **Dictionnaire des gloses en vieux breton**. Paris: Klincksieck, 1964.
Reference Source

George K (ed.): **An Gerlyver Meur. Cornish-English, English-Cornish Dictionary.** 3rd Edition. Llandysul: Gomer Press, 2020.

GPC Online: **Aberystwyth: University of Wales Centre for Advanced Welsh and Celtic Studies**. 2014.
Reference Source

Graves E: **The Old Cornish Vocabulary.** Columbia University PhD dissertation, 1962.

Greene D: **Gemination.** *Celtica.* 1956; **3**: 284–289.

Greene C: **The spirant mutation in Brythonic.** *Celtica.* 1966; **7**: 116–119.

Griffith A: **On the Old Irish third palatalisation and the 3sg. present of the copula.** *Ériu.* 2016; **66**: 39–62.
Publisher Full Text

Griffith A, Stifter D, Toner G: **Early Irish Lexicography - A Research Survey.** *Kratylos.* 2018; **63**: 1–28.
Publisher Full Text

Hamp E: **Varia III 3.** *acus ~ ocuis. Ériu.* 1981; **32**: 159–161.
Reference Source

Hamp EP: ***sm.* in Gaulish.** *Romance Philology.* 1974; **28**: 17–20.

Hamp EP: **tyst.** *The Bulletin of the Board of Celtic Studies.* 1974b; **26/1**: 33.

Hamp EP: **Welsh *dichlyn'.*** *The Bulletin of the Board of Celtic Studies.* 1974c; **26**: 138.

Hamp EP: ***diffoddi* and *differaf/diffryt.*** *The Bulletin of the Board of Celtic Studies.* 1974d; **26**: 309–309.

Hamp EP: **On some sources of Primitive Breton I.** *Études Celtiques.* 1974e; **14**(1): 195–199.
Reference Source

Harvey A: **Aspects of lenition and spirantization.** *Cambr Medieval Celt Stud.* 1984; **8**: 87–100.
Reference Source

Hayden D, Stifter D: **The lexicography and etymology of OIr. *eclas*.** *North American Journal of Celtic Studies.* 2022; **6**(2): 236–250.
Publisher Full Text

Hill E: **Silbische Liquiden vor Nasalen im Inselkeltischen und das Problem der Nasalpräsentien vom Typ air. sernaid,. kymr.** *sernaid kymr -sarnu Keltische Forschungen.* 2012; **5**: 157–184.
Reference Source

Hill E: **Untersuchungen zum inneren Sandhi des Indogermanischen. Der Zusammenstoß von Dentalplosiven im Indoiranischen, Germanischen, Italischen und Keltischen.** [= Münchner Forschungen zur historischen Sprachwissenschaft 1]. Bremen: Hempen, 2003.
Reference Source

Irslinger BS: **Abstrakta mit Dentalsuffixen im Altirischen**. Heidelberg: Universitätsverlag Winter, 2000.
Reference Source

Isaac G: **The Chronology of the Development of Brittonic Stops and the Spirant Mutation.** *J Celt Linguist.* 2004; **8**(1): 49–85.
Reference Source

Isaac G: **Brittonic Voiceless Spirants Again.** *J Celt Linguist.* 2008; **12**(1): 17–37.
Reference Source

Jackson K: **Language and History in Britain. A chronological survey of the Brittonic Languages 1st to 12th c. A.D**. Edinburgh University Press, 1953.
Reference Source

Jackson K: **Gemination and the Spirant Mutation.** *Celtica.* 1960; **5**: 127–134.
Reference Source

Jackson KH: **A Historical Phonology of Breton**. Dublin Institute for Advanced Studies, 1967.
Reference Source

Janson T: **Vocative and the grammar of calls**. In: Barbara Sonnenhauser und Patrizia Noel Aziz Hanna (eds.), *Vocative! Addressing between System and Performance. [= Trends in Linguistics. Studies and Monographs 261]*. Boston - Berlin: Mouton de Gruyter, 2013; **261**: 219–234.
Publisher Full Text

Jasanoff J: **The origin of the Celtic comparative type OIr. tressa, MW trech stronger.** *Die Sprache.* 1988–90; **34**: 171–189.
Publisher Full Text

Jasanoff J: **Double nasal presents.** *Indo-European Linguistics.* 2022; **10**: 88–106.
Reference Source

Jordán Cólera C: **Lengua y Epigrafía Celtibéricas. 2 vols**. Zaragoza: Prensas de la Universidad de Zaragoza, 2019.
Reference Source

Jørgensen A: **Celtic.** In: Thomas Olander (ed.), *The Indo-European Language Family. A Phylogenetic Perspective.* Cambridge: Cambridge University Press, 2022; 135–151.
Publisher Full Text

Jørgensen AF: **A possible substrate feature in some Celtic and Gallo-Romance bird names**. In: Guus Kroonen (ed.), *Sub-Indo-European Europe: Continuity and discontinuity in the linguistic landscape of prehistoric Europe*. Mouton De Gruyter, forthc.

Kim RI: **The Celtic feminine numerals 3' and 4' revisited.** *Keltische Forschungen.* 2008; **3**: 143–167.
Reference Source

Kluge F: **Die germanische consonantendehnung.** *Paul und Braune Beiträge zur Geschichte der deutschen Sprache und Literatur.* 1884; **9**: 149–186.
Publisher Full Text

Koch JT: **Neo-Brittonic Voiceless Spirants from Old Celtic Geminates.** *Ériu.* 1989; **40**: 119–128.
Reference Source

Kortlandt F: **Phonemicization and Rephonemicization of the Old Irish Mutations.** *Ériu.* 1982; **33**: 73–83.
Reference Source

Kroonen G: **Etymological Dictionary of Proto-Germanic [= Leiden Indo-European Etymological Dictionary Series 11]**. Leiden - Boston: Brill, 2013: **11**.
Reference Source

Kuiper FBJ: **Gothic *bagms* and Old Icelandic *ylgr*.** *North-Western European Language Evolution.* 1995; **25**: 63–88.
Publisher Full Text

Kümmel MJ: **Das dünkt mich dunkel: Germanische etymologische Probleme.** In: Bjarne Simmelkjær Sandgaard Hansen, Benedicte Nielsen Whitehead, Thomas Olander and Birgit Anette Olsen (eds.), *Etymology and the European Lexicon. Proceedings of the 14th Fachtagung der Indogermanischen Gesellschaft, 17-22 September 2012.* Copenhagen. Wiesbaden: Reichert. 2016; 219–233.

Kümmel M, Rix H: **Lexikon der Indogermanischen Verben. Die Wurzeln und ihre Primärstammbildungen. Zweite, erweiterte und verbesserte Auflage**. Wiesbaden: Reichert, 2001.
Reference Source

Kuryłowicz J: **Morphological Gemination in Keltic and Germanic**. In: Ernst Pulgram (ed.), *Studies Presented to Joshua Whatmough on his Sixtieth Birthday*. 'S-Graven-hage: Mouton, 1957; 131−144.

Lash E: **A quantitative analysis of *e/i* variation in Old Irish *etar* and *ceta*.** *Ériu.* 2017; **67**: 141−167.
Publisher Full Text

Lechleitner G, Remmer U: **Celtic Recordings - Ireland, Wales, Brittany, Isle of Man and Scotland (1907-09) [= Sound Documents from the Phonogrammarchiv of the Austrian Academy of Sciences. The Complete Historical Collections 1899-1950.** Series 5/2: The Collections of Rudolf Trebitsch. OEAW PHA CD 14]. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 2004.

Lühr R: **Fälle von Doppelkonsonanz im Keltischen: Zur Frage ihrer Genese.** *Sprachwissenschaft.* 1985; **10**(3−4): 274–346.
Reference Source

Macalister RAS: **Corpus Inscriptionum Insularum Celticarum (CIIC).** Dublin: Stationery Office, 1945; **1**.
Reference Source

Martinet A: **Celtic Lenition and Western Romance Consonants.** *Language.* 1952; **28**(2): 192–217.
Publisher Full Text

Masson O: **Géminations expressives dans l'anthroponymie grecque.** *Bulletin de la Société de Linguistique de Paris.* 1986; **81**: 217–229.
Publisher Full Text

Matasović R: **Etymological Dictionary of Proto-Celtic [= Leiden Indo-European Etymological Dictionary Series 9]**. Leiden - Boston: Brill, 2009; **9**.
Reference Source

Matasović R: **Language of the bird names and the pre-Indo-European substratum.** In: Romain Garnier (ed.), *Loanwords and Substrata. Proceedings of the Colloquium Held in Limoges (5th– 7th June, 2018).* [= Innsbrucker Beiträge zur Sprachwissenschaft 164]. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck, 2020; 331–344.

Matras Y: **The borrowability of structural categories.** In: Yaron Matras and Jeanette Sakel (eds.), *Grammatical Borrowing in Cross-Linguistic Perspective.* Berlin - New York: De Gruyter Mouton, 2007; 31–74.
Publisher Full Text

Matras Y, Adamou E: **Word classes in language contact.** In: Eva Van Lier (ed.), *The Oxford Handbook of Word Classes.* Oxford: Oxford University Press. In press.

Mayrhofer M: **Indogermanische Grammatik. Band I: 2. Halbband: Lautlehre. [Segmentale Phonologie des Indogermanischen]**. Heidelberg: Universitätsverlag Winter, 1986.

McCone K: **The Indo-European Origins of the Old Irish Nasal Presents, Subjunctives and Futures [= Innsbrucker Beiträge zur Sprachwissenschaft 66]**. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck, 1991; **66**.
Reference Source

McCone K: **Zum Ablaut der keltischen *r*-Stämme**. In: Benedicte Nielsen Whitehead and Jens Elmegård Rasmussen (eds.), In: *Honorem Holger Pedersen. Kolloquium der In-do-germanischen Gesellschaft vom 26. bis 28. März 1993 in Kopenhagen*. Wiesbaden: Ludwig Reichert, 1994; 275–284.

McCone K: **Towards a Relative Chronology of Ancient and Medieval Celtic Sound Change [= Maynooth Studies in Celtic Linguistics 1]**. Maynooth: Department of Old and Middle Irish, St. Patrick's College, Maynooth, 1996; **1**.
Reference Source

McCone K: **Double Nasal' Presents in Celtic, and Old Irish *léicid* 'leaves.** In: Jay Jasanoff, H. Craig Melchert and Lisi Oliver (eds.), *Mír Curad. Studies in Honor of Calvert Watkins*. [= Innsbrucker Beiträge zu Sprachwissenschaft 92]. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck, 1998; 465−476.

McCone K: **Mögliche nicht-indogermanische Elemente in den keltischen Sprachen und einige frühe Entlehnungen aus indogermanischen Nachbarnsprachen.** In: Olav Hackstein, Gerhard Meiser (eds.), *Sprachkontakt und Sprachwandel: Akten der XI. Fachtagung der Indogermanischen Gesellschaft, 17.– 23. September 2000, Halle an der Saale*. Wiesbaden: Reichert, 2005; 395–435.

Menard M: **Devri: Le dictionnaire diachronique du breton.** Rennes: Kuzul ar Brezhoneg, 2016.

Motta F: **Un relitto morfologico celtiberico.** *AION Linguistica*. 1981; **3**: 271–275.
Reference Source

Nance R: **A new Cornish-English dictionary.** St. Ives: Lanham, 1938.
Reference Source

O'Brien MA: **Short notes. 3. *uchtlach* 'lapful', *littiu* 'porridge.** *Celtica*. 1954; **2**: 353.

O'Brien MA: **Etymologies and notes. 13. O.Ir. *aitenn*; Welsh *eithin*.** *Celtica*. 1956; **3**: 177.

Pedersen H: **Vergleichende Grammatik der keltischen Sprachen. 2 Bände. I. Einleitung und Lautlehre. II. Bedeutungslehre (Wortlehre).** Göttingen: Vandenhoeck & Ruprecht, 1909–1913.

Piette JRF: **French loanwords in Middle Breton.** Cardiff: University of Wales Press, 1973.
Reference Source

Pokorny J: **Indogermanisches etymologisches Wörterbuch.** 2 Bde. Tübingen – Basel: Francke Verlag, 1959.
Reference Source

Prósper BM: **Lenguas y religiones prerromanas del occidente de la península Ibérica.** [= Acta Salmanticensia. Estudios filológicos 295]. Ediciones Universidad de Salamanca, 2002.
Reference Source

Prósper BM: **Estudios sobre la fonética y la morfología de la lengua celtibérica.** In: Francisco Villar, Blanca María Prósper, *Vascos, celtas e indoeuropeos. Genes y lenguas*. [= Acta Salmanticensia. Estudios filológicos 307]. Ediciones Universidad de Salamanca, 2005; 153−364.

Repanšek L: **Keltska dediščina v toponimiji Jugovzhodnega Alpskega prostora.** Ljubljana: Založba ZRC, 2016.
Publisher Full Text

Rodway S: **Welsh *bresych* 'cabbages'.** *Indo-European Linguistics and Classical Philology*. 2019; **22/2**: 1165–1169.

Ruiz Darasse C, *et al.*: **Recueil informatisé des inscriptions gauloises.** Bordeaux, 2022.

Russell P: **A Footnote to Spirantization.** *Cambridge Medieval Celtic Studies*. 1985; **10**: 53–56.
Reference Source

Russell P: **The Celtic Preverb **uss* and Related Matters.** *Ériu*. 1988; **39**: 95−126.
Reference Source

Schaffner S: **Lateinisch *rutilus*, rötlich, gelbrot, goldgelb', altirisch *ruithen* 'Strahl, Glanz' und mittelkymrisch *rwt* Rost, Korrosion.** *Die Sprache*. 2016–7; **52/1**: 102–123.
Reference Source

Schumacher S: **The Development of Proto-Celtic **au* in British Celtic.** In: Simon Rodway, Jenny Rowlands, and Erich Poppe (eds.), *Celts, Gaels, and Britons.* Studies in Language and Literature from Antiquity to the Middle Ages in Honour of Patrick Sims-Williams [= Medieval Texts and Cultures of Northern Europe 35]. Turnhout: Brepols, 2022; 187–209.
Publisher Full Text

Schmitt R: **Morphologie der Namen: Vollnamen und Kurznamen bzw. Kosenamen im Indogermanischen; Entwicklung der Namen in älteren indogermanischen Sprachen.** In: Ernst Eichler, Gerold Hilty, Heinrich Löffler, Hugo Steger und Ladislav Zgusta (eds.), *Namenforschung. Ein internationales Handbuch zur Onomastik*. 2 Bände. Berlin–New York: De Gruyter, 1995; **419–427**: 616–636.
Publisher Full Text

Schrijver P: **Animal, vegetable and mineral: some Western European substratum words.** In: Alexander Lubotsky (ed.), *Sound Law and Analogy. Papers in Honor of Robert S. P. Beekes on the occasion of his 60th birthday.* Amsterdam: Rodopi, 1997; 293–316.

Schrijver P: **On the development of vowels before tautosyllabic nasals in Primitive Irish.** *Ériu*. 1993; **44**: 33–52.
Reference Source

Schrijver P: **Studies in British Celtic Historical Phonology.** [= *Leiden Studies in Indo-European* 5] Amsterdam − Atlanta: Rodopi, 1995.
Reference Source

Schrijver P: **Spirantization and Nasalization in British.** *Studia Celtica.* 1999; **33**: 1–19.

Schrijver P: **Varia V. Non-Indo-European surviving in Ireland in the first millenium A.D.** *Ériu*. 2000; **51**: 195–199.
Reference Source

Schrijver P: **The etymology of Welsh chwith and the semantics and morphology of PIE *k(ʷ)sweibʰ-.** In: Paul Russell (ed.), *Yr Hen Iaith. Studies in Early Welsh.* [= Celtic Studies Publications 7]. Aberystwyth: Celtic Studies Publications, 2003; 1–23.

Schrijver P: **Varia I. More on non-Indo-European surviving in Ireland in the first millennium A.D.** *Ériu.* 2005; **55**: 137–144.
Reference Source

Schrijver P: **Old British.** In: Elmar Ternes (ed.), *Brythonic Celtic − Britannisches Keltisch. From Medieval British to Modern Breton*. [= Münchner Forschungen zur historischen Sprachwissenschaft 11], Bremen: Hempen Verlag, 2011; 1–84.

Schrijver P: **Sound Change, the Italo-Celtic Linguistic Unity, and the Italian Homeland of Celtic.** In: John T. Koch and Barry Cunliffe (eds.), *Celtic from the West 3. Atlantic Europe in the Metal Ages: Questions of Shared Language*. [= Celtic Studies Publications 19]. Oxford: Oxbow Books, 2016; 489–502.

Schrijver P: **The Development of Proto-Celtic **st* in British Celtic.** In: Simon Rodway, Jenny Rowland, and Erich Poppe (eds.), *Celts, Gaels, and Britons*. *Studies in Language and Literature from Antiquity to the Middle Ages in Honour of Patrick Sims-Williams.* [= Medieval Texts and Cultures of Northern Europe 35]. Turnhout: Brepols, 2022; 187–209.

Schumacher S: **Die keltischen Primärverben. Ein vergleichendes, etymologisches und morphologisches Lexikon.** *Unter Mitarbeit von Britta Schulze-Thulin und Caroline aan de Wiel* [= Innsbrucker Beiträge zur Sprachwissenschaft 110]. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck, 2004.
Reference Source

Sims-Williams P: **Dating the Transition to Neo-Brittonic: Phonology and History, 400–600.** In: Alfred Bammesberger and Alfred Wollmann (eds.), *Britain 400–600: Language and History.* Heidelberg: Winter, 1990; 217–261.
Reference Source

Sims-Williams P: **The problem of spirantization and nasalization in Brittonic Celtic.** In: Alexander Lubotsky, Jos Schaeken, Jeroen Wiedenhof (eds.), *Evidence and Counter-Evidence. Essays in Honour of Frederik Kortlandt Volume I:* Balto-Slavic and Indo-European Linguistics [= Studies in Slavic and General Linguistics 32]. Leiden: Brill, 2008; **32**: 509–525.
Reference Source

Sims-Williams P: **Irish Influence on Medieval Welsh Literature.** Oxford University Press, 2011.
Reference Source

Sjoestedt ML: **L'aspect verbal et les formations a affixe nasal en celtique.** Paris: Librairie Champion, 1926.
Reference Source

Sommerfelt A: **Consonant Quantity in Celtic.** *Norsk Tidsskrift for Sprogvidenskap.* 1954; **17**: 102–118.

Stifter D: **Chapter 4. Early Irish.** In: *The Celtic Languages.* 2nd Edition. Edited by Martin J. Ball and Nicole Müller, Abingdon – New York: Routledge, 2009; 55–116.
Reference Source

Stifter D: **Study in Red.** *Die Sprache.* 1998; **40/2**: 202–223.

Stifter D: **The Invisible Third. The Basque and Celtic Words for 'Swallow'.** *Ériu.* 2010; **60**: 145–157.
Publisher Full Text

Stifter D: **The textual arrangement of Alise-Sainte-Reine [L-13].** *Zeitschrift für celtische Philologie.* 2011; **58**(1): 165–181.
Publisher Full Text

Stifter D: **Lack of Syncope and other nichtlautgesetzlich Vowel Developments in OIr. Consonant-Stem Nouns. Animacy Rearing its Head in Morphology?** In: Thomas Krisch and Thomas Lindner (eds.), *Indogermanistik und Linguistik im Dialog. Akten der XIII. Fachtagung der Indogermanischen Gesellschaft. 2008 Salzburg.* Wiesbaden: Ludwig Reichert Verlag, 2011b; **2011**: 556–565.

Stifter D: **Lenition of **s* in Gaulish?** In: Benedicte Nielsen Whitehead, Thomas Olander, Birgit A. Olsen and Jens Elmegård Rasmussen (eds.), *The Sound of Indo-European. Phonetics, Phonemics and Morphophonemics.* Copenhagen: Museum Tusculanum Press, 2012; 523–544.
Reference Source

Stifter D: **Vocative for nominative**. In: *Vocative! Addressing between System and Performance.* Hrsg. von Barbara Sonnenhauser und Patrizia Noel Aziz Hanna [= Trends in Linguistics. Studies and Monographs 261]. Boston – Berlin: Mouton de Gruyter, 2013; 43–85.
Publisher Full Text

Stifter D: **The Phonology of Celtic**. In: *Handbook of Comparative and Historical Indo-European Linguistics. An International Handbook.* Edited by Jared S. Klein, Brian Joseph, Matthias Fritz [= Handbücher zur Sprach- und Kommunikationswissenschaft 41/2]. Berlin – New York: Walter de Gruyter, 2017; 1188–1202.
Publisher Full Text

Stifter D: **Varia II. The origin of time.** *Ériu.* 2017b; **67**: 219–226.

Stifter D: **The stars look very different today.** *Ériu.* 2018; **68**: 29–54.
Reference Source

Stifter D: **The etymology of Breton *gad* 'hare'.** *Études Celtiques.* 2021; **47**: 171−177.
Reference Source

Stifter D: **12. With the Back to the Ocean: the Celtic Maritime Vocabulary**. In: Kristian Kristiansen, Guus Kroonen & Eske Willerslev (eds.), *The Indo-European Puzzle Revisited.* Cambridge University Press, 2023a; 172–192.
Publisher Full Text

Stifter D: **Prehistoric layers of loanwords in Old Irish**. In: Guus Kroonen (ed.), *Sub-Indo-European Europe: Continuity and Discontinuity in the Linguistic Landscape of Prehistoric Europe.* Mouton De Gruyter, forthcoming.

Stifter D, Braun M, Salomon C, *et al.*: ***Lexicon Leponticum* (Vienna)**. 2009.
Reference Source

Stokes W: **On the assimilation of pretonic *n* in Celtic suffixes.** *Indogermanische Forschungen.* 1891; **2**: 167–173.
Publisher Full Text

Stokes W: **On the assimilation of pretonic *n* in Celtic suffixes.** *Trans Philol Soc.* 1891–3; **22**(2): 297–307.
Publisher Full Text

Strachan J: **The Compensatory Lengthening of Vowels in Irish.** *Trans Philol Soc.* 1891−4; **22**(1): 217−254.
Publisher Full Text

Stüber K: **The Historical Morphology of n-Stems in Celtic.** Maynooth: Department of Old and Middle Irish, 1998.
Reference Source

Stüber K: **Individualisierendes **-on-* in Namen von Göttern und Göttinnen.** *International Journal of Diachronic Linguistics and Linguistic Reconstruction.* 2004; **1**: 1–17.

Stüber K: **Die Verbalabstrakta des Altirischen**. 2 vols. [= Münchner Forschungen zur historischen Sprachwissenschaft 15]. Bremen: Hempen, 2015.
Reference Source

Stüber K, Zehnder T, Remmer U: **Indogermanische Frauennamen.** Heidelberg: Universitätsverlag Winter, 2009.
Reference Source

Testen D: **Stem-final **-KK-* in Celtic terms for 'pig'.** *Ériu.* 1999; **50**: 161–164.
Reference Source

Thes. = Stokes, Whitley, John Strachan (eds.), **Thesaurus Palaeohibernicus. A Collection of Old-Irish Glosses Scholia Prose and Verse.** 2 vols. Cambridge University Press: 1901-3.

Thomas PW: **The Brythonic Consonant Shift and the Development of Consonant Mutation.** *Bulletin of the Board of Celtic Studies.* 1990; **37**: 1–42.

Thurneysen R: **A Grammar of Old Irish. Revised and enlarged edition.** Dublin: Dublin Institute for Advanced Studies, 1946.
Reference Source

Toner G, *et al.*: **An Electronic Dictionary of the Irish Language, based on the Contributions to a Dictionary of the Irish Language (Dublin: Royal Irish Academy, 1913–1976)**. 2019.

Van Sluis P: **Lenited voiceless stops in Middle Welsh: phonology and orthography.** PhD-thesis, Aberystwyth University, 2019.
Reference Source

Van Sluis P: **Substrate alternations in Celtic.** In: Guus Kroonen (ed.), *Sub-Indo-European Europe: Continuity and Discontinuity in the Linguistic Landscape of Prehistoric Europe.* Mouton De forthc.

Van Sluis P, Anders Richardt J, Guus K: **European Prehistory between Celtic and Germanic: the Celto-Germanic isoglosses revisited.** In: Kristian

Kristiansen, Guus Kroonen & Eske Willerslev (eds.), *The Indo-European Puzzle Revisited.* Cambridge University Press, 2023; 193–243.
Reference Source

Vendryes J, Bachellery É, Lambert PY: **Lexique étymologique de l'irlandais ancien**. Dublin - Paris: DIAS - CNRS, 1959.
Reference Source

Weiss M: **The Inner Revolution: Old but Not That Old.** In: Matilde Serangeli and Thomas Olander (eds.), *Dispersals and Diversification*. Linguistic and Archaeological Perspectives on the Early Stages of Indo-European [= Brill's Studies in Indo-European Languages & Linguistics 19]. Leiden: Brill, 2020; 263–286.
Publisher Full Text

Wheatley D, Iosad P: **Acoustic correlates of the fortis/lenis distinction in early 20th century Donegal Irish.** handout of a paper delivered at the Tionól of the School of Celtic Studies, Dublin Institute for Advanced Studies, 2021.
Reference Source

Wmffre I: **Post-Roman Settlement in Wales: New Insights from a Recent Study of Cardiganshire Place-Names.** In: Karen Jankulak and Jonathan M. Wooding (eds.), *Ireland and Wales in the Middle Ages*. Dublin: Four Courts Press, 2007; 46–61.
Reference Source

Wodtko DS: **Monumenta Linguarum Hispanicarum. Band V.1. Wörterbuch der keltiberischen Inschriften.** Wiesbaden: Reichert, 2000.
Reference Source

Wodtko DS, Irslinger B, Schneider C: **Nomina im Indogermanischen Lexikon.** Heidelberg: Universitätsverlag Winter, 2008.
Reference Source

Zair N: **The Reflexes of the Proto-Indo-European Laryngeals in Celtic.** [= Brill's Studies in Indo-European Languages & Linguistics 7]. Leiden − Boston: Brill, 2012; 7.
Reference Source

Zimmer S: **Studies in Welsh Word-Formation.** Dublin: Dublin Institute for Advanced Studies, 2000.
Reference Source

Zupitza E: **Über doppelkonsonanz im Irischen.** *Zeitschrift für vergleichende Sprachforschung auf dem Gebiete der Indogermanischen Sprachen* 36/2. 1900; 202–245.
Reference Source

Open Peer Review

Current Peer Review Status: ✔ ✔ ✔ ✔

Version 2

Reviewer Report 08 June 2024

https://doi.org/10.21956/openreseurope.18478.r38030

© **2024 Jørgensen A.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Anders Jørgensen**
Disciplinary Domain of Humanities and Social Sciences, Faculty of Arts, Department of Archaeology and Ancient History, Uppsala University, Uppsala, Sweden

The author has taken the peer reviews into account when preparing the second version. I have nothing to remark this time around so nothing stands in the way.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Comparative and historical linguistics pertaining to Celtic (mainly Breton) and the Indo-European languages more broadly.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 03 April 2023

https://doi.org/10.21956/openreseurope.16648.r30786

© **2023 Roma E et al.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

✔ **Elisa Roma**
[1] Dipartimento di Studi Umanistici - Sezione di Linguistica teorica e applicata, Università degli Studi di Pavia, Pavia, Italy

[2] Dipartimento di Studi Umanistici - Sezione di Linguistica teorica e applicata, Università degli Studi di Pavia, Pavia, Italy
Gianguido Manzelli
[1] Dipartimento di Studi Umanistici, Universita degli Studi di Pavia, Pavia, Lombardia, Italy
[2] Dipartimento di Studi Umanistici, Universita degli Studi di Pavia, Pavia, Lombardia, Italy

Overview
This article presents consistent and comprehensive evidence to support a scenario of the development of the reconstructed consonantal phoneme inventories of Proto-Indo-European and Proto-Celtic that motivates the emergence of phonologically long (aka geminate) consonants in the Celtic branch and the later restructuring of length oppositions in Insular Celtic languages.

The article argues extensively and mostly convincingly for the emergence of long consonants (predominantly resonants and secondarily voiceless stops) in Proto-Celtic mainly as a result of assimilation of consonant clusters, both original clusters and, notably, clusters arising through word-formation processes and therefore morpheme encounters. These developments are dealt with in detail distinguishing common phonetic outcomes within the Celtic branch from divergent outcomes of original clusters (Sections 3 and 4) and different chronologies and kinds of word-formation processes involved (Sections 5 and 6). The paper also highlights the relatively smaller evidence for long voiced plosives in Proto-Celtic and in earlier attested Celtic languages such as Gaulish, a fact which might be related to the divergent developments of voiceless and voiced stops in lenition contexts in the two Insular Celtic branches.

The author suggests that reconstructible long consonants for which etymologies cannot be grounded on attested forms in Indo-European languages can be due to other factors, such as affective connotations and sound symbolism, especially in word classes and semantic fields where these phenomena are purportedly more frequent (address forms of kinship terms and personal names, adjectives denoting physical qualities). What cannot be arranged in any of these files is reasonably, albeit tentatively, ascribed to prehistoric borrowing (lexical copying) from unknown substratal languages. A merit of the article is that it thoroughly engages with the reconstruction and motivation of forms for which long consonants might be based on affective connotations, sound symbolism, and even borrowing (Sections 9, 11, 13), so that, in an attempt to avoid circularity, these residual groups are not just unaccounted for left-overs.

Doubtful issues and problematic cases are not concealed (Sections 10 and 12).

The comments that follow are therefore not to be meant as major criticims, but rather as suggestions that, it is hoped, may help improve the clarity and persuasiveness of the paper. They are grouped under three headings: 1) request for consistent terminology, 2) comments on specific points, 3) typos and minor slips.

Reference is made below to pages in the pdf file and to numbered sections; abbreviations are the same as the paper's, if not given in full.

1) In general, there are a few **terminological inconsistencies**, probably due to cumulative drafting, that could be addressed, to the reader's benefit:

I suggest that a single term, either “plosive” or “stop” or “occlusive”, be used throughout. If these terms are not synonyms in the author's usage, this should be clarified. Similarly for “fricatives” and “spirants” (e.g. p. 20 § 6.2 “spirant sounds”). “Obstruents” is yet another term that in some loci but possibly not throughout the paper is used for stops/ plosives/ occlusives (e.g. cf. p. 22 “a geminate voiced obstruent resulted”, p. 36 “intervocalic obstruents and fricatives”), while obstruent is in fact a term that includes both stops and fricatives.

Either “velar” or “guttural” should be chosen (again, there does not seem to be any difference between their usage, so a single term would be preferable).

For the mid-high front vowel in Goedelic, ɪ is used on p. 23, which is the IPA symbol for mid-high front (unrounded), while on p. 22 (§ 7 and § 7 (4)) a different symbol ɩ, possibly rather used for high central unrounded, is used instead.

On a slightly different note, although the widespread terminology “mediae and mediae aspiratae” (p. 6) for the consonant inventory of PIE has a long tradition, it could be improved through greater consistency, allowing conversion with the terms used for the daughter languages and in general.

2) Comments on specific points:

p. 4 (Preliminaries): “**ballo*-, which is meant as equivalent to a phonetic analysis as [bal:o-].” The phonetic transcription [bal:o-] (rather [bal:o]-) here seems to imply that a long consonant in a reconstructed form is not considered phonologically relevant (e.g. an allophone between vowels in this example); this is not generally the case, as reconstructed forms are rather closer to broad, phonemic transcriptions (put between slashes //). Anyway, this point should be better clarified. Also, it should be clarified if any stress and syllabic border is meant or implied in reconstructed forms.

p. 4 (Preliminaries): “The main contrast between the two classes of Celtic stops is considered to be between ‘voiced’ (= D) and ‘unvoiced/voiceless’ (= T) consonants”. It should be made clear from the outset (and not simply implied or taken for granted) whether the main phonological contrast between the two classes of stops (voiced vs. unvoiced) for the purposes of the paper is deemed corresponding, i.e. not a different opposition, to a phonetic contrast between singleton and long (= lenis and fortis respectively). This is not so, but what follows here seems to suggest the contrary.

p. 4 (Preliminaries) and throughout: The label "quasi-" in qPIE and qPC is slightly misleading in this context, as it does not correspond to usual practice both for attested and reconstructed languages: in fact, modifiers distinguishing different phases usually point to chronological order (pre-, post-, early, late, old, middle, modern), while "quasi" refers to an approximation from either direction, or is rather synonymous with pre-. Therefore, quasiPC could be meant as a phase of late Indo-European approximating the separation of the Celtic branch. Clearer labels would be useful.

p. 6, Table 1: p is inserted in the bilabial fricative case, but I suppose this is a slip: a fricative should be represented by a symbol for a fricative (e.g. for a voiceless bilabial fricative [ɸ]); [p] is a plosive that possibly lies behind a fricative consonant at this stage. This development from PIE has not been mentioned here before, however (see the paragraph listing the main changes in the consonant system from PIE to Proto-Celtic - beginning with “After” and ending with “Table 1 in

Section 1", p. 6). So, either p should be put in the right column (plosive) or fricativisation of the bilabial plosive should be mentioned and a fricative put in this slot in Table 1 (see § 2.1 (14) for an instance of this development).

p. 7 § 2.1 (9) "**salH̯†nH̯ī-*": The use of the symbol † to signal the position where a vowel has undergone syncope in the prehistory of Irish should be mentioned.

p. 8 § 2.2 (4) "Gaul. **cammīnos*, which survives as a loanword in Ital. *cammino*, Fr. *chemin*, etc. 'path, way'": "as a loanword" sounds odd for a reconstructed Gaulish word that survives in French and Italian.

p. 13 § 3.7 (2): "... for 'eye' is PIE **suh₂lih₁* 'two suns'". Beside the translation, a clarification such as "the PIE dual form..." would be useful here, since when an inflected, non-nominative, non-singular (and non-masc. for adjectives) PIE form is given, its case/number (gender) must be specified.

p. 14 § 3.9.1 (3): "constitutes the root of OIr. impersonal verb *do·cuäs* 'has gone'." It is rather an impersonal verb FORM, traditionally a singular passive (perfect) form, something which is different from an impersonal verb (i.e. a verb that is not construed with a nominative argument). The point is not totally irrelevant here since the Indo-European formation with **-to-* lies regularly behind OIr. preterite/perfect passive forms. From the point of view of both their etymology and the Old Irish paradigms, the examples in § 3.9.2 (4) and § 3.9.5 (3) are the same; it is only the English translations that need to resort to "impersonality" with an intransitive verb such as 'to go'. Perhaps a different translation, anyhow, could be envisaged, e.g. 'there has gone', 'someone has gone'.

p. 16 § 3.10 (3) "not ***-aggo-* > OIr. ***-ac*": rather OIr. **-<*ac*> = -/ag/ (see § 4.1 for = /d/ through the stage **dd*)

p. 17 § 5.1 (3) "one can think of a formation from the root **h₃reḭH-* 'to whirl', perhaps from a hypothetical nasal-infixed verb **rinati* 'to run', comparable with Goth. *rinnan* < **h₃ri-n-H-*, with the addition of the adjectival suffix **-no-* to the present stem. Alternatively, the formation could derive from the noun **reḭno-* 'great flowing mass of water' < **h₃reḭ(H)-no-*, with ablaut and with the addition of a second nasal suffix." The two alternatives appear to be highly speculative: can similar formations be invoked to support either hypothesis?

p. 17 § 5.1 (6) "Old Irish has two distinct verbs that influenced each other formally: in PIE **sn̥-n(e)-h₂-* 'to obtain' [...] PIE **suů̯h₂-* 'to sound' > **suu̯ana-* acquired geminate **nn* probably after its model, i.e., **suu̯anne/o-* > OIr. *seinnid* 'to sound, make music'." This kind of analogy at an early stage is frankly hard to believe, if the verbs were unrelated and, at the time when gemination was taken over by the second verb, length was distinctive for **n*. Although mutual influence between the two verbs can be conceived, either a different explanation could be envisaged for geminate *n* in *seinnid* 'play music', or a later analogy should be invoked.

p. 18 § 6.2 "An equation encompassing Gaulish and Irish..." Why "Gaulish and Irish" if Galatian is also involved?

p. 21 § 6.3 (1) "Alternatively **kuu̯etur-* could be an archaism preserved in this compound, or *petor-* is simplified from the younger Gaulish form **petuu̯or-* with influence from the cardinal **kuu̯etuu̯ores* '4'."

Since these alternatives are simpler, it would be better to mention them first, and then put forward the hypothesis which implies two successive metatheses.

p. 22 (end of § 6.3) “There was a contrast between *s* and a strong sibilant…” If this is mirrored in the opposition between *s* and *ts*, it seems rather an instance of fricative vs. affricate. Otherwise, this point could be better qualified.

p. 22 (end of § 6.3) “On the whole, geminated voiced stops were rare except across the morpheme boundary.” If that is so, it could be wiser to put geminated voiced stops on a different level as opposed to geminated voiceless stops: Table 2 puts all geminated consonants except nasals and liquids between brackets, which means that they are all considered allophones of the corresponding singletons, but if geminated voiced stops are so infrequent and limited to morpheme boundaries, one wonders whether they can be put on a par with the other allophones and should not be considered simply consonant clusters (where sub-phonemic alternations were rather between assimilated and non-assimilated clusters, than between short and long consonants). The point is that consonant encounters in word formation do not automatically create geminated allophones, but only (allowed) consonant clusters. The absence of true geminated allophones for voiced stops at this stage could be linked to their later divergent developments (§ 6.2).

p. 22 § 7 “fell together as *é*.” Rather <é> = /eː/; as elsewhere, the spelling conventions of the different languages should not be taken for granted when describing phonetic changes. Therefore, also after “initial *éc*- and *ét*-”, add = /eːg/- and /eːd/-.

p. 22 § 7 (4) *ceta*- is also a preverb with the meaning 'first' (perhaps more commonly), so a different example would be preferable.

p. 23, fn. 3 “3rd singular of the copula”: insert “present indicative”

p. 24 Caption of Table 3: “preverbal 'in'” looks strange, as forms between single inverted commas are usually translations (is this what is meant here?)

On a different note, the reconstructed forms in the Table are not straightforward, since allomorphy of the preverb is not clearly grounded on the basis of reconstructed contexts but rather on the basis of OIr. outcomes and in particular stress positioning. An independent reason, either phonetic or other, should be put forward for each previous stage allomorphy. Also, at this point it is not entirely clear how relevant the reconstruction of these allomorphs is for the main purpose of the paper. Since the discussion is meant to justify why **in*+voiceless stop does not turn into *i/e*+voiced stop in OIr. in some cases, the starting point and the detour of the discussion should take a different stance.

p. 24 fn. 5 This reconstruction assumes that **et* is the particle that occurred in main clauses between preverb and verb-root and after a simple inflected verbal form. This option should be clarified.

p. 24 (§ 7) “**igg*- before a back vowel > **egg*-.” Why a back vowel and not a lower non-front vowel? Does this occur with a following *u*?

p. 25 § 8.1 “the outcome *ƒ* (phonologically equivalent to a geminated **u̯!*) is identical to that of the internal group **-su̯-*, but is not phonetically similar to single **u̯*, which rather gets lost.” It gets lost if internal, but not if initial.

p. 26 § 8.2 While it is generally acknowledged that gemination did not play a role for the opposition between lenition and non-mutation, as it did in Goedelic, it is commonly held that it played a role for Welsh fricativisation of voiceless consonants (voiceless spirantisation; see the summary in Willis 2009 and § 14.2 in this article). At least some lines here about this development would clarify the picture and the author's stance in its regard and in relation to Section 14.2.

p. 27 § 9.2 “in the root or in the semantically meaningful portion of the names”. Suffixes and endings are also meaningful (as all morphemes, by definition), so here it would be better to put the morphological predilection in different terms (e.g. semantically intense part?). Usually content vs. function word is used for lexemes, but that would not easily work for morphemes in name formation.

p. 27 § 9.2 “in Gaulish short names or in names formed with suffixes.” This should be clarified: it is not entirely clear to what kind of names these are contrasted (names formed by composition, supposedly?).

p. 27 § 9.2 “Dithematic names”. Does this mean compound?

p. 27 § 9.2 “came to contain a relatively high proportion of geminate stops, in contrast to the rest of the lexicon, where gemination was most prominent among resonants.” This appears to be true for voiceless geminate stops, but not voiced ones (as per Table 4).

p. 28 § 10 “The reverse of this rule or tendency is the simplification or degemination of geminates in pretonic position”. This simplification would not be warranted in the second scenario just mentioned, i.e. that gemination after penultimate stressed syllable is due to Latin stress position rules.

p. 29 § 11.1 “restrict the category of symbolic gemination in Celtic to two types of words: onomatopoetic nouns (in fact a single example) and adjectives.” Cross-linguistic parallels for these kinds of expressive gemination would be useful for the sake of the argument.

p. 29 § 11.3.1 “Only a subset of adjectives with geminates is of more or less good Indo-European inheritance. They are being repeated in this panorama for the sake of the overall argument, even though they have already been discussed in previous sections.” It is not entirely clear why forms where gemination emerges as the outcome of sound change are listed in this section where expressive gemination in adjective bases is deemed to occur and where one expects to find instances of adjectives where “phonologically obscure gemination” is identified (beginning of Section 11.3). Rather than supporting the overall argument, it puzzles the reader.

p. 30 § 11.3.2 (9) “qPC **kokko-*” This word (with geminate consonant) is attested in Greek and Latin (and attributable to Galatian, Freeman 2001: 16), and probably refers primarily to the kermes oak berry (Delamarre 2003: 120 ‘grain rouge, cochenille’), so one wonders whether it can be an instance of expressive gemination in adjectives. The development of the colour meaning is secondary.

p. 30 § 11.3.2 (11) **messamo-*. It is not entirely clear to me why this item is put here. Its formation is different from the other cases in this Section, as gemination in all likelyhood occurs as a result of word formation.

p. 30 § 11.3.2 (12) "Early ModIr. *cittach* 'lefthanded, clumsy'." Are there no other related words?

p. 31 § 11.3.2 (15) OIr. *tricc*. Should not this go in the following Section 11.3.3?

p. 31 § 11.3.3 "They relate to basic physical concepts; to physical defects; and they are symbolic of swiftness and agitation." Could taboos have played a role?

p. 31 end of § 11.3.3 "Another reason for ascribing these words to Urschöpfungen within Celtic, and not to loans from unknown languages, has to do with the typological tendency of adjectives to be less amenable to borrowing than nouns." This is a crucial point in the argument and therefore, although it is commonplace, it deserves non-trivial mention and updated references (e.g. Matras 2007, Matras & Adamou in press).

p. 35 § 14 "*mak^wk^wo-, or rather sequences of velars + labiovelars, i.e., *$makk^wo$-." From a phonetic point of view, it is rather unlikely that secondary articulations like labialisation occurred twice, so to say. Generally, geminate consonants are produced with a longer constriction duration than singletons (Ladefoged and Maddieson 1996); phonological gemination of complex articulations such as affricates, for example, usually consist of a longer stop gesture with delayed release (e.g. [t:s]) rather than two successive articulations stop+fricative ([tsts], Ladefoged and Maddieson 1996: 92). Anyhow, typological parallels for gemination including two successive labialisations should be mentioned in support of the first option. (The articulation of Modern Irish palatal vs. neutral tense (long) resonants, in varieties where the length distinction is preserved, could also be compared, although this is not an exact phonetic parallel). [Note that this comment argues for the opposite view as Martin Kümmel's review. *kk^w is meant to stand in contrast to single *k^w, rather than to *k^wk^w]

p. 35 § 14.1 "However, after the contrast lenition : non-lenition had emerged as the central phonological opposition in Insular Celtic, the gemination of stops lost its phonological significance and turned into a phonetic feature that was concomitant with non-lenition, indicated by the bracketed length marks in Table 6 below." If the emergence of lenition "may well at first have been a side effect of the strong opposition between single vs. geminate consonants", after the phonologisation of lenition, geminate stops could rather lose their phonetic status of long or tense consonants, but the effects of their earlier phonologisation were maintained and continued in fact in the phonological status of (non-lenited) stops. In other words, if the opposition between single vs. geminate stop turned into fricative vs. stop, this is not a real instance of de-phonologisation, as would be the case if it were maintained only in some environments, rather a re-phonologisation. This at least is in line with what is stated at the beginning of Section 8.1.

3) Typos and minor slips:

p. 5 (previous research) Stokes > Stokes'
p. 5 (previous research) assmilation > assimilation
p. 10 § 3.1 (2) oder > or

p. 20 § 6.2 (3) bye-form > by-form
p. 22 (end of § 6.3) “in the system above”: Table 2 is in fact below.
p. 22 § 7 allophon > allophone
p. 30 § 11.3.1 (8) (OIr.) *mall* is missing
p. 30 § 11.3.2 (6) acymr. > OW
p. 30 § 11.3.2 (6) einäugig > one-eyed (?)
p. 33 § 13.1 (10) substanivised > substantivised
p. 36 (end of § 14.1) obervations > observations

References

1. Delamarre X: Dictionnaire de la langue gauloise: une approche linguistique du vieux-celtique continental. *Paris: Errance*. 2003.
2. Freeman P: The Galatian language: a comprehensive survey of the language of the ancient Celts in Greco-Roman Asia Minor. *Lewiston, NY: Mellen Press*. 2001.
3. Ladefoged P, Maddieson I: The sounds of the world's languages. *Oxford: Blackwell*. 1996.
4. Matras Yaron: The borrowability of structural categories. *In: Yaron Matras and Jeanette Sakel (eds.), Grammatical Borrowing in Cross-Linguistic Perspective. Berlin, New York: De Gruyter Mouton*. 2007. 31-74 Publisher Full Text
5. Matras Y, Adamou E: Word classes in language contact. *In: Eva Van Lier (ed.), The Oxford Handbook of Word Classes. Oxford: Oxford University Press*. In press. Reference Source
6. Willis D: Old and Middle Welsh. *In: M. J. Ball, and N. Müller (eds), The Celtic languages (2nd ed.). London, New York: Routledge*. 2009. 117-160

Is the work original in terms of material and argument?

Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?

Yes

Is the work clearly and cogently presented?

Yes

Is the argument persuasive and supported by evidence?

Yes

If any, are all the source data and materials underlying the results available?

Yes

Does the research article contribute to the cultural, historical, social understanding of the field?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Celtic Linguistics (Goedelic)

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 10 Apr 2023

David Stifter

Thank you very much for your many useful comments! David

Competing Interests: No competing interests were disclosed.

Reviewer Report 03 March 2023

https://doi.org/10.21956/openreseurope.16648.r30783

© 2023 Jørgensen A. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Anders Jørgensen
[1] Disciplinary Domain of Humanities and Social Sciences, Faculty of Arts, Department of Archaeology and Ancient History, Uppsala University, Uppsala, Sweden
[2] Disciplinary Domain of Humanities and Social Sciences, Faculty of Arts, Department of Archaeology and Ancient History, Uppsala University, Uppsala, Sweden

The article provides a thorough treatment of the various origins of geminate consonants in the older Celtic languages. The article goes into great detail and provides explicit etymologies for a great number of words. As such, it will undoubtedly be used not only as a handbook of gemination in Celtic, but also as an etymological supplement for the Celtic languages for words which just happen to contain geminates.

The article proceeds methodically, going through the various input clusters responsible for geminates in Celtic. The prose is straightforward and easy to follow. Of special interest is the discussion of possible substrate vocabulary, partially identifiable by geminates which do not allow further segmentation.

While I have no difficulty giving the paper my full endorsement, there are a few minor things that might improve the text (page number refer to the pdf version of the article):

p. 3: In the list of abbreviations, "OBret." is missing

p. 4: Quoting Cornish from the GPC might not be ideal. It would be advisable to check the attested spellings in Graves (1962), The Old Cornish Vocabulary (for Old Cornish), Low (2022), A Dictionary of Middle Cornish (for Middle Cornish minus the more recently discovered texts) and George (2020), An Gerlyver Meur (for Old, Middle and Late Cornish).

p. 4: Caution is advisable when quoting from Favereau's dictionary. He modernizes Old Breton (with accompanying phonetic transcription according to dialect!) and makes headwords out of material only *possibly* attested in e.g. place-names. It is always a good idea to check rarer words in Hemon's GIB and online in DEVRI.

p. 6, § 2.1, example (4) and following: It might be a good idea to briefly mention Hill (2010-12). It relates directly to the discussion of words with syllabic *l*+*n* here and in the nasal presents below. If Hill is correct, the proposed reconstructions here will not work.

p. 7, § 2.1, example (9): It may be worth mentioning the existence of MBret. *sallaff* 'to salt', adj. *sall* 'salted'. A derivation from French *saler* 'to salt' is possible (with Bret. *sall* 'salted' being back-formed from the verb). However, the Breton *-ll-* is surprising and the match to the proposed PCelt. form is intriguing. Could be cognate with OIr. *sall* instead.

p. 8, § 2.2, example (3): It would be better to include Middle Breton *cam* 'step' here instead of the Modern Breton derivative (especially since the Breton suffix *-enn* may cause confusion for the reader).

p. 10-11, § 3.1, example (8): The lack of syncope seems to be secondary if the reconstruction in the preceding entry is to be trusted.

p. 11, § 3.1, example (13): Reference to the discussion of Egurtzegi and Ariztimuño (2013) would be suitable here, since it appears to eliminate some of the evidence in favour of substratal status.

p. 14, § 3.9.4: delete hyphen in the heading.

p. 16, § 4.2, example (4): Does a Modern Breton *izhin* actually exist?

p. 17, § 3.7, example (3): W *melltith* is not a good example of the borrowing of Lat. *-ld-* as Brit. **-llt-* (shouldn't this rather be Brit. **-lt-*?). As shown by MBret. *malloez* < *maledictio͞, millizyen* < *maledictio͞ne-,* Bret. *millig-,* MCorn. *mylyg-* < *maledĭc-,* Welsh *melltith* is probably secondary, with learned *-d-* (devoiced by the preceding *-ll-*), as also seems to be the case with MBret. *bennoez, binizyen* < *benedictio͞ne-, binnig-* < *benedĭc-* as opposed to W *bendith, bendig-*. Note that the medial cluster in *melltith*, if one really takes it at face value, cannot have merged with older **-lt-* (from PCelt. **-lt-* and Latin *-lt-, -ld-*), because this gives MW *-ll-* medially, cf. e.g. **cal'dā̆ria* > **kaltɔr* > MW *callawr,* MBret. *cauter.*

This leaves us with *mellt* as the only reasonably solid example of preserved **-ld-* (as Brit. **-lt-*). Given that we have examples of vacillation between *-ll* and *-llt* in Middle Welsh (MW *gwellt* next to OW *guel,* MW *guell,* MW *gwyllt* next to MW *guyll, gwyll,* MW *deall* next to *dallt*), it might rather be the case that the final *-t* in *mellt* is unetymological.

p. 18, § 6.2: Breton *a-dreñv* means 'behind'.

p. 20, discussion of the outcome of voiced geminates in Brit.: Allowing an analogical treatment of clusters across morpheme boundaries may allow us to treat the outcome of **zd* as having passed through the stage **dd*, only to become devoiced to **tt* and spirantized and thus being part of a general devoicing of voiced geminate clusters (cf. Jørgensen 2022). This requires that e.g. Bret. *gad*

is a borrowing (presumably from Irish).

p. 21, example (4): It would make sense to standardize Breton *gwrec'h*, *grec'hent* to either forms in *gwr-* or *gr-*.

p. 30, § 11.3.2, example (1): It seems much more likely that the *-an* in *bychan,* etc. is the diminutive suffix, cf. the meaning.

p. 30, § 11.3.2, example (4): This contradicts the change of **gg* > **kk* > **x* presented earlier in the text. I have suggested (Jørgensen 2022: 145) that the actual reflex of **buggo-* in Breton is *bouc'h* 'blunt' and that *bouk* is a borrowing from Irish.

p. 33, § 13.1, example (26): It may be better to leave out W *croen* and OCorn. *croin*, since they cannot continue **krokken-* directly anyway, whatever the explanation. Also, as far as I can see, a reconstruction **krokkenno-* would fit both Irish and Breton better (MCorn. *crohen* tells us nothing and Gallo-Lat. *crocina* does not fit anyway and may have suffix substitution). The word consistently rhymes as having /-enn/ in Middle Breton and is mostly spelled in accordance with this. It does not otherwise behave like a singulative in *-enn*, forming its plural with internal *i*-affection, so it does not appear to have been influenced by this suffix.

References

1. Low I: A Dictionary of Middle Cornish. *Cardiff: Goldcrest Publications*. 2021.
2. George K: An Gerlyver Meur. *Kesva an Yeth Kernewek, Cornish Language Board*. 2020.
3. Graves E: The Old Cornish Vocabulary. *Columbia University PhD dissertation*. 1962.
4. Hill E: Silbische Liquiden vor Nasalen im Inselkeltischen und das Problem der Nasalpräsentien vom Typ air. sernaid, kymr. -sarnu. *Keltische Forschungen*. 2012; **5**: 157-184
5. Jørgensen A: Celtic. 2022. 135-151 Publisher Full Text
6. Egurtzegi A, Ariztimuño B: Remarks on the Etymology of the Basque Word for ‘Swallow’ and Its Potential Relation to Celtic. *ÉRIU*. 2013; **63** (-1): 79-90 Publisher Full Text

Is the work original in terms of material and argument?

Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?

Yes

Is the work clearly and cogently presented?

Yes

Is the argument persuasive and supported by evidence?

Yes

If any, are all the source data and materials underlying the results available?

Yes

Does the research article contribute to the cultural, historical, social understanding of the

field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Comparative and historical linguistics, pertaining to Celtic (mainly Breton) and the Indo-European languages more broadly.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 08 Mar 2023

David Stifter

Thank you very much for your very helpful comments!

Competing Interests: No competing interests were disclosed.

Reviewer Report 22 February 2023

https://doi.org/10.21956/openreseurope.16648.r30784

© 2023 Stüber K. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Karin Stüber
[1] University of Zurich, Zurich, Switzerland
[2] University of Zurich, Zurich, Switzerland

Geminate consonants in lexemes attested in Celtic languages have been used in the past as cues for non-Indo-European origins of the words in question and in consequence for their status as borrowings from substrate languages. These alleged loanwords, in turn, were sometimes used for conclusions on phonological properties of those substrate languages.

The article discusses the historical phonology of geminate consonants in the Celtic languages. In a systematic fashion the evidence for etymologically justified geminates, arising by assimilation of consonant clusters, is presented first for Proto-Celtic and then for the individual Celtic languages. This is followed by a chapter on non-etymological gemination, e. g. in onomastics or with expressive force.

The article offers a very useful collection of all the relevant material and a thorough discussion thereof. Finally, the question whether geminates can be a marker of substrate loan words is addressed.

The author's conclusion is that, broadly speaking, geminate resonants (like ll, rr) are usually phonologically justified in inherited lexemes, whereas words containing geminate stops tend to be etymologically obscure. In the case of nouns, there is good cause to understand them as loanwords. On the other hand, the author identifies a number of adjectives describing physical properties containing geminate stops. In their case, he argues that those geminates might have arisen for onomatopoetic reasons.

In my view, the article is a valuable contribution to Celtic phonology. The idea of taking word class into account when considering phonological properties, when transferred to other languages, promises interesting results

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Yes

If any, are all the source data and materials underlying the results available?
No source data required

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Historical linguistics, Celtic linguistics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 22 Feb 2023

David Stifter

Thank you very much!

Competing Interests: No competing interests were disclosed.

Reviewer Report 16 February 2023

https://doi.org/10.21956/openreseurope.16648.r30785

© 2023 Kümmel M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Martin Joachim Kümmel
[1] Friedrich-Schiller-Universität Jena, Jena, Germany
[2] Friedrich-Schiller-Universität Jena, Jena, Germany

The paper treats a topic never systematically treated as such, and so it is by definition innovative, even if much of its content takes up earlier research concerning individual points. The question it tries to answer follows from the absence of geminates in the Proto-Indo-European ancestor of Celtic and the existence of geminates in Celtic languages: Geminates must have been introduced by phonological changes somewhere in the (pre)history of Celtic, so where did they come from? The same question must be asked for every other branch of Indo-European that developed geminates, and it has not been investigated systematically for all of them, and so this article could serve as a model for other branches.

The main source for geminates was assimilation of different consonant clusters, and so the article discusses a large number of such clusters and their development in Celtic; it also includes a discussion of the outcome of **i* preceding assimilated **NT*-clusters in Goidelic, arguing for a variant **in* beside **en, *eni*, **ande* for „preverbal *in*“. It is shown that gemination of resonants was probably the first and arose by assimilation of some clusters already in Proto-Celtic (PC). In addition, some clusters with stops also appear to have been assimilated everywhere and thus possibly in PC, namely **ndn, *nnd* and even **dk*. Many more changes cannot be ascribed to Proto-Celtic but belong to the parallel histories of Celtic subbranches and show similar tendencies after their separation. In addition to the results of various assimilations, cases of „onomastic“ and „expressive“ gemination and the impact of possible borrowings are discussed. It is shown that different factors "conspired“ for the rise of geminates, including support from unknown substrate languages which probably had geminates.

Previous research is adequately discussed, and the argumentation is generally sound; unclear cases are named as such, and dubious proposals are rightly rejected. Of course, as always in such comprehensive overviews, not all cases can be discussed in all details but there are no serious shortcomings.

Remarks on single points are given under the heading numbers of the original paper.

1. It remains questionable if "strongly affective“ words with geminates can really be "projected back to the protolanguage“. The problem, as already shown by the Celtic examples, is that such words do not follow the sound laws, and hence it is impossible to demonstrate that they must be inherited. Instead, the correspondence between Gothic *atta*, Gaulish *atta/attū*, Hittite *atta-* etc. does not presuppose a common ancestor **attV-*, as little as its similarity to Common Turkic *ata* presupposes a common ancestor. By the way, the Hittite word does not necessarily have an

original geminate **tt* (distinct from single **t*), since *tt* would be the outcome of **t* anyway, and the Germanic word, being an *n*-stem, could easily owe its geminate to Kluge's Law. In the light of „onomastic" gemination (section 9), it seems equally possible that a potential PIE word still had **t* which underwent gemination only *einzelsprachlich* (but not, e.g., in Slavonic **at-ika-*).

2.1 What exactly is the process of "analogical introduction" of consonantal gemination after the model of one suffix? Was this a proportional analogy **-lo- : *-llo- = *-ko- : X*?

2.2 For the examples with **T-s-m°* in suffixes, it might be made clearer why the individual cases must be reconstructed with **s* and not simply with **T-m°*, since the **s* is never directly visible nor supported by parallels in other branches.

2.3 (1-2) OE *pint-*, MLG *klint* need not presuppose **nd-* if my proposal of some special developments in Germanic NT-clusters (Kümmel 2017) is correct which would allow to explain Germanic **nt* also from **nd^{h}*; alternatively, an explanation through Kluge's Law would be possible, viz. **glend^{h}nó-* > Germ. **klentta-* > **klinta-*.

2.4 (1-4) Original **dk* was apparently not assimilated but just simplified to **k* in decade formations like PC **wikantī* '20', **trikont-* '30' < (q)PIE **(h$_1$)wi-dkm̥tih$_1$, *tri-dkómt-*, where Indo-Iranian has **nkˆ* (Sanskrit *viṃśatí-*, *triṃśát-*; Ossetic Digoron *insæj, ærtin*) while other IE languages (with slightly different formations) show loss of **d* with compensatory lengthening of the vowel (e.g., Latin *uīgintī*, Greek Doric *wīkati*) if they continue such a form or sequence (unlike remade forms like Greek *triákonta*). So it seems that the assimilation in words formed in Celtic follows a new pattern here.

(5) The difference of **dk* > **kk* and **tk* > **kt, *dg* > **gd* is striking but it might be useful to address the question of a conditioning for the metathesis in **arxto-, *gdonjo-* (and **gdū* 'earth'). Maybe just **dk* > **tk*, followed by **tk* > *kk* /V_V (but not **dg* > **gg*?) and **tk, *dg* > **kt, *gd* /#,C_?

3.4 The example of **tisr-* with some segmental reflex of **s* in Celtic and the other examples with complete loss do not show a uniform treatment, or do they? This might be made clearer.

3.7 (1) There might be an alternative etymology for *gwell* etc., from a **wel-*root, cf. Germanic **wel-* 'good' etc. (so Matasović 2009 s.v.).

3.9 For the intermediate stage **t^{s}* to yield later *ss*, it must have been either **ts* or **tt^{s}*, so one of these should be written instead; otherwise it would have been monomoraic and become bimoraic again.

3.9.5 For the development of **st* > *ss* an intermediate stage [ts] is unnecesssary, so its asumption depends on (Gaulish) attestations of *ðð/θ* etc. Are there examples other than *tuððos/tuθos* for this? In any case, such a development in Gaulish does not necessarily imply that Insular Celtic had the same metathesis.

3.10 Fn. 2 If Proto-Celtic really had (acquired and) preserved **pp* in contrast to simple **p*, wouldn't it be useful to include this possible (but very rare) phoneme in the phoneme tables?

4.1 Would it be possible to set up a relative chronology by which **zd* > **dd* > British (and Gaulish)

**tt* was earlier than the rise of younger **dd* (which was not devoiced)? Devoicing of geminates is a rather natural change (and probably part of Kluge's Law in Germanic), and the examples of a voiceless outcome under 6.2 seem to show the same, and might be older, as discussed there.

6.1 What could be the reason why original **d^hsd^h* > **dzd^h* does not develop like original **tst* > Celtic **ts*? With only one example, some kind of analogical influence is difficult to exclude.

6.2 Here the question arises what kind of "morpheme boundaries" are meant: formations with prefixes are originally compounds (at least they are in Sanskrit where external sandhi is applied), so it is actually boundaries of phonological words, not (word-internal) morphemes. Could this make a difference for the outcome of clusters, and would it be possible to assign conflicting results to full vs. incomplete univerbation?

11.3.1 (13) The vocalism of Slav. **munaga-* has been argued to be regular from a zero grade (Matasović 2004) - still not directly comparable to Celtic **men-* and Germanic **man-* but not really "irregular".

14. There is hardly any real difference between **k^wk^w* and **kk^w*, as it is quite improbable that it was actually pronounced without lip rounding in the first half of the long [k]-gesture and with rounding in the second, and it is even more improbable that such a [kk^w] could stand in contrast to [k^wk^w].

14.1 The recordings of Trebitsch have been published in Lechleitner & Remmer (2004), so this edition should be cited here.

14.2 Tables 7 and 8: Why is labial-velar /g^w/ present in table 7 but gone in table 8? Is there not still a unlenited : lenited correlation with /w/ in the attested languages, more or less like /b/ : /β/? Probably, the glide should be classified as labio-velar rather than „bilabial", and one may also consider adding another labio-velar phoneme /h^w/ (cf. Primitive Irish /x^w/), especially since this possibly does not always go back to a cluster **hw* (at least according to Jørgensen 2012).

The relative length of voiceless obstruents may be argued to have remained relevant until the split of vowels (reflected) in (Modern) Welsh into a long/tense in originally open syllables (preceding simple voiced consonants and some fricative-stop clusters) and short/lax variant in originally closed syllables (preceding clusters, original geminates, voiceless consonants and *m, ng*), so it must have played a role in phonology.

References

1. Jørgensen AR: Palatalization of *sk in British Celtic. *In Nielsen Whitehead, B; Olander, T; Olsen, BA; and Rasmussen, JE (eds.), The Sound of Indo-European. Phonetics, Phonemics and Morphophonemics. Selected papers from the conference held in Copenhagen, 16-19 April 2009 (Copenhagen Studies in Indo-European 4). Copenhagen: Museum Tusculanum*. 2012. 209-222
2. Kümmel MJ: Das dünkt mich dunkel: Germanische etymologische Probleme. *In Hansen, BSS; Olander, T; and Olsen, BA (eds.), Etymology and the European lexicon. Proceedings of the 14th Fachtagung der Indogermanischen Gesellschaft, 17-22 September 2012, Copenhagen. Wiesbaden: Reichert*. 2012. 219-233
3. Lechleitner G, Remmer U (eds.): Celtic Recordings - Ireland, Wales, Brittany, Isle of Man and Scotland (1907-09). (Sound Documents from the Phonogrammarchiv of the Austrian Academy of Sciences. The Complete Historical Collections 1899 - 1950. Series 5/2: The Collections of Rudolf

Trebitsch. OEAW PHA CD 14). *Wien: Verlag der Österreichischen Akademie der Wissenschaften*. 2004.
4. Matasović R: The Proto-Indo-European syllabic resonants in Balto-Slavic. *Indogermanische Forschungen*. 2004. 337-354
5. Matasović R: Etymological dictionary of Proto-Celtic (Leiden Indo-European Etymological Dictionary Series 9). *Leiden/Boston: Brill*. 2009.

Is the work original in terms of material and argument?
Yes

Does it sufficiently engage with relevant methodologies and secondary literature on the topic?
Yes

Is the work clearly and cogently presented?
Yes

Is the argument persuasive and supported by evidence?
Yes

If any, are all the source data and materials underlying the results available?
No source data required

Does the research article contribute to the cultural, historical, social understanding of the field?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Indo-European linguistics; historical phonology; phonological typology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 16 Feb 2023
David Stifter

Thank you very much for this helpful and insightful report, David

Competing Interests: No competing interests were disclosed.

Comments on this article

Version 1

Reader Comment 03 Feb 2023

Matthias Donners, Friedrich-Schiller-University, Jena, Germany

Dear David, there are two things I noticed.

In Section 2.1. (17), there is a small typo. The translation of Sucellus says "good-stiker", instead of "striker".

Secondly, in Section 11.3.2. (11), you compared Gaul. messamobi with OIr. messam. Although the etymology and form of the Gaulish word with -ss- is probably correct, it is not actually attested on the Lezoux plate inscription (L-66), 4th line, but rather mesamobi with single -s-, independent of the differences in various lectures of this inscription (cf. RIG 2.2: 176-178). In light of gussou in Line 7 of the same inscription, I think mesamobi with a single -s- is to be taken seriously, although non-spelling of geminates is attested (cf. e.g. E. Dupraz, ZcP 62. (2015): fn. 20). As I said, the etymology of the form containing the suffix *~-samo- is probably valid, and the Old Irish form speaks unambiguously for the geminate.

Best regards, Matthias

Competing Interests: No competing interests were disclosed.

Author Response 02 Feb 2023

David Stifter

Hardly published, I have changed my mind about section 3.7 (**rl* > **ll*). A third, good example is OW *guolleuin* (Juvencus glosses), MW *gollewin* "west" < **u̯or-lugu-īno-* (*uel sim.*), i.e. the preverb **u̯or-* "on, upon" + a derivative of the word seen in W *lleu* "light, brightness". In ModW *gorllewin*, the preverb *gor-* < **u̯or-* has been re-introduced.

Competing Interests: No competing interests were disclosed.