

# An Outline of the Standard Written Form (SWF) for Cornish

Albert Bock  
Benjamin Bruch

## 1. The Role of the Standard Written Form

The Standard Written Form for Cornish, as agreed on 15 December 2007 by the Ad-Hoc group, comprising users of Cornish nominated by the language groups, represents a common ground for users of all existing orthographies and speakers of all varieties of Revived Cornish. The SWF is not meant to replace other spelling systems, but rather to provide public bodies and the educational system with a universally acceptable, inclusive, and neutral orthography.

The SWF incorporates features drawn from a number of different Cornish orthographies, including Unified Cornish, Kernewek Kemmyn, Modern Cornish, Unified Cornish Revised, Kernowak Standard, and Kernewek Dasunys. The SWF follows a combination of phonetic and phonemic principles in spelling rather than a strict morpho-phonemic approach. The two principal effects are a closer relationship between sound and spelling and a stronger resemblance to orthographic systems designed by native speakers of their respective languages, like the orthography of Modern Welsh.

The SWF is much more inclusive of variant forms than any previous Cornish orthography. Care has been taken to construct a system which speakers of all varieties of Revived Cornish can learn to use quickly and easily, in a manner which suits their linguistic and aesthetic preferences. Even those users who do not wish to use the SWF actively should find it relatively easy to read.

This document does not represent a complete specification of the SWF. It is rather an explanation of the agreement reached by the AHG. A few details are still open and will be settled after discussions with the linguistic advisors, at which point a full specification will be distributed by the Partnership. This will include more detailed explanations of specific issues, rules for converting from other orthographies to the SWF, a basic wordlist, and paradigms of prepositions and common verbs.

## 2. Principles; linguistic, political and practical considerations

The Standard Written Form largely follows the principles for a common orthography first set out in the draft for Kernewek Dasunys. These are:

- a) **Inclusivity** – Users of all varieties of Revived Cornish should be able to write as they speak.
- b) **Accessibility** – The SWF should be easy for speakers, learners, and teachers to learn and use.
- c) **Accuracy** – The SWF should reflect the pronunciation of both traditional and Revived Cornish.
- d) **Authenticity** – The SWF should not look “foreign” to present-day speakers of Cornish.

A fifth principle has been added to these, based on the statement of the Cornish Language Commission:

- e) **Continuity** – The SWF should, as much as possible, involve the smallest possible number of changes for the largest possible number of speakers, i. e. minimise the distance to travel for users of all orthographies.

### 3. Main Forms, Variant Forms, and Side Forms

In order to be inclusive of all varieties of Revived Cornish, the SWF will allow a small number of variants in spelling. To reduce the burden on teachers and learners, the number of permitted variants will be kept to a minimum. Most of the variation between Middle and Late Cornish forms is bridged using *umbrella graphs* (on which see below). The remaining variant spellings can be grouped according to their status as either **Variant Forms** or **Main Forms** and **Side Forms**.

#### 3. 1. Variant Forms

Variant Forms are equal in status to one another, and can be described as equivalent Main Forms. They are unambiguous in that there is a clear correspondence between pairs of variant graphs. The status of Variant Form generally is accorded to variant graphs which indicate a historical sound change: a difference in pronunciation between the varieties of spoken Cornish. In general, these correspond to sound changes which occurred between Middle and Late Cornish.

#### Examples of Variant Forms

<i>MC</i>		<i>LC</i>	<i>examples</i>		
<mm>	~	<bm>	<b>tamm</b>	~	<b>tabm</b> 'piece'
<nn>	~	<dn>	<b>penn</b>	~	<b>pedn</b> 'head'
<-i>	~	<-ei>	<b>chi</b>	~	<b>chei</b> 'house'
<ew>	~	<ow>	<b>klewes</b>	~	<b>klowes</b> 'hear'
<s>	~	<j>	<b>kerensa</b>	~	<b>kerenja</b> 'love'

#### 3. 2. Main and Side Forms

Side Forms are not wholly equal in status to Main Forms. They generally reflect speakers' aesthetic preferences about spelling rather than an actual difference in pronunciation. In a few cases, forms which are historically attested but less commonly used in Revived Cornish have been included as Side Forms rather than Variant Forms, since they will not be presented to learners in introductory material.

Side Forms are officially recognised, and may be freely used in writing and publishing, including in officially funded publications. They will be considered correct by Cornish spell-checking software, and students may likewise use them in examinations and written work. Although the Main Forms will be given preference in textbooks and official documents, literary texts which use Side Forms are also admissible for use in schools. In order to respect the link between contemporary and historical forms of the Cornish language, the use of such literary texts in teaching advanced students is indeed encouraged. It is also likely that many place-names will be written in forms that reflect Side Form spellings using <c>, <q>, and <wh>.

The principal Side Forms in Cornish spelling involve the sounds written <k> and <hw> in the Main Forms. These may also be written more in accordance with attested historical practice, using the graphs <c>, <qu>, <x>, and <wh>. In traditional Cornish, the sounds /k/, /kw/, /ks/ and /m/ (or /hw/) were written according to the same rules used in English, French and many other Western European languages. According to these rules, the sound /k/ is written <c> before the vowels <a, o, u> and the consonants <l, r, k>. The sequence /kw/ is written <qu>, and the sequence /ks/ (or /qz/) is written <x>. Elsewhere, the sound /k/ is written <k>. In this document, Side Forms are given in curly brackets:

<b>kans</b>	<b>{cans}</b>	'hundred'	<b>kwarter</b>	<b>{quarter}</b>	'quarter'
<b>krev</b>	<b>{crev}</b>	'strong'	<b>taksi</b>	<b>{taxi}</b>	'taxi'
<b>hweg</b>	<b>{wheg}</b>	'sweet'			

#### 4. Inclusivity: Accounting for Middle and Late Cornish Forms

Spoken Revived Cornish has two main variants, one based on Middle and one on Late Cornish. The SWF recognises Revived Late Cornish as a variant of equal standing. Even though the orthography on the whole leans much more towards Middle Cornish, Late Cornish variants are accounted for:

- **orthographically**, through the use of Variant Forms and umbrella graphs;
- **morphologically** and **syntactically**, through the recognition of Late Cornish forms and syntactic structures as being equally correct as their Middle Cornish counterparts; and
- **lexically**, through the inclusion of Late Cornish lexical items in official dictionaries.

Late Cornish forms will be spelt more or less phonemically in closed class words like auxiliary verbs and prepositions, where paradigms differ considerably between MC and LC in any case. In open class words, umbrella spellings (see below) will be given preference. Variant spellings like ⟨bm⟩, ⟨dn⟩, and ⟨-ei⟩ will be used in both classes.

In order to keep the written representations of the dialects of Revived Cornish close enough to one another to ensure mutual comprehension, a number of variant pronunciations are bridged by **umbrella graphs**. These umbrella graphs are typically based on MC spellings because with a few exceptions, phonemic distinctions were lost rather than created during the evolution of LC from MC.

##### Umbrella graphs in the SWF

<i>Umbrella graph</i>	<i>MC pronunciation</i>	<i>LC pronunciation</i>	<i>examples</i>
⟨u⟩	[y(:)]	[i(:)]	<b>rudh</b> [ry:ð], [ri:ð]
⟨eu⟩	[œ(:)]	[e(:)]	<b>keus</b> [kœ:z], [ke:z]
⟨y⟩ when stressed	[ɪ(:)], [i(:)]	[e(:)]	<b>pryv</b> [pri:v], [pre:v]

#### 5. Features of the SWF as agreed by the AHG

##### 5. 1. Vowel Length

The AHG agreed that vowel length in monosyllables would be tied to the nature of the following consonant or consonants. This system is very close to that of Welsh and to those of other Cornish orthographies which indicate vowel length.

Vowels in stressed monosyllables are long:

- In final position: **ro** ‘gift’, **da** ‘good’, **bre** ‘hill’
- Before a single voiced consonant: **mab** ‘son’, **hir** ‘long’, **mil** ‘thousand’
- Before the clusters ⟨st, sk⟩: **lost** ‘tail’, **Pask** ‘Easter’
- Before a fricative: **hav** ‘summer’, **nos** ‘night’, **kath** {**cath**} ‘cat’

Vowels in stressed monosyllables are short:

- Before a voiceless stop: **hat** ‘hat’, **top** ‘top’
- Before a consonant written double ⟨mm, nn, ll, rr, ff, ss⟩: **penn** ~ **pedn** ‘head’, **pell** ‘far, long’
- Before a consonant cluster other than ⟨st, sk⟩: **park** ‘field’, **kans** {**cans**} ‘hundred’

There are a few exceptions to these rules, generally involving loanwords.

## 5. 2. Vocalic Alternation

The SWF acknowledges vocalic alternation, a sound change whereby stressed ⟨y⟩ in monosyllables and unstressed ⟨y⟩ in final syllables often becomes ⟨e⟩ when a suffix is added. This change affects words in ⟨yw⟩ as well as words in ⟨y⟩. Common examples include the following:

### Stressed ⟨y⟩ in root becomes ⟨e⟩:

<b>*blydh-</b>	-	<b>bledhen</b>	‘year’
<b>bydh</b>	‘will be (3sg.)’	<b>bedhav</b>	‘will be (1sg.)’
<b>bydh</b>	‘be! (2sg.)’	<b>bedhewgh</b>	‘be! (2pl.)’
<b>bys</b>	‘world’	<b>besyow</b>	‘worlds’
<b>dydh</b>	‘day’	<b>dedhyow</b>	‘days’
<b>gwydh</b>	‘trees’	<b>gwedhen</b>	‘tree’
<b>pryv</b>	‘worm’	<b>preves</b>	‘worms’
<b>prys</b>	‘time’	<b>presyow</b>	‘times’

This stressed ⟨y⟩ is pronounced [ɪ:], [i:] by speakers of MC and [e:] by speakers of LC.

### Stressed ⟨yw⟩ becomes ⟨ew⟩:

<b>byw</b>	‘alive’	<b>bewnans</b>	‘life’
		<b>bewa</b>	‘live’
<b>lyw</b>	‘rudder’	<b>lewya</b>	‘steer’
<b>skyw</b>	‘shelter (n.)’	<b>skewya</b>	‘shelter (v.)’

This stressed ⟨yw⟩ is pronounced [ɪʊ] by speakers of MC and [ɛʊ] by speakers of LC. Some polysyllabic words with ⟨ew⟩ in MC will use the Variant Form ⟨ow⟩ in LC. LC speakers would therefore write **bownans** rather than **bewnans**.

### Unstressed final ⟨y⟩ in root becomes ⟨e⟩:

<b>enys</b>	‘island’	<b>enesow</b>	‘islands’
		<b>enesek</b>	‘isolated’
<b>benyn</b>	‘woman’	<b>benenes</b>	‘women’
<b>melyn</b>	‘yellow’	<b>melender</b>	‘yellowness’
<b>menydh</b>	‘mountain’	<b>menedhyow</b>	‘mountains’
<b>gorthyp</b>	‘answer’	<b>gorthebi</b>	‘answer’
		<b>gorthebow</b>	‘answers’

Stems in SWF ⟨i⟩ are not affected by Vocalic Alternation, according to the recommended pronunciation of KK, UC, and RLC. Thus, the SWF will write **tir**, pl. **tiryow** ‘land’. Only UCR has Vocalic Alternation in some of these stems, e.g. **myr** (SWF **mir**) ~ **meras**, ‘look’. Here, the SWF follows majority usage and does not show Vocalic Alternation in words like **mires**, **hwilas**, etc. Consideration can be given to including {**meres**, **whelas**} as Side Forms to reflect UCR usage and occasional LC attestations with [ɛ].

## 5. 3. Vowels: ⟨oo⟩ and ⟨o⟩

The SWF spells the sound(s) represented by KK ⟨oe⟩ as ⟨oo⟩ when the vowel in question is long and ⟨o⟩ when it is short, as shown below:

*Short*

<b>tomm ~ tobm</b>	‘hot’
<b>gallos</b>	‘be able to’
<b>kavos {cavos}</b>	‘find’

*Long*

<b>koos {coos}</b>	‘wood’
<b>loor</b>	‘moon’
<b>goon</b>	‘moor’

The vowel ⟨oo⟩ will be used in cases where KK has a long ⟨oe⟩, with the exception of a small number of words which seem to have undergone a different phonological development from the **loor**, **goon** words, based on their spellings in LC. Words like the following will therefore be spelled with ⟨o⟩: **kon {con}** ‘dinner’, **tron** ‘nose’, **on** ‘lamb’, **gor** ‘knows’, **hwor** ‘sister’, **kor {cor}** ‘wax’, **noth** ‘naked’.

**5. 4. Diphthongs**

In general, ⟨iw⟩ is used where both RMC and RLC have [iʊ], [ɪʊ], while ⟨yw⟩ is used in cases where RMC [ɪʊ] corresponds to RLC [ɛʊ]. Examples include:

⟨iw⟩ for MC and LC [iʊ], [ɪʊ]:

<b>diw</b>	‘two (f.)’
<b>diwvron</b>	‘breasts’
<b>liw</b>	‘colour’
<b>liwya</b>	‘dye’
<b>niwl</b>	‘fog’
<b>piw</b>	‘who’

⟨yw⟩ for MC [ɪʊ], LC [ɛʊ]:

<b>byw</b>	‘alive’
<b>lyw</b>	‘rudder’
<b>yw</b>	‘is’

Because of Vocalic Alternation, ⟨yw⟩ can only occur in stressed monosyllables. In polysyllabic words, ⟨yw⟩ will become ⟨ew⟩, just as ⟨y⟩ often becomes ⟨e⟩:

<b>bewnans</b>	‘life’
<b>lewya</b>	‘steer (v.)’

Words that have [ɛʊ] in both MC and LC are spelled ⟨ew⟩:

<b>blew</b>	‘hair’
<b>dew</b>	‘two (m.)’
<b>ewn</b>	‘correct’
<b>lew</b>	‘lion’
<b>rew</b>	‘ice’
<b>tew</b>	‘fat’
<b>tewl</b>	‘dark’

Note that in LC, ⟨ew⟩ sometimes becomes ⟨ow⟩ in stressed polysyllables, producing Variant Forms like:

<b>bewnans ~ bownans</b>	‘life’
<b>klewes ~ klowes</b>	‘hear’
<b>{clewes ~ clowes}</b>	

<b>tewlel ~ towlel</b>	‘throw’
<b>dewdhek ~ dowdhek</b>	‘twelve’

In addition, the SWF recognises the digraph ⟨uw⟩. At present, there are two different suggestions as to what sound or sounds ⟨uw⟩ represents and what words in the SWF should be spelled with ⟨uw⟩. There is unanimity as far as the words **Duw** ‘God’, **duwon** ‘misery’, and **ruw** ‘king’ and their compounds are concerned. Other possible cases have been referred to the linguistic advisors and arbiter.

### 5. 5. Sonorants: ⟨mm ~ bm, nn ~ dn, ll, rr⟩ and ⟨m, n, l, r⟩

The SWF indicates geminate or long liquids ⟨ll, rr⟩ in writing where they are actually pronounced long in conservative Middle Cornish or as fortis in later forms of the language. The positions where this happens are those where pre-occlusion of [m:] and [n:] occurs, which is to say in stressed syllables only. Therefore ⟨mm/bm, nn/dn, ll, rr⟩ will be used in monosyllabic words and their compounds as well as in the stressed syllables of polysyllabic words.

<b>tamm ~ tabm</b>	‘piece’
<b>bronn ~ brodn</b>	‘breast’
<b>pell</b>	‘far, long’
<b>berr</b>	‘short’
<b>lemmyn ~ lebmyn</b>	‘now’
<b>ranna ~ radna</b>	‘divide’
<b>challa</b>	‘jaw’
<b>gorra</b>	‘put’

In unstressed and pre-tonic syllables, and also in consonant clusters, the SWF writes single ⟨l, r, m, n⟩, except in compounds where the initial element retains its double (or pre-occluded) consonant because of secondary stress. This approach has several benefits: it ties the orthography more closely to the spoken language, and it indicates at first glance where pre-occlusion can and cannot occur. This can be illustrated by the following examples, where pre-occlusion is lacking in LC. In such cases, it should be noted, the MC and LC spellings are often identical, since there is no need for a Variant Form in ⟨bm, dn⟩:

<b>pluven</b>	‘feather’	<b>not</b>	*pluedn
<b>gwedren</b>	‘glass’	<b>not</b>	*gwedredn
<b>mynsa ~ menja</b>	‘would like’	<b>not</b>	*mednja
<b>tomder</b>	‘warmth’	<b>not</b>	*tobmder

The SWF spells **pluven** with a single ⟨-n⟩ because phonetically the suffix has a single [n]. Double consonants are not necessary to indicate vowel length in the unstressed final syllables of words like **pluven** or **gwedren**, since all vowels in unstressed position are short by default.

<b>pluvennow ~ pluednow</b>	‘feathers’
-----------------------------	------------

The plural form **pluvennow ~ pluednow** is written with ⟨nn⟩ or ⟨dn⟩ to emphasise that the [n:] is historically long/fortis and that it is pre-occluded in Late Cornish. This double consonant also clearly indicates that the stressed vowel which precedes it is short.

In compound words, pre-occlusion occurs in syllables which bear secondary stress, indicating that these syllables were longer than unstressed ones and could contain geminate sonorants as well as long vowels.

Therefore, the SWF writes (ll, rr, mm, nn) rather than (l, r, m, n) in this position:

<b>kammneves</b> ~ <b>kabmdhavas</b>	‘rainbow’
{ <b>cammneves</b> ~ <b>cabmdhavas</b> }	
<b>pellgowser</b>	‘telephone’

There is a small class of MC words which lack pre-occluded Variant Forms. Some of these seem to have been borrowed from English after pre-occlusion occurred, e.g. **jynn** ‘engine’; others like **gonn** ‘knows’, were replaced by alternatives in LC and thus are only attested in their non-pre-occluded MC variants.

#### 5. 6. Voiceless stops: (-pp-, -tt-, -kk-) and (-p, -t, -k)

The SWF does not write (-pp -tt -kk) word-finally because this would suggest a physically impossible pronunciation, such as [hat:] for **hatt**. It is impossible to pronounce geminates not followed by a vowel unless they are ejectives (a class of sounds which does not exist in Cornish). Therefore, in the SWF, vowels are always short before (p t k) apart from a few exceptional, easily recognisable loanwords like **strok**, **stret**. Word-internally, (-pp- -tt- -kk-) are retained in words like **klappya** {**clappya**} and plurals like **hattow**, **hattys**, where the stop may be pronounced as a geminate. In short, this represents a move towards a more phonetic, ‘spell-as-you-speak’ approach.

#### 5. 7. Fricatives: (gh) and (h)

MC /x/ in words like **sygh** ‘dry’ and **margh** ‘horse’ is written (gh) in the SWF.

Between vowels, or between (r, l) and a vowel, the spelling (h) will be used in both MC and LC variants:

<b>sygh</b>	‘dry’	<b>sehes</b>	‘thirst’
<b>flogh</b>	‘child’	<b>flehes</b>	‘children’
<b>margh</b>	‘horse’	<b>marhek</b>	‘knight’

One exception is the word ‘small,’ which is **byhan** in MC but **bian** in LC.

The digraph (gh) thus always represents [x] for speakers of MC and [h] for speakers of LC, whereas (h) represents [h] for both.

#### 5. 8. Fricatives: Word-final (-dh), (-v) vs. (-th), (-f)

It has been agreed that word-final fricatives should be spelt in a way which reflects their status of voicedness / voicelessness. Thus, the SWF writes **bodh** ‘will’, **klav** {**clav**} ‘sick’, but **eth** ‘eight’, **dalleth** ‘begin’, **hanaf** (~ **hanath**) ‘cup’, etc. Where evidence from the traditional Cornish corpus is ambiguous, Breton and Welsh cognates will be examined.

#### 5. 9. Sibilants: (c) and (z)

Loanwords from Norman French like **cita** will be spelt with initial and medial (c). This reflects the idea that the sound represented by (c) was probably still an affricate at the time of borrowing, and that it was pronounced as voiceless [s] in both MC and LC. This is necessary to distinguish it from voiced initial and medial /s/ in LC, which is spelt (s). This principle is not extended to final position where the SWF will write (s) in words like **plas**, since writing word-final (c) would likely lead learners to mispronounce it as [k].

Thus, the SWF writes:

<b>cita</b>	‘city’	[ˈsɪːtɑ, ˈsɪtə]
<b>cider</b>	‘cider’	[ˈsɪdɛr, ˈsɔɪdər]
<b>plas</b>	‘place’	[plɑːs]
<b>gras</b>	‘grace, thanks’	[grɑːs]

The possibility of introducing ⟨z⟩ for [z] in certain contexts has been referred to the linguistic advisors and the arbiter for ruling.

### 5. 10. ⟨dhyworth⟩

The preposition ‘from’ is spelt **dhyworth** in the SWF, in order to reflect

- that the soft mutation of the initial consonant (< **a-dhyworth**) became generalised early on and
- that the first syllable was syncopated in LC **dhort**, indicating early reduction of the unstressed vowel.

## 6. Graphs used in the Standard Written Form

The table on the following page lists the graphs used in the SWF, along with an IPA transcription which reflects their pronunciation in all varieties of spoken Revived Cornish. This list is primarily descriptive, not prescriptive, although it excludes a few features like diphthongised long [eː] > [eɪ] and [oː] > [ɔɔ] which are often heard in the speech of learners. This should not be misunderstood as an approach of complete *laissez-faire*. However, the political question of a prescribed standard pronunciation for the two main variants (MC and LC) is not one which the SWF itself should attempt to solve.

For a discussion of Variant and Side Forms, see above.



## Graphs used in the SWF and their pronunciation in spoken revived Cornish

<i>Consonants</i> ~ <i>Variant Forms</i> { <i>Side Forms</i> }	<i>IPA</i>	<i>Vowels</i> ~ <i>Variant Forms</i> { <i>Side Forms</i> }	<i>IPA</i>
b	[b]	a	[a(:) æ(:) ʌ(:)]
d	[d]	e	[e(:) ε(:)]
dh	[ð]	eu	[œ(:) ø(:) e(:) ε(:)]
f	[f v]	i ~ -ei (stressed, final)	[i(:) i] ~ [-əi]
ff	[f f]	y	[i(:) i: e: ε]
g	[g]	o	[ɔ(:) o: ʌ]
gh	[x x: h]	oo	[o: ʊ: u:]
h	[h]	ou	[u(:) ʊ]
hw {wh}	[ʍ hw]	u	[y(:) i(:) ʏ i]
j	[dʒ]		
k {c, q}	[k]	<i>Diphthongs</i>	<i>IPA</i>
l	[l]	aw	[aʊ əʊ]
ll	[l: l]	ew	[εʊ]
m	[m: m]	iw	[iʊ iʊ]
mm ~ bm	[m: bm m]	yw	[iʊ εʊ]
n	[n]	ow	[ɔʊ] ~ [u:]
nn ~ dn	[n: dn n]	uw	[yʊ iʊ]
p	[p]	ay	[aɪ æɪ]
r	[r r ɹ]	ey	[εɪ]
rr	[r: ɹ]	oy	[ɔɪ]
s	[s z]		
ss	[s: s]		
sh	[ʃ]		
t	[t]		
th	[θ]		
tth	[θ: θ]		
v	[v]		
w	[w]		
y	[j]		
z (? - advisors & arbiter)	[z]		