Word-Initial Clusters in Welsh: A Typological Analysis

Tomasz Czerniak

Abstract

This paper applies the typological division between 'TR-only' and 'anything-goes' languages developed by Scheer (1999b, 2004, 2007) to the data from the Bangor dialect of Welsh collected from Fynes-Clinton (1913) pertaining to the well-formedness of word-initial clusters. It appears that the most commonly found clusters are those of rising sonority (TR) and those preceded by [s] (ST, STR). However, there are a number of initial clusters that may surface due to vowel deletion and whose sonority profile fails to satisfy Sonority Sequencing (that is, the second member is equally or less sonorous than the first one). This places Welsh in the 'anything-goes' category, which has theoretical implications: the article proposes that the beginning of the word is *not* translated into structural information in Welsh. In other words, phonological domains do not begin with an empty CV unit.

Keywords: *Welsh, initial clusters, syncope, Government Phonology, typology*

1. Introduction

The aim of this paper is to investigate the left edge of the word in Welsh from the typological perspective offered by Government Phonology (GP) in general and the Lateral Theory of Phonology (CVCV) in particular. A typological demarcation line will be drawn between languages whose consonants must contract a relation within initial clusters and those languages where it is optional. In order to place Welsh on one side of the fence, the phonotactic constraints on the left periphery of the word have to be analysed. Anderson's (2011) melody-based analysis of initial clusters in Irish reached the conclusion that Irish and English belong to two opposite typological groups. The present analysis of Welsh will shed more light on the affiliation of the Celtic languages in general.

Section 2 investigates the well-formedness restrictions on word-initial sequences with a particular focus on attested clusters in 2.1., on clusters resulting from initial vowel deletion in 2.2. and on the strength of the initial consonant in 2.3. Section 3 introduces the theoretical assumptions developed in CVCV concerning the typology in question. Sections 3.1. and 3.2. deal with the structural configuration of clusters breaking the Sonority Sequencing Principle, namely falling sonority and s-initial clusters respectively. Section 4 examines the application of the mechanisms described in 3 to the data from Welsh. Section 5 summarises the discussion and suggests the direction for further research.

2. The phonotactics of the left edge of the word in Welsh

The ensuing sections aim at providing an accurate account of what types of clusters are allowed at the left periphery of the word in Welsh. Data are specific to the Bangor variety, which is a representative of the northern dialect group, and are taken from Fynes-Clinton (1913 [1995]).

2.1. Admissible clusters

One of the facts about the phonology of Welsh is that in the initial position clusters of rising sonority prevail. Additionally, a single stop or a stop-liquid cluster may be preceded by the sibilant [s] (cf. Awbery 1984: 86-87 and 2010: 369-370). For convenience, we will refer to those types of clusters as TR and STR respectively, where T stands for an obstruent, R for a sonorant and S for the alveolar sibilant. Let us consider the following examples:

(1)	a.	[ˈpjodan]	*pioden*	‘magpie’
		[plant]	*plant*	‘children’
		[pnaun]	*prynhawn*	‘afternoon’
		[ˈprədlon]	*prydlawn*	‘punctual’
		[tlaud]	*tlawd*	‘poor’
		[treː]	*trêf*	‘town’
		[kloː]	*clo*	‘lock’
		[knaɨ]	*cnau*	‘hazel-nuts’
		[kroːg]	*crog*	‘cross’
		[kwəir]	*cyweir*	‘a thrashing’
		[blaːs]	*blâs*	‘taste’
		[ˈbnavɨd]	*bynafyd*	‘to hurt’
		[braud]	*brawd*	‘brother’
		[bwiː]	*bwi / buoy*	‘dolphin’
		[djaul]	*diawl*	‘devil’[1]
		[dleːd]	*dlêd*	‘debt’
		[druːg]	*drwg*	‘bad’
		[ˈgjalam]	*gwialen*	‘rod’
		[gliːn]	*glin*	‘knee’
		[gnəɨd]	*gwneud*	‘to make’
		[griʃa]	*grisiau*	‘stairs’
		[gweld]	*gweld*	‘to see’
	b.	[spaːr]	*sbâr*	‘spare’
		[spjanɨ]	*chwibanu*	‘to whistle’[2]
		[sploit]	*sbloit*	‘exploit’
		[spranɨ]	*ysbranu*	‘to refuse (of food)’
		[skavaɬ]	*scafell*	‘eye-brow’
		[sklatar]	*sglater*	‘slater’
		[skravaɬ]	*ysgrafell*	‘curry-comb’
		[skwaːr]	*ysgwâr*	‘square’

[1] The cluster [dj] is often turned into an affricate [dʒ] (Awbery 2010: 369).

[2] This is the only lexical item that begins with [spj] and is an alternative pronunciation for [χwibjanɨ]. Similarly, *sbloit* and *sbladdrio* [splaðrjo] ‘blather’ are the only examples of [spl].

[steːn]	*ystên*	'milking-pail'
[strɨːd]	*ystrŷd*	'street'

We can see that initial clusters may incorporate stops, which are followed by glides, liquids or nasals. However, not all combinations are attested. For instance, [pj] is an attested cluster, whereas [bj] is not. Conversely, [gj] can be found in the initial position in the absence of its voiceless counterpart [kj]. In sum, [w] does not combine with [p, t, d], [j] with [t, k, b], [n] with [t] and [d], while [l] and [r] combine with all stops. It is noteworthy that [kn] and [gn] are by far the most frequent stop-nasal clusters in the initial position, while [tn] and [dn] are unattested. Clearly, the place of articulation plays a role in the well-formedness of clusters and the coronal stop-nasal sequence gap may be viewed as systematic rather than accidental.

As for the STR clusters, they are much more restricted: ST+w clusters are limited to a few lexemes starting with [skw]; [spj] can hardly be said to be well-formed since its existence is unattested save one notable exception of an alternative pronunciation of *chwibanu*; the liquid [l] is found only after [sk] and two lexemes with [sp], while [n] cannot follow ST clusters. The only sonorant capable of following all ST clusters is [r].

Interestingly, ST and STR were not perceived as well-formed initial clusters at all times. Initial ST clusters developed a prosthetic [ə], which is spelt *y* in Modern Welsh. However, it is pronounced only when stressed (schwa is regularly stressed in Welsh in the penultimate syllable) and was never counted as a syllable in the poetry of Old and Middle Welsh if the preceding word ended in a vowel. Finally, the definite article in Middle Welsh was pronounced [ə] before consonants and [ər] before vowels and [s] (see Morris-Jones 1913: 26-27 and Willis 2010: 123). Thus, it might be concluded that STR clusters suffered a slight misadventure on their way through history. They were fine in Old Welsh, then Middle Welsh saw them as ill-formed and tried to introduce a prosthetic vowel which did not catch on but whose remnants survived into Modern Welsh in spelling and [ə]-zero alternations.

STR clusters aside, Welsh only allows clusters of rising-sonority in the initial position, which is in accord with the Sonority Sequencing Principle (cf. Harris 1994: 56 and the reference therein), which says that an optimal syllable should consist of a sonority peak and that the further away a segment is from the peak, the less sonorous it should be. Therefore a TRVRT syllable is optimal since V (a vowel) constitutes the sonority peak, it is surrounded by two sonorants and the whole sequence is flanked by two obstruents (e.g. *prank*, *trend*, *friend*, *quilt*). An initial cluster conforming to the Sonority Sequencing Principle will be temporarily referred to as a 'well-formed branching onset', while one whose sonority profile deviates in any way (RR, TT, RT) will be considered ill-formed.

It behoves us to mention other extremely infrequent, yet existent three- and four-member initial consonant clusters in Welsh. They are listed in (2) below:

(2)	a.	[gwlɑːd]	*gwlâd*	'country'
		[gwraig]	*gwraig*	'woman'
		[gwniŋan]	*cwningen*	'rabbit'
	b.	[kljaran]	*cliaren*	'a term of reproach'
		[krjadɨr]	*creadur*	'creature'
	c.	[gwnjadrag]	*gwniadwraig*	'sempstress'

The clusters in (2a) are often simplified through the deletion of [w] in northern dialects (e.g. [ɡlɑːd]), while in *cwningen* it is the whole cluster that often ends up deleted (i.e. [niŋɑn]) (cf. Awbery 2010: 370). Interestingly, Hannahs (2013: 35-36) notes that the [ɡw, kw, χw] sequences could be viewed as velar consonants with secondary labial articulation [ɡ^{w}, k^{w}, χ^{w}]. However, this possibility will not be investigated here due to space limitations.[3] The combination liquid-glide, as in (2b), is quite common in the word-medial position due to the high frequency of suffix-initial [j]. Finally, four-consonant clusters as in (2c) are highly infrequent in Welsh. There are a handful of such instances usually at morpheme boundaries but hardly ever in the initial position.

The Sonority Sequencing Principle is adhered to only by the examples in (1a) and (2b), while those in (1b), (2a) and (2c) seem to violate it. Firstly, [s] is more sonorous (or at least no less sonorous) than any stop while being farther away from the peak in STR clusters, therefore it cannot be regarded as a well-formed branching onset. Secondly, liquids and nasals are less sonorous than glides, yet they are closer to the peak in (2a) and (2c). (2b) conforms to Sonority Sequencing in that segments decrease in sonority with the distance from the peak of the syllable. However, branching onsets in the following sections will be restricted to two positions, which makes the clusters from (2b) ill-formed.

Although the clusters from (1b) are cross-linguistically frequent and merit a typological category of their own (see Kaye 1996), those from (2) will be viewed as problematic yet important for the typological distinction made in this approach.

2.2. Syncope and resulting clusters

Literature on the phonotactic constraints in Welsh is limited to a handful of publications (Awbery 1984, 1986, 2010; Iosad 2012; Hannahs 2013) whose treatment of syncopated forms is rather negligible. For instance, although both Iosad (2012: 104) and Hannahs (2013: 115-119) discuss a prosody-related phenomenon they call 'antepenultimate deletion', their analyses are restricted to the vowel that stands in the absolute initial position, that is #_. No attention is paid to #C_C vowel deletion, which is present in the language and whose results create clusters that are otherwise unattested.

Prior to the analysis of the syncopated clusters a word of comment on the Welsh prosodic system is in order. Firstly, regular stress is fixed on the penultimate syllable (e.g. [ˈpjɔdɑn]). It may be shifted to the final syllable in morphologically complex expressions (e.g. [pɑrˈhɑi] *parhau* 'to continue', [peŋˈɡliːn] *pen-glin* 'knee') or to any other position in borrowings (e.g. [ˈpɑrɑɡrɑf] *paragraff* 'paragraph') (Hannahs 2013: 45ff). Secondly, long vowels are encountered exclusively in stressed syllables and, importantly for North Welsh, the syllable must be final (Awbery 1994: 75). Thirdly, Griffen (1979) observes that secondary stress is assigned to the initial syllable of a four-syllable word (e.g. [ˌɬefeˈθəirjɑ] *llyffetheiriau*

[3] The problem of [ɡ^{w}, k^{w}, χ^{w}] deserves a separate analysis. It will be omitted here since it contributes little to the present typological analysis: if they are single consonants with double articulation, the analysis ought to focus on sub-segmental structures. If, on the other hand, they are clusters of rising sonority, they should be considered together with other TR clusters (see section 3.1.).

'shackles'). However, the antepenultimate syllable receives not tertiary but quaternary stress regardless of whether there are four or just three syllables in the word. Thus, the antepenultimate syllable falls prey to deletion on account of its weak stress level. Interestingly, this takes place only if the antepenultimate syllable is initial at the same time ([ɬefeˈθəirjɑ] >*[ɬefˈθəirjɑ]). The phenomenon of antepenultimate deletion or syncope is illustrated with the examples below:

(3)	a.	['əskaɬ]	/	['skaɬan]	*ysgall* / *ysgallen*	'pl. / thistle'
		['əskau]	/	['skawan]	*ysgaw* / *ysgawe*	'pl. / elder-tree'
		['əskuɨð]	/	['skuɨða]	*ysgwydd* / *ysgwyddau*	'shoulder / pl.'
		['**a**vol]	/	['vola]	*afol* / *afolau*	'apple / pl.'
		['**e**draχ]	/	['draχa]	*edrych* /*edrycha*	'to look / fut.'
	b.	['k**a**lon]	/	['klondid]	*calon* / *calondid*	'heart / cheering'
		['k**a**rjad]	/	['krjada]	*cariad* / *cariadau*	'love / pl.'
		['f**e**nast]	/	['fnestri]	*ffenestr* / *ffenestri*	'window / pl.'
		[ɬefeˈθəirja]	/	['ɬfeθar]	*llyffetheiriau* / *llyffethair*	'pl. / shackle'
		['s**u**ppar]	/	['spera]	*swper* / *swperau*	'supper / pl.'
	c.	[dəˈveθa]	~	['dveθa]	*difetha*	'to waste'
		[kərˈbinjon]	~	[kriˈbinjon]	*cibinion*	'rankings'
		[kəˈnevin]	~	['knevin]	*cynnefin*	'accustomed'
		[pəˈnelin]	~	['pnelin]	*penelin*	'elbow'
		[ɬəˈfanta]	~	['ɬfanta]	*llyffanta*	'to wander about'
	d.	[amgan]	/	[mgenaχ]	*amgen* / *amgenach*	'alternative / different'
		[əm]+[kɨð]	/	[mgiðjad]	*ym* + *cudd* / *ymguddio*	'oneself + concealed'/ 'to hide oneself'
		[dolig]	~	[mdolig]	*Nadolig*	'Christmas'
		[əmgernjal]	~	[mgernjal]	*ymgernial*	'to quarrel'
		[əmdreχgar]	~	[mdreχgar]	*ymdrechgar*	'energetic'
		[əmdreχfa]	~	[mdreχfa]	*ymdrechfa*	'struggle'
		[əmdreχɨ]	~	[mdreχɨ]	*ymdrechu*	'to make great effort'

The examples in (3a) are instances of alternations. Whenever the phonological situation changes (i.e. a suffix is added, which reassigns stress to the newly formed penultimate syllable), the vowel that finds itself in the unfavourable antepenultimate position ends up being deleted. The resulting clusters are now word-initial (#VCC > #CC) and can be said to follow Sonority Sequencing as they have an ST or TR sonority profile. For the time being we will assume that ST clusters are a special instance of Sonority Sequencing observance and that initial [ɬ] belongs to this category. In other words, the clusters that surface due to syncope follow the same phonotactic restrictions as those described in 2.1.

Analogical alternations are listed in (3b), where the vowel is deleted when the prosodic configuration changes and the syllable receives quaternary stress. Importantly, the vowel does not follow the word boundary in this case. It separates two consonants which then form a word-initial cluster (#CVC > #CC). Again, Sonority Sequencing is preserved and the resulting clusters have a TR and ST sonority profile, which agrees with the word-initial phonotactics.

(3c) includes alternations of a different type. The vowel is already in the antepenultimate syllable but its pronunciation is optional. The list includes alternative phonetic realisations of the same lexical items, which might indicate that the systemic rejection of antepenultimate

vowels is not well-established in the system and it is possible that we are facing an ongoing diachronic change: vowel deletion is known to have existed before ST clusters and is frequent in the language across dialects. It is introduced by analogy to other vowels even inside clusters if there is an alternation caused by the stress shift. Lexical items that possess a vowel with quaternary stress but do not show signs of stress shift may or may not delete the vowel. Be that as it may, this type of syncope can create clusters whose sonority profile is more ambiguous. First, [dv] can be interpreted as a TT cluster, since both members are obstruents, or a TR cluster, since the second member is more sonorous than the first one. Second, the [ɬf] cluster can be assumed to be an ST cluster (we have already allowed the lateral-fricative to occupy this special place), since it is a sibilant followed by an obstruent, or a TT cluster since both members are fricatives. The remaining sequences fall into the category of TR clusters.

Perhaps the most revealing clusters are listed in (3d). The list includes both alternations resulting from stress shift as well as alternative pronunciations with and without initial clusters. As for the sonority profile of the resulting clusters, not only do they violate Sonority Sequencing (nasals are more sonorous than stops) but in addition they are composed of three members, which will exclude them as candidates for branching onsets. Interestingly, the only nasal allowed to occupy the initial position in such a cluster is [m]. More interestingly, it never shares the place of articulation with the following stop, which is atypical of nasal-stop clusters in general (Harris 1994: 69) and Welsh clusters in particular (Awbery 1984: 86).

If the spelling is to be taken seriously as linguistic evidence of diachronic change, it could be assumed that the ongoing processes have not yet found new orthographic conventions. For example, the spelling of English *night* indicates the presence of a fricative-stop cluster, while Polish *góra* ['gura] 'mountain' suggests that the first vowel has a mid round quality. Neither of these spelling conventions reflects the phonetic shape of words without historical bias. Thus, the alternations in (3d) and the modern spelling tradition combined provide evidence that the representations of ['mgenɑχ], ['mdolig] or ['mdreχkɑr] should incorporate an alternating vowel which is marked 'ø' [ø'mgenɑχ], ['mødolig] or [ø'mdreχkɑr].

To sum up, Welsh allows ST and TR clusters in the initial position in citation forms, that is words unaffected by concatenative morphology and subsequent phonological processing. Marginally, clusters of the TRR type are also found in the word-initial position. Consonant sequences which arise through the syncope of the initial-syllable vowel conform to the phonotactic restrictions in most cases, that is they are ST and TR clusters, but some dubious sequences like [dv] or [ɬf] may also surface. The clusters that introduce the most information on the well-formedness of the initial sequences are [mg], [md] and [mdr] since they clearly violate the Sonority Sequencing Principle.

2.3. Relative strength of the initial consonant

The last aspect of the phonotactic setting in Welsh which needs investigation is the relative strength of consonants in the word-initial position. In order to say anything about the strength of the initial position, or lack thereof, we must compare the distributional freedom of single consonants in different positions within the word, which will be done in section 2.3.2.

However, no analysis of initial consonants with respect to their strength or weakness would be complete without at least a mention of a consonant-related phenomenon common to all Celtic languages, namely the Initial Consonant Mutation (ICM).

2.3.1. Initial Consonant Mutation

This section will include only the most rudimentary information on ICM in Welsh. A fuller description can be found in Ball and Müller (1992), while for purely phonological accounts of ICM the reader is referred to Cyran (2010: 45-71) and Hannahs (2013: 120-149). There are three mutation types in Welsh that affect the quality of the initial consonant, namely Soft Mutation (SM), Nasal Mutation (NM) and Aspirate/Spirant Mutation (AM). They are triggered by morpho-syntactic conditions rather than by any phonological factors. The oft-quoted explanation is that the masculine, feminine and plural possessive pronouns have the same phonetic shape but influence the following consonant differently. This is illustrated below:

(4)	a.	[i kɑːθ]	[i gɑrð]	[i trol]	[i kweːst]
		eu cath	*eu gardd*	*eu troll*	*eu cwest*
		'their cat'	'their garden'	'their car'	'their inquest'
	b.	[i gɑːθ]	[i ɑrð]	[i drol]	[i gweːst]
		ei gath	*ei ardd*	*ei droll*	*ei gwest*
		'his cat'	'his garden'	'his car'	'his inquest'
	c.	[i χɑːθ]	[i gɑrð]	[i θrol]	[i χweːst]
		ei gath	*ei ardd*	*ei droll*	*ei gwest*
		'her cat'	'her garden'	'her car'	'her inquest'

The initial consonant in (4a) remains unaffected (that is it retains its shape from the citation form), while those in (4b) are (4c) are mutated. The masculine possessive pronoun triggers SM, which voices the voiceless stops and causes [g] to be deleted. The nouns following the feminine possessive pronoun undergo AM, which turns voiceless stops into voiceless fricatives with the corresponding place of articulation. It fails to target voiced stops, hence *gardd* remains unchanged. To conclude, it is the changing semantic value of the pronoun, not its phonological properties, that affects the following consonant. The precise changes that are triggered by mutations are listed below:

(5)	a.	Soft Mutation	target:	p	t	k	b	d	g	m	ɬ	r̥ʰ	tʃ
			result:	b	d	g	v	ð	--	v	l	r	dʒ
	b.	Nasal Mutation	target:	p	t	k	b	d	g				
			result:	m̥ʰ	n̥ʰ	ŋ̊ʰ	m	n	ŋ				
	c.	Aspirate Mutation	target:	p	t	k							
			result:	f	θ	χ							

Soft Mutation has the largest number of both target consonants and triggering contexts. Nasal Mutation affects only stops and contributes nasal resonance to their pronunciations

without affecting their voicing. Finally, Aspirate Mutation targets only voiceless stops. Traditionally, all of these changes (except [m] > [v] perhaps) can be viewed as lenition processes, thus mutations affect the strength of the initial consonants by depleting it. Voiceless stops become voiced, spirantised or nasalised, which moves them closer to vowels on the sonority scale; voiced consonants spirantise, nasalise or delete, which also increases their sonority. The voiceless liquids lose their voicelessness, which turns them into fully-fledged sonorants. Consequently, mutations will be treated as processes of weakening rather than strengthening capacity.

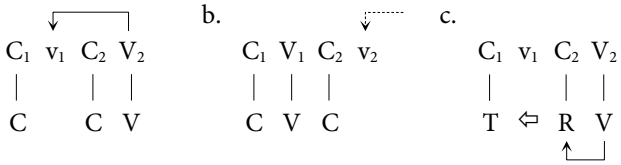
2.3.2. Strong vs. weak positions

Harris (1994: 194-225) analyses different weakening phenomena that affect the pronunciation of English [t] in various dialects and comes to the conclusion that it is mostly lenited word-finally (e.g. [pʊt] > [pʊʔ]), before a consonant (e.g. [ˈhɒtkiː] > [ˈhɒʔkiː]) and foot-initially between two vowels (e.g. [ˈprɪti] > [prɪɾi]). It is true of many consonants in many systems that they experience weakening processes in similar (_#, _C, V_V) contexts. A theory which will be used to justify the typological categorisation of Welsh is described in detail by Scheer (2004: 117-147) and it capitalises on the correlation between positional strength and weakness, and lateral relations between segments. Although the principles behind consonantal interactions will be largely simplified on the following pages due to space limitations, it should suffice to demonstrate that the theory offers a non-arbitrary representation of ill- and well-formed initial clusters along with implications for the representation of a domain.

The Lateral Theory of Phonology represents the syllabic structure using strictly alternating onsets and nuclei represented as C's and V's, hence its shorthand name CVCV. The constituents may or may not be filled with melodic content, thus a certain amount of tolerance for empty constituents must be accounted for. Since only nuclei can be active sources of phonological enforcement, while silence (or lack of any sonority) is an inherent property of a consonant, the emptiness of a C position will not be discussed any further. As far as the emptiness of a nucleus is concerned, it can be and must be justified in one of the following ways: a nucleus may remain empty if it is governed by the following melodically realised nucleus (6a). A nucleus may remain empty if it is domain-final and the parametric setting of the system in question allows Final Empty Nuclei (FEN) (6b). A nucleus may remain empty if it is circumscribed by a governing relation between two consonants in a branching onset (6c) (Scheer 2004: 67)[4]:

[4] CVCV uses two opposing lateral forces which marshal inter-constituent relations: (a) Government is a force spoiling the melody. If an onset is governed, it is likely to undergo lenition (lose some melodic material. If a nucleus is governed, it remains uninterpreted phonetically. (b) Licensing is a lateral force granting melodic 'health' to a position. If a constituent is licensed it is likely to support more melodic material. Infrasegmental Government is also a lateral force in this sense since it is a relation between two consonants which silences the empty nucleus straddled by the two members (Scheer 2004: 161).

(6) a. b. c.

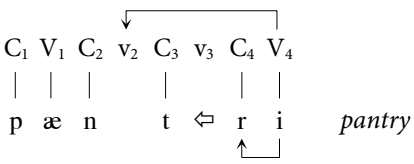


The nucleus v_1 in (6a) is empty because it is governed by the nucleus immediately to its right. The direction of government is marked with the arrow above the constituents whereas empty nuclei are shown with a lowercase 'v'. Governed nuclei are encountered inside clusters which cannot be regarded as branching onsets. For instance, the structure of the English word *bandit* [bænødɪtø] includes an empty nucleus between [n] and [d], which is governed by the following [ɪ]. Importantly, empty nuclei which never surface in vowel-zero alternations can be governed by FEN's, hence the well-formedness of *band* [bænødø]. Such empty nuclei are found in coda-onset clusters (RøT) or 'bogus' clusters (TøR) but not in branching onsets (TR).

The graph (6b) represents a domain-final nucleus that is empty, which cannot be taken for granted. Languages like Italian or Japanese do not have consonant-final words, which, according to this theory, means that the Final Empty Nucleus parameter is switched off and only (6a) and (6c) could potentially silence a nucleus. English, on the other hand, is a FEN language, thus it allows consonant-final words like *ban* [bænø], *bad* [bædø], *band* [bænødø] or *bandit* [bænødɪtø].

Finally, what will be of greatest interest to us is that nuclei remain empty inside an Infrasegmental Government (IG) domain (6c). Sonorants, which are deemed more complex than obstruents in melodic terms (see Scheer 1999a), are the only consonants capable of governing their neighbours. In order to do that, they have to be licensed by the following nucleus which is not itself governed. Such empty nuclei are invisible to phonological processes – they neither require government nor block its application across a TR cluster. An example from English illustrates the application of Government across a TR cluster:

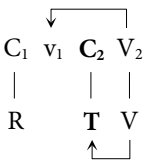
(7)

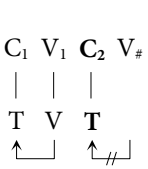


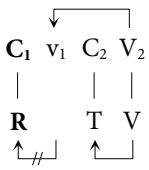
pantry

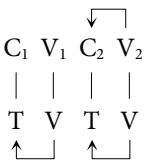
The nucleus V_4, which is filled with melody, has two duties to perform. First, it licenses the preceding T ⇦ R governing relation (the arrow underneath), which silences the nucleus v_3. Second, it governs the preceding nucleus v_2, which cannot be silenced by an ill-formed branching onset [nt].

We have thus introduced the two lateral forces which are discharged by a nucleus, namely government (silencing nuclei) and licensing (allowing inter-consonant relations). CVCV assigns two antipodal effects to the two antagonistic lateral mechanisms. If a position is licensed, its segmental complexity (melodic expression) is secured. If, on the other hand, a position is governed, its segmental complexity is depleted (Scheer 2004: 148-180).

(8) a. 

b. 

c. 

d. 

The consonant C_2 in (8a) is licensed by the following nucleus V_2 but is not governed, since government is absorbed by the empty nucleus v_1. Thus, C_2 is licensed but ungoverned, which is recognised as the strong (post-coda) position. The domain-final empty nucleus $V_\#$ in (8b) is devoid of licensing potential, hence the consonant C_2 is found in a weak (word-final) position. Similarly, the consonant C_1 in (8c) is unlicensed since the following nucleus is governed and therefore not a lateral actor. This configuration is considered a weak (coda) position. Finally, the representation in (8d) contains two melodically filled (laterally active) nuclei. Since V_2 has no nuclear target to hit, it concentrates its governing capacity on the preceding onset. Therefore, certain languages display weakening phenomena in the intervocalic position. For the sake of convenience (and since it is not immediately relevant to the present discussion) cases where a position is both licensed and governed, or neither licensed nor governed at the same time, will be ignored.

Thus, CVCV acknowledges two weak positions for consonants, namely intervocalic (VCV – governed) and Coda (Cø – unlicensed) and one strong position dubbed the Coda Mirror (øC – licensed, ungoverned). Be that as it may, the positional factors are not the only aspects which translate into the strength or weakness of a consonant. Nevertheless, in order not to exceed the scope of the present investigation, we will rely on relative strength, as put forward by the Coda Mirror theory. Let us now compare the distribution of Welsh consonants with respect to their freedom to occupy the word-initial, word-final, intervocalic and pre-consonantal positions within a word.

(9)	a.	[**p**ɑdɑr]	*pader*	‘the Lord’s prayer’
		[**t**ɑvod]	*tafod*	‘tongue’
		[**k**enedl]	*cenedl*	‘nation’
		[**b**ɑːχ]	*bach*	‘small’
		[**d**iktɑr]	*digter*	‘anger’
		[**g**elɨn]	*gelyn*	‘enemy’
		[**tʃ**erk]	*siarc*	‘shark’
		[**dʒ**ob]	*job*	‘job’
		[**f**əðlon]	*ffyddlon*	‘faithful’
		[**θ**əlɨ]	*erthylu*	‘to miscarry’
		[**s**ɑːf]	*saff*	‘safe’
		[**ʃ**iːr]	*sir*	‘shire’
		[**ɬ**ivjo]	*llifio*	‘to saw’
		[**χ**udɨ]	*chwydu*	‘to vomit’
		[**h**egɑl]	*hegl*	‘foot’
		[**v**el]	*fel*	‘like’
		[**m**ɨːsk]	*mysg*	‘midst’
		[**n**ivɑr]	*nifer*	‘number’
		[**l**ɑrdjo]	*lardio*	‘to break down’
		[**r**ɑːs]	*ras*	‘race’

	[**r̥ʰ**uːd]	*rhŵd*	‘rust’
	[**w**imblad]	*gwimled*	‘gimlet’
	[**j**eχid]	*iechyd*	‘health’
b.	[ka**p**ten]	*capten*	‘captain’
	[do**k**tor]	*doctor*	‘doctor’
	[ki**b**ðaɬ]	*cibddall*	‘dull of comprehension’
	[kənɬəi**d**va]	*cynlleidfa*	‘congregation’
	[ki**g**vran]	*cigfran*	‘raven’
	[kre**f**tur]	*crefftwr*	‘craftsman’
	[krə**v**dur]	*cryfder*	‘strength’
	[noi**θ**ter]	*noethder*	‘nakedness’
	[go**s**tag]	*gosteg*	‘silence’
	[gwa**ɬ**tog]	*gwalltog*	‘hairy’
	[nə**χ**tod]	*nychdod*	‘asthenia’
	[a**ð**vad]	*addfed*	‘ripe’
	[ɬa**m**bad]	*Llanbedr*	‘place name’
	[a**n**duɨol]	*andwyol*	‘harmful’
	[ka**ŋ**kar]	*cancr*	‘cancer’
	[e**l**van]	*elfen*	‘element’
	[du**r**nod]	*diwrnowd*	‘day’
c.	[ta**p**]	*tap*	‘tap’
	[rəi**t**]	*reit*	‘right’
	[sli**k**]	*slic*	‘slick’
	[maː**b**]	*mâb*	‘son’
	[dleː**d**]	*dlêd*	‘debt’
	[deː**g**]	*dêg*	‘ten’
	[wa**tʃ**]	*wats*	‘watch’
	[sosi**dʒ**]	*sosej*	‘sausage’
	[hu**f**]	*hwff*	‘rough push’
	[kaː**θ**]	*cath*	‘cat’
	[druː**s**]	*drŵs*	‘door’
	[erχi**ɬ**]	*erchyll*	‘abominable’
	[moː**χ**]	*môch*	‘pig’
	[ɬaː**ð**]	*llâdd*	‘to kill’
	[**r̥**ʰigu**m**]	*rhigwm*	‘rhyme’
	[meli**n**]	*melin*	‘mill’
	[dre**ŋ**]	*dreng*	‘rude’
	[əsko**l**]	*ysgol*	‘school’
	[ɬaθa**r**]	*llathr*	‘bright’
d.	[ʃo**p**a]	*siopao*	‘shops’
	[ɬe**t**i]	*llety*	‘inn’
	[ku**k**u]	*cwcw*	‘cuckoo’
	[di**b**an]	*diben*	‘purpose’
	[ka**d**aχ]	*cadach*	‘cloth’
	[e**g**ar]	*eger*	‘a bore’
	[kra**tʃ**an]	*cratsian*	‘to creak’
	[pro**dʒ**akt]	*project*	‘project’
	[ho**f**i]	*hoffi*	‘to like’
	[ɬe**θ**ar]	*llethr*	‘slope’
	[ɬə**s**an]	*llysan*	‘bilberry’

[fɑʃun]	*ffasiwn*	'fashion'
[aɬan]	*allan*	'out'
[biχɑs]	*buches*	'a number of cows together'
[gwɑˈhɑːn]	*gwahan*	'apart'
[dovi]	*dofi*	'to tame'
[ɑðɑ]	*Adda*	'Adam'
[dəmɑ]	*dyma*	'here is / are'
[anadl]	*anadl*	'breath'
[aŋan]	*angen*	'necessity'
[kalon]	*calon*	'heart'
[karɑ]	*carrai*	'lace'

This rather lengthy list is an exhaustive index of which consonants may and may not occupy a particular position in a Welsh word. (9a) contains examples of consonants in the word-initial position. The only consonants exempted from the initial position are [ð] and [ŋ]. Nonetheless, all the examples in (9) are citation forms, which means that mutated reflexes were not taken into consideration. Thereupon, all [d]-initial and [g]-initial nouns will supply relevant examples when Soft Mutated and Nasal Mutated respectively.

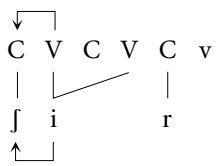
The pre-consonantal coda position illustrated by (9b) bans a slightly higher number of consonants. First, the dental stop [t] appears before other stops or fricatives exclusively in foreign words and morphologically complex expressions where it is separated from the following consonant by a morpheme boundary. Second, the affricates, which are fairly new structures in Welsh, are not found before other consonants. Third, the fricatives [ʃ] and [h] are not found pre-consonantally. Fourth, glides [w] and [j] as well as the voiceless nasals [m̥ʰ n̥ʰ ŋ̊ʰ] will be omitted. The occurrence of the former depends on whether we interpret the off-glides in diphthongs as vocalic or consonantal (see Awbery 1984, 1986), while the latter are confined to the initial position by Nasal Mutation.

The word-final coda (9c) allows a somewhat broader spectrum of consonants. The only segments banned from this position are fricatives [ʃ], [h] and [v], the last one being a characteristic feature of North Welsh phonology.

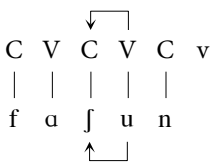
Finally, the intervocalic position allows all consonants (except for the aforementioned glides and voiceless nasals) to appear.

To conclude, the word-initial position patterns with the intervocalic one as far as the distributional properties of consonants are concerned by allowing most (if not all) consonants. The pre-consonantal and the word-final codas share the property of being more restrictive in general, and excluding the same fricatives in particular. Thus, distributionally speaking, the word-initial and intervocalic positions can be regarded as one type, while the pre-consonantal and word-final positions are another. Let us attempt a theoretical justification:

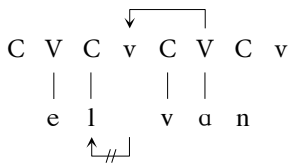
(10) a.



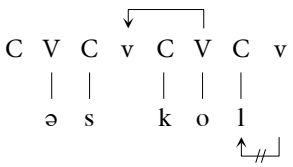
b.



c.



d.



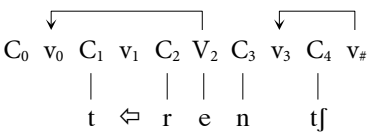
A governed consonant is represented in (10a) and (10b) and we can see that government has no particular effect upon the melodic expression of the consonant. It is true that if there are consonant weakening phenomena in a language, they are likely to concern a governed consonant but no such requirement is imposed. In other words, it is possible for a governed consonant to remain unchanged and this is exactly what happens in Welsh. However, when a consonant is devoid of licensing as in (10c) and (10d), its melodic expression may be suppressed. That is why unlicensed positions can host a more restricted range of consonants in Welsh than the licensed ones.

3. Typological Divisions in the Lateral Theory of Phonology

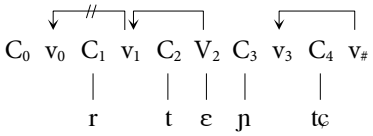
3.1. Anything-goes vs. TR-only

The Lateral Theory of Phonology (known as LTP or CVCV) offers a non-arbitrary typological division based on the presence vs. absence of an inter-consonantal relation between the two members of a cluster. According to CVCV, all languages of the world belong to either of the following categories: 'anything-goes' languages, which allow clusters of rising sonority (TR), even sonority (TT, RR) and falling sonority (RT) in the initial position, or 'TR-only' languages, which allow only the first type. Scheer (1999a, 2004: 95-116 and 2007) assumes that this is not an observational fact about languages but it provides a theoretical device which expects languages to belong to either type. It is assumed that the left-edge of the word in a language might be marked for phonology with a phonological object that is an empty CV unit (cf. Lowenstamm 1999). The empty nucleus of this unit is visible to phonological processing and therefore requires government so as to remain silent. Since, as shown in (7), TR clusters are the only ones that do not block the application of government, it is these clusters that can follow the initial empty CV. The graphs in (11) compare the representations of the English (TR-only) word *trench* and Polish (anything-goes) *rtęć* 'mercury'.

(11) a.



b.



In (11a), the nucleus v_1 is circumscribed by the IG relation contracted by the initial clusters, so the laterally active nucleus V_2 has a green light to target the initial CV with government. Contrarily, there is no IG relation in the RT cluster in (11b). The nucleus v_1 is governed, which makes it laterally inactive, therefore the initial CV cannot be silenced. The typological difference is thus between languages that allow a phonologically visible empty nucleus v_1 and those where it is either filled with melody or taken care of by IG.

The presence of the initial empty CV manifests itself in the following way: firstly, the only type of clusters allowed initially are branching onsets (IG) for reasons explained above. Secondly, the vowel in the first syllable never alternates with zero – it must remain laterally active in order to silence the initial CV. Thirdly, the first consonant finds itself in a strong position – it follows an empty nucleus, thus it is in a licensed but ungoverned position.

What appears to be problematic for this typology is that TR-only languages can also allow STR clusters which should block the application of government. These will be briefly discussed in the next section.

3.2. Magic Licensing

The special status of s+C clusters has to be taken into account in Standard Government Phonology (e.g. Kaye 1990, 1996 and Harris 1994) as well as its daughter theories (Scheer 2004; Cyran 2010). Kaye (1996) observed that s+C clusters behave in languages in a special way. For example, they are the only fricative-stop clusters for many languages (e.g. Japanese). What is more, they are the only fricative-stop clusters in the initial position for many languages (e.g. English, German, Italian). English long vowels are allowed to precede s+C clusters (e.g. *task* [tɑːsk], *paste* [peɪst]).[5]

In Standard Government Phonology, they were rejected as possible candidates for branching onsets for melodic reasons. Obstruents are said to be more complex than sonorants (they contain more elements) and this asymmetry allows them to contract a constituent governing relation, which is head-initial. That is T governs R (cf. Harris 1994: 170-177). Since [s] is more sonorous (less complex) than stops, it cannot be regarded as a possible governor in s+C branching onsets. For this reason, Kaye (1996) proposed that s+C be viewed as coda-onset clusters. Although it remedied the ill-formedness of STR clusters, it introduced further complications – why does [s] belong to a syllable which does not exist? The syllable in which [s] is a coda was assumed to be headed by an empty nucleus. Nevertheless, the emptiness or silence of that nuclear position remained unaccounted for. Kaye (1996) realised the shortcomings of that analysis, resorting to the acknowledgement of such structures with no apparent justification. Therefore, s+C clusters came to be known as 'magic clusters' in the

[5] This is also true of nasal-stop alveolar clusters (e.g. *branch* [brɑːntʃ], *paint* [peɪnt]).

Government Phonology tradition, for the mechanism responsible for silencing the empty nucleus was 'magic licensing'.

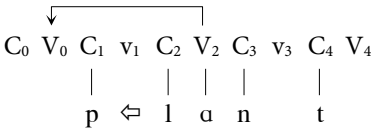
The Lateral Theory of Phonology offers no consolation in this matter. The major difference is that s+C clusters are represented as søC ones with an empty nucleus demanding government to remain inaudible. Still, the emptiness of the initial CV before STR or ST clusters begs to be resolved.

To sum up, languages can be divided into four categories according to the typologies explained in section 3: Those that mark the beginning of the word with an empty CV versus those that do not do so, and those languages where the 'magic licensing' parameter is set to 'on' versus those where it is set to 'off'. The ensuing section views Welsh from these vantage points.

4. Welsh as an anything-goes language

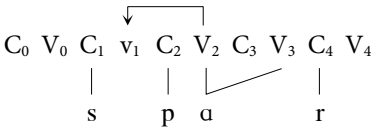
The examples in (1) above include TR, ST and STR clusters, which seems to suggest that the 'magic licensing' parameter is set to 'on' and takes care of the initial CV:

(12) a.



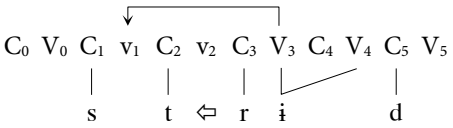
government across an IG domain

b.



'magic licensing' of C_0V_0

c.



government across an IG domain and 'magic licensing' of C_0V_0

As illustrated above, the existence of the initial CV can be justified if we assume that TR clusters indeed form branching onsets and that ST clusters do magically take care of the preceding empty nucleus. However, the shaky status of s+C clusters and the history of the prosthetic vowel testify otherwise. It seems that 'magic licensing' was ungrammatical around Old/Middle Welsh, when the prosthetic vowel was introduced, and its remnants can still be detected in vowel-zero alternations before ST and in orthographic conventions. If we assume then that ST clusters do not trigger 'magic licensing', there are two important implications: first, the initial CV is not there in Welsh. Second, RT and CTR clusters should be well-formed initially as well.

Let us entertain this possibility for a moment. If there is no initial CV, we should find evidence for an even and falling sonority slope in initial clusters, vowel instability at the left edge and a lack of strength of the initial consonant.

There are at least two pieces of evidence suggesting that Welsh allows RT clusters initially. First, there are examples of falling-sonority clusters in Welsh which result from vowel syncope

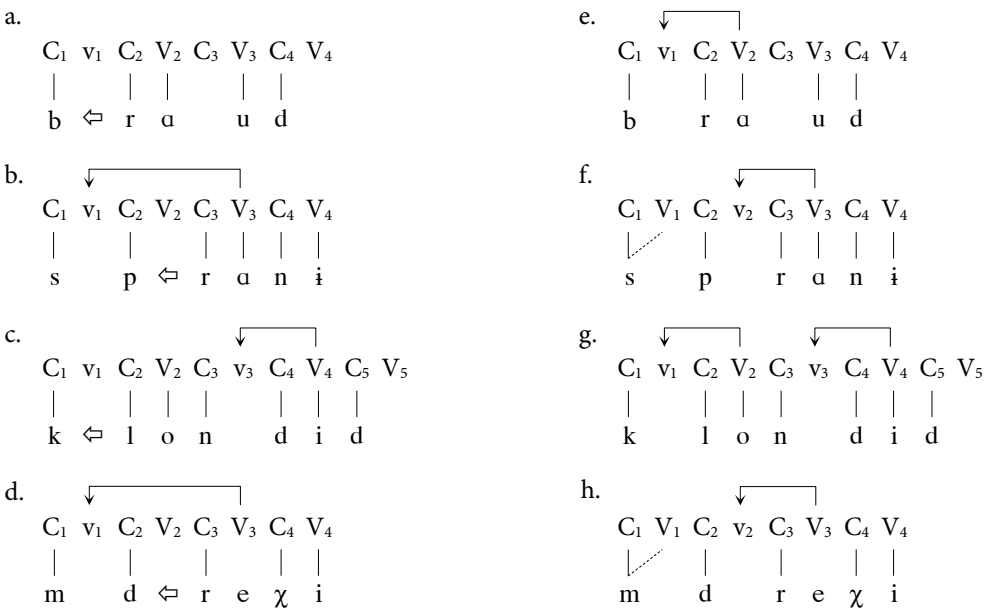
(3d). Interestingly, they all start with [m]. Additionally, the handful of TRR clusters in (2) can hardly be assumed to render a branching onset. Second, stops in TR clusters are turned into nasals by Nasal Mutation, which means that (at least in mutation contexts) NR clusters are well-formed. Thus, there is no proof that Welsh is a TR-only language.

As far as the stability of the initial vowel is concerned, Welsh deletes the vowel in the antepenultimate syllable provided that it is the first syllable of a word. The vowel is thus silenced by government and could not possibly guarantee the existence of an initial CV.

Moreover, the word-initial position offers no particular strength to the consonant. At least, there is no greater strength initially than can be found intervocalically. Interestingly, though, morpho-phonological processes may deplete the segmental expression of the word-initial consonant but they do not seem to enhance it.

Taking all that into consideration, we may propose the structure of the left edge of a word in Welsh as follows:

(13)



The graphs in (13a-d) represent an assumption where TR clusters may form branching onsets in Welsh, while those in (13e-h) illustrate a situation without any branching onsets whatsoever. The former option requires little justification, whereas the latter one is slightly more controversial. The example in (13c) should be approached with caution since it is clearly a syncopated form. This paper does not make a serious commitment to the assumption that there are no branching onsets in Welsh but merely acknowledges such a possibility.

The branching consonants in (13f) and (13h) represent trapped consonants (Scheer 2004: 283-364). It should be emphasised that only sonorants may extend their melody to a neighbouring nuclear position. [s] must be assumed to be a special obstruent that behaves like a sonorant with respect to occupying neighbouring constituents. The nucleus which accommodates the sonorant melody is no longer empty and requires no government from the following vowel. Moreover, [m] fails to behave like a coda consonant refusing to adopt the place of articulation of the following onset. That is because it is not a coda in (13h) – it is

followed by a melodically filled nucleus, which makes it an intervocalic nasal (in terms of lateral relations).

The possibility that branching onsets do not exist in Welsh can be reinforced by two additional facts. Firstly, the first member (T) of a branching onset undergoes various weakening phenomena in morpho-syntactic contexts including nasalisation, spirantisation or even deletion. It is hard to accept that [gn], [gl], [gr] and [gw] are branching onsets in the light of their realisation as [n], [l], [r] and [w] respectively in Soft Mutation. Secondly, vowels in South Welsh cannot be long before clusters of whatever sonority slope. If more than two consonants follow a vowel, it must remain short (Awbery 1984: 66). In that respect TR clusters and RT clusters behave alike, therefore it is feasible to assume that they should have the same structure, that is CøC.

5. Conclusion

The aim of this paper was to investigate word-initial clusters in Welsh from a typological point of view. A vantage point was offered by the Lateral Theory of Phonology, which classifies languages as either TR-only (CV-initial) or anything-goes (non-CV-initial). All in all, there is enough evidence in Welsh to seriously question the existence of an initial CV: The descriptive accounts that state that the only clusters found in the initial position in Welsh are TR and STR have been proven inaccurate – we should also include TRR, TT, RT and RR (the last one at least as a result of Nasal Mutation). Furthermore, a number of initial ST and TT clusters are the result of vowel syncope. In many cases it is the word-initial vowel that ends up being deleted stranding the clusters without initial support, although there are instances where it is a #CVC sequence that surfaces as #CC due to syncope. Finally, it cannot be asserted that the initial consonant in a word in Welsh is of a particular strength. This is crucial from the theoretical perspective since it sheds light on the structural configuration, that is the initial consonant behaves like other governed consonants, ergo its position in a Coda Mirror is possible but not evident. Finally, the lack of an initial CV (and the employment of trapped consonants) allows us to assume that branching onsets in Welsh are completely ill-formed. Nevertheless, this assumption requires more extensive research concerning the behaviour of consonant clusters in the medial and final positions as well as internal and external sandhi phenomena.

References

Anderson, Cormac. 2011. How English is Irish? The typology of Irish initial consonant clusters. In *Formal and Historical Approaches to Celtic Languages*, edited by Krzysztof Jaskuła, 13-33. Lublin: Wydawnictwo KUL.

Awbery, Gwenllian M. 1984. Phonotactic constraints in Welsh. In *Welsh Phonology. Selected Readings*, edited by Martin J. Ball, and Glyn E. Jones, 65-104. Cardiff: University of Wales Press.

Awbery, Gwenllian M. 1986. *Pembrokeshire Welsh. A Phonological Study*. Llandysyl: National Museum of Wales.

Awbery, Gwenllian M. 2010. Welsh. In *The Celtic Languages*, edited by Martin J. Ball, and Nicole Müller, 359-426. London and New York: Routledge.

Ball, Martin J., and Nicole Müller. 1992. *Mutation in Welsh*. London and New York: Routledge.

Cyran, Eugeniusz. 2010. *Complexity Scales and Licensing in Phonology*. Berlin and New York: Mouton de Gruyter.

Fynes-Clinton, Osbert H. 1913 [1995]. *The Welsh Vocabulary of the Bangor District*. Oxford: Oxford University Press.

Griffen, Toby D. 1979. On phonological stress in Welsh. *Bulletin of the Board of Celtic Studies* 28, 106-112.

Hannahs, S. J. 2013. *The Phonology of Welsh*. Oxford: Oxford University Press.

Harris, John. 1994. *English Sound Structure*. Oxford: Blackwell.

Iosad, Pavel. 2012. *Representation and Variation in Substance-free Phonology. A Case Study in Celtic*. University of Tromsø, PhD dissertation.

Kaye, Jonathan. 1996. Do you believe in magic? The story of s+C sequences. In *A Festschrift for Edmund Gussmann from his Friends and Colleagues*, edited by Henryk Kardela, and Bogdan Szymanek, 155-176. Lublin: Wydawnictwo KUL.

Lowenstamm, Jean. 1999. The beginning of the word. In *Phonologica 1996. Syllables!?*, edited by John Rennison, and Klaus Kühnhammer, 153-166. The Hague: Holland Academic Graphics.

Morris-Jones, John. 1913. *A Welsh Grammar. Historical and Comparative*. Oxford: Clarendon Press.

Scheer, Tobias. 1999a. A theory of consonantal interaction. *Folia Linguistica* 32, 201-237.

Scheer, Tobias. 1999b. On constraints vs. non-circular approaches to word-initial clusters. In *Phonologica 1996. Syllables!?*, edited by John Rennison, and Klaus Kühnammer, 289-304. The Hague: Holland Academic Graphics.

Scheer, Tobias. 2004. *A Lateral Theory of Phonology. What is CVCV, and Why Should it be?* Berlin and New York: Mouton de Gruyter.

Scheer, Tobias. 2007. On the status of word initial clusters in Slavic (and elsewhere). *Formal Approaches to Slavic Linguistics* 15, 346-364.

Willis, David. 2010. Old and Middle Welsh. In *The Celtic Languages*, edited by Martin J. Ball, and Nicole Müller, 117-160. London and New York: Routledge.