The ATR/Laryngeal connection and emergent features*

Pavel Iosad
The University of Edinburgh
pavel.iosad@ed.ac.uk

Revised version: 16th July 2021

In this paper I consider an unusual sound change and synchronic alternation in south-eastern dialects of Welsh, traditionally referred to as 'provection' (Welsh *caledíad*). This phenomenon involves the 'devoicing' of stops following stressed vowels. I analyse it as the result of a typologically rare kind of vowel-consonant featural interaction, where 'devoicing' is triggered by tense vowels. This is unexpected typologically, since tongue root advancement – often a correlate of 'tense' vowel quality – is more frequently associated with *voicing* rather than devoicing in obstruents. I argue that in Welsh the phonologization of the quality distinction in stressed vowels has created a system where such an interaction is dispreferred.

I suggest that the unusual consonant-vowel interaction in south-eastern Welsh is facilitated by the phonological ambiguity of the 'tenseness' contrast in vowels. Although the distinction between 'tense' and 'lax' vowel pairs is made consistently in most Welsh dialects, it is often fairly inert phonologically. This makes it possible for it to become involved in a variety of phonological processes. I survey the diverse ways in which quantity-quality interactions in vowels found across Welsh dialects, and argue that the attested multiple paths of phonologization support a bottom-up, substance-free approach to phonologization. I suggest that the particular strength of the substance-free model lies in its ability to account for different – sometimes drastically so – phonological behaviour emerging from ambiguous data.

The organization of this paper is as follows. In section 1 I describe the phenomenon of provection, survey the existing literature, and address the issue of whether provection involves neutralization of laryngeal contrast in stops. Section 2 presents the proposal for the synchronic

*The bulk of this paper was presented at the GLOW workshop *Phonological Specification and Interface Interpretation* at Hogeschool-Universiteit Brussel (April 2014). Related material was also presented at the Symposium on Historical Phonology (Edinburgh, January 2014), the 8th North American Phonology Conference (Montreal, May 2014), the 21st and 22nd Welsh Linguistics Seminars (Gregynog, July 2014 and July 2015), and the 1st Conference on Linguistic Diversity in Wales (Aberystwyth, July 2014). Thanks to the audiences at these fora, in particular Gwen Awbery, Daniel Currie Hall, Michaela Hejná, Bill Idsardi, Ineke Mennen, Andrew Nevins, Ros Temple, Bert Vaux, and Alan Yu. For help with fieldwork in South Wales, thanks to Christine Jones and Mererid Hopwood (University of Wales Trinity Saint David) and to Diarmuid Johnson (Menter Rhos-y-Gilwen). Thanks to Josef Fruehwald for invaluable statistical and programming advice. I am particularly grateful for the comments and suggestions related to the draft by three reviewers and Brett Miller. All errors of fact or interpretation remain my own.

analysis of provection, and section 3 addresses a number of assumptions on which this analysis rests. In section 4 I offer a hypothesis for the diachronic origin of this pattern, with a particular focus on the role of preaspiration as a precursor to the misparsing that led to provection. Section 5 offers a defence of the substance-free approach to the Welsh data over one that relies on a phonetically defined [(±)tense] feature. Section 6 is a brief conclusion.

1 The problem

1.1 Nature and extent of the phenomenon

This paper focuses on a feature characteristic of south-eastern dialects of Welsh and known in the literature as 'provection' (in Welsh, *caledíad* or *caledu*, literally 'hardening'). It is traditionally described as the 'devoicing' of 'voiced' stops after a stressed vowel.[1]

The process is both a historical sound change and a synchronic alternation. For clarity, the consonants of interest are highlighted. These examples, as all others in this paper unless indicated otherwise, are taken from the comprehensive description of the dialect of Nantgarw (Taff Valley) by C. H. Thomas (1993). The examples are given in the original transcription and in standard Welsh orthography.[2] Numbers indicate page references: by default these are to volume 2 of the book, which contains the glossary. I reproduce C. H. Thomas' transcriptions exactly in this section.

(1)	a.	[ke'**g**ina]	*ceginau*	'kitchens' (151)
	b.	['ke**k**ɪn]	*cegin*	'kitchen' (151)
(2)	a.	['govɪ**d**]	*gofid*	'regret' (413)
	b.	[go'vɪ**t**jo]	*gofidio*	'to regret' (414)

Historically, forms such as those in (1b) and (2b) should be compared to cognates found in other varieties of Welsh, which contain voiced stops following the stressed vowel: cf. ['kʲe**gʲ**in] 'kitchen', [go'vi**d**jo] 'to grieve' in Northern Welsh (Bangor [Fynes-Clinton 1913: s. vv.]).

Traditionally, provection is prevalent in the central and eastern parts of South Wales, in the old counties of Glamorgan and Monmouthshire. As described by C. H. Thomas (1975), it is (or was) strongly represented in the south-east, both in the Valleys (Thomas cites Tafarnau Bach in Sirhowy Valley, Merthyr Tydfil in the upper Taff Valley, Hirwaun and Rhigos, both in the upper Cynon Valley) and in the more low-lying parts of Glamorgan (Nantgarw, now a northern suburb of Cardiff; Ely Valley to the west of Cardiff; and Coity near Bridgend). Provection is also found in west Glamorgan, in the area around Swansea (e. g. Watkins 1967) and to the north in

[1] Somewhat confusingly, the literature on Welsh also uses 'provection' to refer to a number of different phenomena, some of which have nothing to do with the one discussed here (Greene 1967), see in particular section 1.3 below. Sometimes, the provection discussed in this paper is referred to as 'southern provection' (*caledíad y De*).

[2] While Welsh orthography is relatively straightforward, the following correspondences should be noted: <u> = [i] (in southern varieties, which are the focus of this paper), <w> = [u] or [w], <y> = [ə] or [i] (in southern varieties); <c> = [k], <dd> = [ð], <ng> = [ŋ] (rarely [ŋg]), <f> = [v], <ff> = [f], <ll> = [ɬ], <rh> = [r̥ʰ], although usually just [r] in southern dialects.

the Tawe and Loughor Valleys, and in the south of Breconshire to the north of the Valleys (e. g. G. E. Jones 2000). There are transitional zones in eastern Carmarthenshire (west of Llanelli) and northern Breconshire, where provection is found in a smaller number of lexical items and often in variation with unprovected forms.

Provection is treated in various levels of detail in the available descriptive literature on the relevant dialects. A large proportion of this work is summarized by C. H. Thomas (1975), to which we should add C. H. Thomas (1993), a book-length treatment of the dialect of Nantgarw with an extensive glossary. These descriptions all agree that the use of provection in these dialects was variable when the data was collected (mostly in the mid 20th century or later). Traditional dialectological descriptions are rarely explicit on the exact nature of the variation, but several studies of provection in a variationist sociolinguistic framework offer more insight: S. E. Thomas (1983) (with a short English summary in S. E. Thomas [1988]) for Ystalyfera in the upper Tawe Valley, B. Thomas (1990) for Pontrhydyfen in the Afan Valley, Flowers (1994) for Crynant in the Dulais Valley, M. C. Jones (1998: chap. 2) for Rhymney, and Brake (2011) for Treorchy in the Rhondda Fawr valley. Finally, Spooner (2016) conducted a small-scale acoustic study of provection in Ystalyfera.

In the remainder of this paper, I shall refer to the two series of Welsh stops as 'fortis' (the series traditionally transcribed as [p t k]) and 'lenis' (traditionally transcribed as [b d g]). They are often described in the literature as 'voiceless' and 'voiced' respectively; however, as I discuss in particular in section 3.1 below, this may be misleading. Occasionally the two series are referred to as 'aspirated' and 'unaspirated', but this can lead to confusion in the description of provection (section 1.3), so I eschew this usage as well, in favour of the less committal 'fortis' and 'lenis' terminology. However, I avoid using the terms 'fortition' and 'lenition'; while they may be adequate as descriptions of a segments acquiring 'fortis' or 'lenis' quality, they may introduce unintended theoretical connotations (see e. g. Honeybone [2008] for an overview).

1.2 Contexts for provection

As recorded in existing sources, provection is highly variable along several dimensions. First, its lexical incidence varies significantly from dialect to dialect: words with provection in one variety may lack it in another. Second, a dialect may allow both provected and unprovected versions of the same word. Third, provection is generally a recessive feature: variationist studies show a decrease in its use in apparent time, and stylistic grading. Some studies (e. g. S. E. Thomas 1983, Brake 2011) also find a link with education, tying the retreat of provection to the spread of the standard language at the expense of local varieties ('dialect death' [M. C. Jones 1998]).

Even so, it is possible to make some phonological generalizations. First, it can only occur following a *penultimate* stressed vowel (stress in Welsh polysyllables is overwhelmingly penultimate), and is almost never found in monosyllables: words like ['tæ**ːd**] 'father' never appear as *['tæ**ːt**]. An exception noted by Flowers (1994) in Crynant is found in the imperative forms of the verbs *rhedeg* 'run', *hwb* 'push', *plygu* 'fold' and *dywed* 'say', respectively ['ri(ː)t] 'run!', ['u(ː)p] 'push!', ['pli(ː)k] 'fold!', and ['gwe(ː)t] 'say!'.[3]

[3]Forms of the verb 'say' with initial [gw] are widespread across South Wales; they are not suppletive, but represent a phonetic development of the stem written as *dywed* in formal registers and *dweud* in less formal ones.

Second, the target consonant may be intervocalic, as in (1), or be followed by one of the glides [j w], as in (2), or sonorants [r l n m] ([ŋ] is not found in this position in Welsh):

(3) a. ['ɛ**p**rɪɬ] *Ebrill* 'April' (349)
b. ['ɛ**k**lʊs] *eglwys* 'church' (341)
c. ['trɔ**t**noθ] *troednoeth* 'barefoot' (817)
d. ['da**t**mɑr] *dadmer* 'thaw (of people)' (276; only example before [m] in Nantgarw)

Among non-sonorants, provection is also fairly common before [v]:

(4) ['dɛ**k**vɑd] *degfed* 'tenth' (279)

Before other consonants, it is very rare or not found at all. There are also very few examples of provection where the 'devoiced' consonant does not follow the stressed vowel directly, as in ['kʊr**t**ab] 'rennet' (cf. Rhigos ['kʊr**d**ab], *cyweirdeb*).

Third, provection never occurs with the lenis fricatives [v ð], which do not become [f θ] in a provection context, even though both segments exist in Welsh:

(5) a. ['kɛlʊ**ð**] *celwydd* 'untruth' (147)
b. ['klu**ð**a] *celwyddau* 'untruths' (147)
c. *['klu**θ**a]

Lenis fricatives can only be involved in provection (at least diachronically) if they follow a provected stop, as in person-inflected forms of the preposition [(h)eb] *heb* 'without': ['ɛ**pθ**o] for *hebddo* 'without him' in Nantgarw (p. 357); [ɛ**pθ**ɪn] for *hebddynt* 'without them' in Rhymney (M. C. Jones 1998).

There are also restrictions on what vowels can trigger provection. In the dialect of Llansamlet (now a suburb of Swansea) provection is triggered only by historical monophthongs: neither diphthongs (['ɬui**d**i] *llwydrew* 'mildew', [ɔi**d**ran] *oedran* 'age') nor monophthongs that derive from historical diphthongs (['pi**d**o] *peidio* 'cease', ['mi**d**u] *meudwy* 'hermit', ['gwi**d**i] *gwaedu* 'bleed') trigger provection (Watkins 1967). In the dialect of Nantgarw, much the same trends obtain; however, diphthongs, most often [ʊi], *can* be triggers (['ɬʊi**t**ro] for *llwydrew*), and there are also some cases of provection after historical diphthongs (['i**p**o] *heibio* 'go past').

Some sources claim that there is a tendency to avoid provection before a so-called svarabhakti vowel, which is inserted to break up word-final rising-sonority clusters (Hannahs 2009, 2011, Iosad 2017c): in Nantgarw, C. H. Thomas (1993) finds 11 examples like ['po**b**ol] *pobl* 'people' (vol. 1, p. 108), ['gwo**b**or] *gwobr* 'prize' (p. 460) with no provection, but only five like ['bɑ**k**al] *bagl* 'crutch' (p. 57) with provection. However, it is not entirely clear whether these generalizations are reliable: of Thomas' 11 forms with no provection, four include either diphthongs (['ɬʊi**b**ɪr] *llwybr* 'path' (p. 528), ['pəu**d**ʊr] 'powder' (p. 673)) or vowels deriving from diphthongs (['ɬi**d**ɪr] *lleidr* 'thief' (p. 519), ['ni**d**ɪr] *neidr* 'snake' (p. 529))

As I shall discuss in more detail below, in all the phonological contexts where provection is possible, it is not, in fact, obligatory: both intervocalic stops and stops before other consonants can remain unprovected, with a variety of conditioning factors.

In particular, as noted above, a single lexical item may show 'doublets', i. e. provected and unprovected forms of the same word, as in [di'o**g**al] or [di'o**k**al] *diogel* 'safe' (p. 309). Very occasionally such etymological doublets may diverge in meaning: in Nantgarw, ['gwo**b**or] without provection means 'prize' (its meaning in the literary language, from where it must have been borrowed according to C. H. Thomas [1993: vol. 1, p. 109]), while the dialect form ['gwa**p**ar] is used in the meaning 'bribe'.

Provection normally creates morphophonological alternations, as in (1) above, but this is not always the case. For instance, in some cases the 'failure' of provection is consistent across the paradigm, as in (6), where provection fails in a lexical item with a svarabhakti vowel.

(6)	a.	['gwi**d**ɪr]	*gwydr*	'glass' (454)
	b.	['gwɪ**d**ra]	*gwydrau*	'glasses' (454)

In other cases, provection may apply only in some forms within a paradigm, as in the following two examples:

(7)	a.	['bɑ**k**al]	*bagl*	'crutch' (57)
	b.	['ba**g**la]	*baglau*	'crutches'[4]

(8)	a.	['gwe**k**ɪl]	*gwegil*	'nape (of the neck)' (443)
	b.	['gwɛ**g**la]	*gwegilau*	'napes' (443)

Here, we find that provection applies despite the presence of a vowel-zero alternation (hence unlike 6), but only does so in forms where the provected consonant is intervocalic: it fails when the stop is followed by a sonorant (even though provection is not impossible before a sonorant).

In yet other cases, historical provection 'overapplies'. In 9, the stop following the [ɔ] is historically lenis, so the presence of [t] looks like provection by comparison with other varieties. However, in the inflected forms we find an apparently provected stop after an unstressed vowel. This creates a situation with no alternation.

(9)	a.	['mɔ**t**rɪb]	*modryb*	'aunt' (569)
	b.	[mə**t**'rəpoð]	*modrybedd*	'aunts' (569)
	c.	*[məd'rəpoð]		

1.3 Provection and neutralization

Provection is sometimes described as the neutralization of laryngeal contrast in stops, for instance by G. E. Jones (1988)[5] and Hannahs (2013). This is important to emphasize, because this view of provection underlies the one theoretical proposal for its analysis known to me, that by Hannahs (2013). He notes that the literature makes numerous references to fortis stops being

[4] This is the plural form given in the dictionary part of C. H. Thomas (1993); a provected form ['bakla] is also given in the text (vol. 1, p. 104). The nature of the variation in this particular item is not made explicit.

[5] G. E. Jones (2000) is more circumspect and speaks of 'alternation' only.

'geminated' or having increased duration after stressed vowels: see, for instance, Fynes-Clinton (1913), Sommerfelt (1925), Williams (1985) and Awbery (1986). It is also well-known that lengthening in obstruents is phonetically inimical to voicing, since the build-up of air in the supraglottal region reduces transglottal airflow (Ohala 1983), and there are typological examples of voiced geminates undergoing (at least partial) devoicing, for instance in Japanese (Hirose & Ashby 2007). Consequently, if stops are longer after stressed vowels, devoicing would not be an unreasonable consequence. Hence, Hannahs (2013) suggests, in dialects with provection lenis stops became lengthened after stressed vowels, merging with existing fortis stops, which are also lengthened in that position.

This account, however, is not tenable. According to existing descriptions provected stops are *not* lengthened in the way that lexically fortis stops are. This is connected to a phonotactic constraint in South Welsh that ties the quantity of the stressed vowel with the quality of the following consonant. Specifically, the following conditions obtain in stressed penultimate syllables in native vocabulary in South Welsh (Awbery [1984]; see also Wells [1979]):

- Before [b d g f θ χ v ð], vowels are long;
- Before [p t k s ʃ ɬ m ŋ w j], vowels are short;
- Before [n l r], vowels may be long or short, with a lexical distribution;

The exception from the pattern is the vowel [ə]: it is always short, which is a general restriction on the patterning of this segment in Welsh. In addition, stressed vowels in South Welsh are short before any sort of consonant cluster, even if it begins with a consonant that otherwise requires a preceding long vowel: ['ɛbrɪɬ] 'April', ['kɛvnɛ] 'backs'.

The nexus of the stressed vowel and the following consonant is characterized by a complementary distribution of length: descriptions agree that consonants are phonetically lengthened after short vowels (see e. g. Awbery 1986, C. H. Thomas 1993).[6] The following examples are from Awbery (1986); the use of the half-length mark is Awbery's.

(10) a. ['karˑɛg] *carreg* 'stone' (10)
b. ['amˑsɛr] *amser* 'time' (10)
c. ['kejˑnɔg] *ceiniog* 'penny' (16)

It is for this reason that provection is generally *not* a neutralizing phenomenon. As both Awbery (1984) and C. H. Thomas (1993) point out, vowels before provected stops remain long, and the stops themselves are short, just as expected after phonologically long vowels (example from Awbery [1984]):

(11) a. ['ɑːkor] *agor* 'open' (40)
b. *['akor]

The behaviour of provected stops with respect to vowel length is in contrast with another 'devoicing' phenomenon, one that is shared by all Welsh dialects that allow a vowel length

[6] Similar facts obtain in monosyllables (Grawunder, Asmus & Anderson 2015)

distinction in penultimate syllables (and, confusingly, sometimes *also* referred to as provection). This process is associated with certain suffixes, most prominently with the equative ([-ɛd]/[-ad]), comparative ([-aχ]), and superlative ([-a]) inflection in adjectives. These suffixes trigger a consonantal alternation that is, at first glance, identical to provection as used elsewhere in this paper, whereby [b d g] become [p t k]:

(12)	a.	['teːg]	*teg*	'fair' (788)
	b.	['tɛkːa]	*tecaf*	'fairest' (788)

In these cases the 'devoicing' of the consonant *is* accompanied by a change in the length and quality of the preceding vowel, precisely unlike the case of *caledfiad*. This process can reasonably be analysed as involving an underlying /h/ at the beginning of the suffixes: the sequence of a lenis stop and [h] coalesce to produce a fortis stop, which, in accordance with regular phonotactics, must be preceded by a short vowel. (See Iosad [2012b: §6.4.4.1] for a detailed analysis that shows why this initial /h/ fails to surface or coalesce when the preceding segment is not a stop.)[7]

There is another aspect in which neutralization in provection is not complete. The fortis stops of Welsh are in fact strongly aspirated in all prosodic positions (Ball 1984, Ball & Williams 2001). Provected stops, however, lack aspiration. S. E. Thomas (1983: p. 48) describes the situation as follows:

> Y mae'r seiniau [b d g] yn seiniau lenis, dianadlog, lleisiol, sef seiniau a gynhyrchir drwy ddefnyddio ynni a thyndra cyhyrau cymharol wan. Wrth galedu fe dry'r sain lenis, ddianadlog, leisiol, yn sain ffortis, ddianadlog, ddilais [...]
>
> The sounds [b d g] are lenis, unaspirated, and voiced: these sounds are produced with relatively weak energy and muscular tension. Provection turns a lenis, unaspirated, voiced sound into one that is fortis, unaspirated, and voiceless [...]

Further, the lack of neutralization between provected stops and lexically fortis ones is confirmed by the fact that the two can in fact be distinguished: sources are quite explicit about the possibility of provected forms undergoing change to assume a shape that renders them indistinguishable from forms containing a fortis stop underlyingly. Specifically, not only is the postvocalic stop itself described as voiceless, but also the stressed vowel is short and lax, as expected before fortis stops. Numerous examples are provided by Watkins (1967: p. 320) for Llansamlet, which I reproduce here without typographic changes. This neutralizing type of provection is distinct from the non-neutralizing type, which is also present in that variety.

[7] An alternative analysis of these facts might involve a floating mora associated to the suffix, which would dock to the consonant and enforce vowel shortening (along the lines of Trommer & Zimmermann 2014). However, this analysis is not tenable because this context does not neutralize quantity contrasts when these are not driven by consonant quality alone: ['gwan] *gwan* 'weak' (p. 430), ['gwanaχ] *gwannach* 'weaker' (vol. 1, part 2, p. 82) contrasting with ['glæːn] *glân* 'clean', [glɑːnaχ] *glanach* 'cleaner' (p. 403).

(13)

a.	['dɪkːɔn]	*digon*	'enough'
b.	['ɬɛtːi]	*lludw*	'ash'
c.	['gʊpːod]	*gwybod*	'to know'
d.	['prɪtːɛ]	*prydau*	'meals'

In these cases, we observe that the stressed vowel is short (and lax), just as in (12) above, and the consonant is long, as expected after a short vowel.

The phonetic study of provection in the Tawe Valley by Spooner (2016) confirms these descriptions. His results show that provected stops have shorter closure duration than fortis ones, and are only rarely, if ever, accompanied by voiceless preaspiration. Fortis stops following stressed vowels, on the other hand, are very frequently preaspirated. (See below section 4 and Morris & Hejná [2020] for more on preaspiration in Welsh.) This is also accompanied by a difference in the duration and quality of the preceding vowel, also in line with the existing descriptions.

The conclusion of this section, therefore, is that provection does not neutralize the contrast between fortis and lenis stops following a stressed vowel, and a different account of the phenomenon is required.

2 Synchronic analysis

In this section I set out a new proposal for the synchronic analysis of provection. I consider in particular its featural interpretation, its interaction with prosodic structure, and the stratal aspects of the phenomenon.

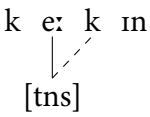
2.1 Provection as vowel-consonant interaction

I propose that provection is an unusual type of vowel-consonant interaction. It arises from a confluence of two factors. First, as discussed above, phonological vowel quantity in Welsh depends in non-trivial ways on the quality of the following consonants; in particular, long vowels normally precede (intervocalic) lenis stops [b d g]. Second, as I discuss below in section 3.2, vowel length in Welsh is consistently associated with a quality distinction, much as in languages like English and German: phonologically long vowels are also tense (see in particular G. E. Jones 1984). If the tense/lax distinction in vowels is phonologized, then the 'core' context for provection – an intervocalic lenis stop following the stressed vowel – can be expressed in terms of adjacency to a [tense] vowel. I propose, then, to identify the feature which encodes 'tense' quality in vowels with the feature associated with fortis stops. I shall call this feature [tense]. Provection arises from rightward spreading of this feature from the vowel to the consonant, resulting in a single token of [tense] linked to two root nodes. I argue that this phonological representation allows us to understand both the phonological contexts for provection and the phonetic realization of this phenomenon.

Consider a lexical item like *cegin* 'kitchen'. In the singular, the vowel of the penultimate syllable is stressed. This circumstance permits it to be long, since it also precedes a lenis stop /g/: ['keːgɪn]. The phonology of the language also requires long vowels to be tense; assume this

is formalized using a privative feature [tense]. This feature undergoes rightward spreading to the following stop to produce the following representation:

(14) Representation with provection



Importantly, the label attached to the feature shared by the vowel and the consonant is essentially arbitrary. The major claim here is that the phonological grammar of the relevant dialects identifies the feature that distinguishes 'tense' and 'lax' versions of stressed vowels with the feature that distinguishes the two series of stops. In most widely used featural theories, of course, the two are treated separately: the former is the vocalic feature [(±)tense] or [(±)ATR], and the latter is a feature like [(±)voice] or [(±)spread glottis]. In this paper I adopt an emergentist (Mielke 2008, 2013, Mielke, Magloughlin & Hume 2010), substance-free (e. g. Morén 2006, 2007, Youssef 2010) conception, where features have no intrinsic phonetic content. Instead, they are identified by the learner on the basis of phonological patterns they participate in, particularly patterns of contrast and morphophonological alternation (Dresher 2014: on this aspect of the approach, see).

The proposal adopted here relies on three premises:

- The distinctive feature in the stop series in Welsh is, in traditional terms, [spread glottis]: 'voiceless' stops are more marked phonologically and 'voiced' ones;
- There is a phonological contrast between two series of vowels: a 'lax' series [ɪ ʊ ɛ ɔ ə a] and a 'tense' series [i u e o ɑ], which interacts with vowel length;
- The lack of neutralization with lexical fortis stops that we find in provection is a (typologically unremarkable) consequence of how doubly linked [tense] is realized in Welsh.

In the following sections I shall argue that all these conditions are in fact met. Before doing so, I present the phonological analysis in detail. As indicated above, I will use the label [tense] to refer both to the vocalic feature distinguishing pairs such as [i] ∼ [ɪ] and to the feature normally referred to as [spread glottis]. I will also use a privative version of this feature, although I make no claims here about the (in)appropriateness of a binary-feature analysis.[8] I will also add length marks for explicitness, even though C. H. Thomas (1993) treats them as unnecessary: [tense] stressed vowels are always long in the Nantgarw dialect.

[8] A reviewer asks whether a privative approach to the 'tenseness' ([±ATR]) distinction is tenable given the evidence for the activity of both [+ATR] and [−ATR] cross-linguistically, and the existence of within-language ternary distinctions (e. g. Kim 2002). I cannot engage with this issue at length here, except to note that under an emergent-feature régime cross-linguistic comparability of featural specification cannot be assured. As for within-language ternary contrasts, see Avery (1996), Ghini (2001), Iosad (2017a) and Sandstedt (2018) for proposals that allow them while keeping the featural system privative.

2.2 Background: foot structure in South Welsh

Before turning to the analysis of provection proper, I will briefly lay out the necessary background around the phonology of consonant and vowel length that will be crucial to the analysis below. As noted in section 1.3, the length of stressed vowels in South Welsh native vocabulary depends on the vowel's segmental context. We can distinguish three classes of contexts (Awbery 1984, Hannahs 2013):

- Stressed vowels are always long before single [b d g f θ χ v ð], before another vowel (i. e. in hiatus), and before single [s ʃ ɬ] if the stressed syllable is word-final;
- Stressed vowels are always short before single [p t k m ŋ], before any consonant cluster, and before single [s ʃ ɬ] if the stressed syllable is penultimate;
- Stressed vowels can be either short or long before [n l r], depending on the lexical item.

In Iosad (2012b: §§6.4.5, 8.2) I offer an extensive OT analysis of the pattern in a south-western (Pembrokeshire) dialect of Welsh, to which I refer the reader for more details. The basic idea (shared by Hannahs 2013) is that those consonants that are preceded by short vowels are moraic and those that are preceded by long vowels are not. The source of the mora can be lexical (as in the case of [n l r]) – what Morén (2001) calls 'distinctive weight' – or the mora can be introduced by the phonological computation (Morén's 'coerced weight'). We need not be concerned here with the precise mechanism by which this weight coercion comes about, but I will assume that consonants preceded by short stressed vowels are moraic on the surface.

The basic pattern is open syllable lengthening: a stressed vowel is long unless it surfaces in a closed syllable. A segment that is normally non-moraic, such as [g], will not block the lengthening of a preceding stressed vowel ([oːgɔv] *ogof* 'cave'), whilst a moraic one, such as [k], will block it ([tɛka] *tecaf* 'fairest'). This analysis is consistent not just with theoretical premises, but also with the phonetic 'half-length' facts reported in section 1.3.

The dialect of Nantgarw described by C. H. Thomas (1993) is different from the description of Pembrokeshire Welsh by Awbery (1986), which was analysed in Iosad (2012b). C. H. Thomas (1993) makes it clear that there are many exceptions to these generalizations. The literature often notes the deviation from these patterns of lengthening and shortening in borrowings from English such as *strôc* 'stroke' with a long vowel before a fortis stop, or *job* 'job' with a short vowel before a lenis one, but the examples given are usually monosyllables. According to C. H. Thomas (1993), however, a non-negligible number of short vowels in stressed penultimate syllables before (unprovected) lenis stops are also attested in Nantgarw:

(15)	a.	['kɔbɪn]	*cobyn*	'cob, pony' (192)
	b.	['kwabad]	*cwabad*	'strike with wet object' (239)
	c.	['rɪgaχ]	*rhigach*	'tease' (684)
	d.	['stɛgɪn]	*stegyn*	'short, stumpy person' (713)
	e.	['d͡ʒɔbɪn]	*jobyn*	'task, job' (339)

C. H. Thomas (1993) counts 97 such instances against 149 examples of unprovected stops after long vowels such as ['paːbɪð] *Pabydd* 'Catholic'. Note that the latter pattern is what is regular

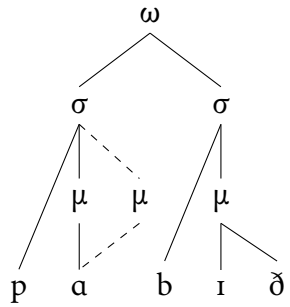
in southern dialects without provection such as Pembrokeshire, but for provecting dialects such as Nantgarw this pattern requires additional explanation, which I provide below in section 2.5. I take these numbers to indicate that in Nantgarw the weight of lenis stops is also distinctive rather than coerced: an input /kɔb⋆μɪn/ with a moraic lenis stop surfaces unchanged (rather than with a long vowel due to weight coercion, as assumed in Iosad [2012b] for Pembrokeshire Welsh).

Similarly, a few words (mostly borrowings) are attested with a long vowel before a fortis stop, a pattern not found in native vocabulary; it is not discussed whether their realization is identical to that of provected stops.

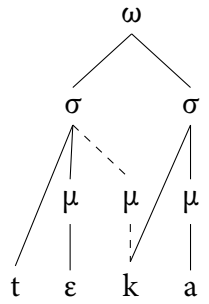
(16) a. ['poːkar] *pocer* 'poker' (vol. 1, p. 69)
b. ['eːko] *eco* 'echo' (vol. 1, p. 69)

With all this in mind, I shall assume the following representations for stressed vowels and their following consonants. Penultimate stressed syllables (the focus of our discussion) are always heavy. They can contain a long vowel (in which case the stressed syllable is open) or a short vowel followed by a moraic consonant. The moraic consonant can be an ambisyllabic geminate (when it is a singleton; this is also proposed by Hannahs [2013]) or a coda within a cluster. The options are laid out in (17), with dashed lines showing when the moraic association of interest is inserted by rule.

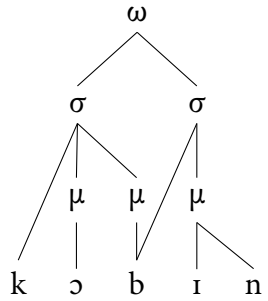
(17) a. Lengthening before nonmoraic consonant: ['pɑːbɪð] *Pabydd* 'Catholic' (p. 621)



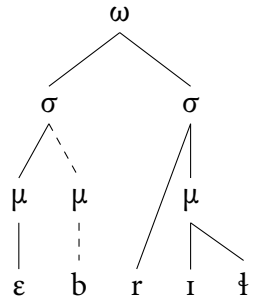
b. No lengthening before moraic consonant, moraicity by coercion: ['tɛka] *tecaf* 'fairest'



c. No lengthening before moraic consonant, underlying moraicity: ['kɔbɪn] *cobyn* 'cob'[9]

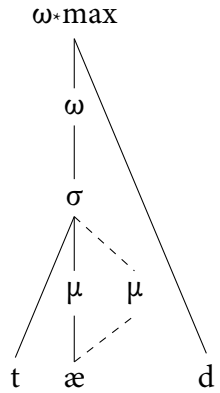


d. No lengthening before moraic consonant, moraicity by position: ['ɛbrɪɬ] *Ebrill* 'April'



The analysis for monosyllables is much the same: vowels do not lengthen when the following consonant is moraic (whether underlyingly or by coercion) and they do when the next consonant is nonmoraic (again, see Hannahs [2013] for a similar approach). As I discuss in more detail in Iosad (2012b), the patterns in South Welsh can be accounted for if we assume that unless a single word-final consonant is moraic, it is extrametrical, in the sense that it does not form part of the stressed syllable. As surveyed by Vaux & Wolfe (2009), there are numerous options in the literature for representing such peripheral 'appendix' segments. For concreteness, I will assume that they are adjoined to the prosodic word via recursion. The options are shown in (18)

(18) a. Final nonmoraic consonant: ['tæːd] *tad* 'father' (p. 783)

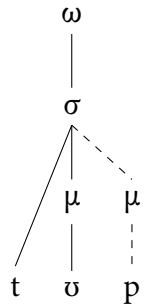


[9]A similar representation has to be posited for words like ['ɬədan] *llydan* 'wide', where the consonant is preferentially nonmoraic, but this preference is overridden by the unviolated ban on long [əː].

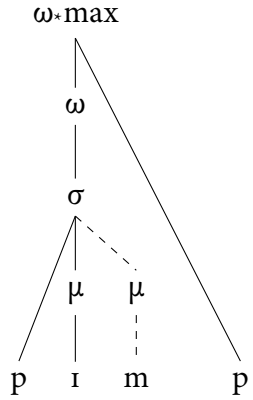
Height	Front	Central	Back
High	iː ɪ	(ɨː ɨ)	uː ʊ
Mid	eː ɛ	ə	oː ɔ
Low		a aː/ɑː	

Table 1: The vowel system of Welsh

b. Final moraic consonant, moraicity by coercion: ['tʊp] *twp* 'stupid' (p. 826)



c. Consonant cluster: extrametricality and moraicity by position: ['pɪmp] *pump* 'five' (vol. 1, part 2, p. 116)



2.3 Background: vowel quantity and quality in Welsh

Another necessary prerequisite is a discussion of the status of [tense] in Welsh vowels outside the context of provection (for more detailed discussion, see Iosad [2017b]). The monophthong system of Welsh is shown in table 1. All Welsh dialects distinguish short and long version of the monophthongs [i u e o a]. Most varieties also have the vowel [ə], which cannot be phonologically long (Hannahs 2013). The distribution of the high central vowel [ɨ(ː)] is limited to northern varieties (Wmffre 2013), and I do not consider it further in this paper. In some varieties (including Nantgarw) the low long vowel can be realized as [æː] (in some varieties also as a centralizing diphthong [eə] or similar) in a final syllable in some lexical items; I ignore this phenomenon here, but see Rees (2015) for some discussion.

As table 1 shows, non-low vowels demonstrate covariation of length and quality of a sort familiar from languages such as English and German: vowels that are long are also less centralized, i. e. 'tense' (see e. g. A. R. Thomas 1966, G. E. Jones 1984). Mayr & Davies (2011) report an acoustic study of both North and South Welsh speakers that confirms the existence of both quantitative and qualitative contrasts, at least in monosyllables. In some varieties, including Nantgarw (C. H. Thomas 1993), descriptions report a phonetic difference between the 'short' and 'long' allophones of /a/, with the 'long' one transcribed as [ɑ] (described as a central vowel, albeit slightly backed compared to 'short' [a]). In this paper, I use [ɑː] when quoting from C. H. Thomas (1993).

Long vowels in Welsh are restricted to stressed syllables – in fact, as we saw in the preceding section, they are restricted to *open* stressed syllables. Since long vowels are also obligatorily tense (in most varieties), vowel tenseness (in stressed syllables) is particularly associated with contexts where the postvocalic consonant does not acquire a mora. In unstressed syllables, the situation is less clear (again, see Iosad [2017b] for more sustained discussion), but the sources generally agree than the distribution of tense and lax vowels in that position is contextually predictable: for instance, it is generally acknowledged that all vowels are tense in hiatus even when unstressed, and many sources state that high vowels are obligatorily tense in final open syllables.[10] Given this lack of contrast between tense and lax vowels in that position, it is not at all clear that phonological [tense] specifications are present in unstressed syllables in South Welsh. In stressed position, however, we can be reasonably certain that, if [tense] is to be represented in the phonology at all, it is present in contexts requiring a long vowel – including, crucially, before [b d g].

2.4 Provection: the core case

The 'core' case of provection is that of intervocalic stops. According to the statistics for incidence of provection by context given by C. H. Thomas (1993), overall stops are more frequently provected than not, but stops in consonant clusters resist it more strongly: in fact, within the cluster category there are (marginally) more non-provected stops than provected ones. The counts are shown in fig. 1.

These numbers suggest that it is the intervocalic stops that require a consistent account first and foremost. Intervocalic position is also important because it is precisely in this context that stressed vowels are preferentially long before [b d g] in South Welsh, as discussed above. I suggested that this preferential association of provection with the context for lengthening is not arbitrary: if length is associated with tenseness, and tenseness is identified with the feature used to implement laryngeal contrast in stops, then provection can be understood as a type of vowel-consonant interaction.

[10] In Nantgarw, the situation in post-tonic syllables can be summarized thus (C. H. Thomas 1993: vol. 1, part 1, pp. 13–20): for high vowels, the distribution where tense vowels are found word-finally and lax vowels are found in closed syllables holds (with some exceptions that I cannot discuss here for reasons of space). In the mid vowels, [ɛ] is not found in this position due to a sound change where it lowers to merge with [a] in final unstressed syllables. The back mid vowel [o] is said to be very close in quality in that position, to the point of occasional merger with (lax!) [ʊ]. The low vowel [a] in final unstressed syllables is described as identical to short (i. e. lax) stressed [a].

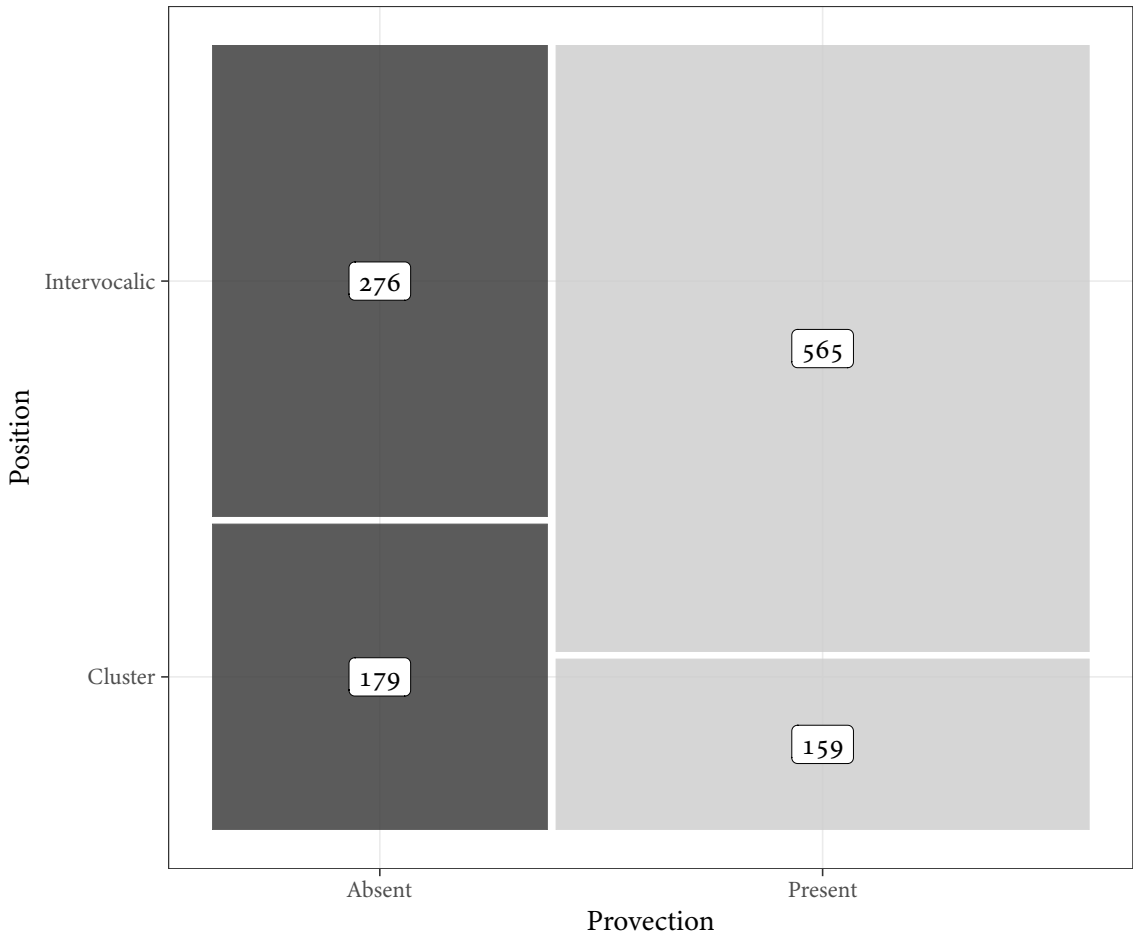


Figure 1: Type frequency of provection in intervocalic stops

In provection-undergoing items such as ['keːkɪn] *cegin* 'kitchen', I propose that the vowel [e] is lexically [tense]: /kegin/. In the unsuffixed singular form, this vowel is stressed and becomes long. This length licenses the presence of a surface [tense] feature, which also spreads rightwards to the stop. In this respect, this vowel is distinct from a non-provecting stressed vowel in a word such as ['niːdɪr] *neidr* 'snake', which is [tense] on the surface but not underlyingly: in such items, the [tense] specification is assigned to stressed vowels predictably, as part of the phonological computation, but this assignment stands in a counterfeeding relation to provection (section 2.5).

In the suffixed form [kɛ'giːna] *ceginau* 'kitchen', the vowel in the initial syllable is lexically [tense], but provection is absent, at least as a regular outcome. I suggest this is connected to the fact that the quantity distinction in Welsh is neutralized in unstressed syllables in favour of short vowels. Given the tight connection between quantity and quality evident under stress, it is plausible that a [tense] contrast should also be neutralized in this position, as discussed in section 2.3 above.

A basic premise of the analysis is that provection occurs if and only if a vowel that is specified for [tense] is followed by a lenis stop. This is highly plausible in stressed position, where vowels before [b d ɡ] are long and hence [tense]. That all long vowels are tense is an exceptionless generalization in South Welsh phonology, and hence I assume it is enforced by the grammar.

As for unstressed position, as noted in section 2.3, it is not clear that [tense] specifications are at all present in unstressed vowels, since they are all phonologically short. Under the privative assumptions made here, the simplest way of enforcing such a neutralization is to posit that an underlyingly [tense] vowel loses this feature and the surface vowel lacks a [tense] specification. If this is so, then the [ɛ] in [kɛ'giːna] is not expected to trigger provection. This corollary is not just an analytical convenience: we might want to extract some predictions regarding the phonetic realization of these vowels. There are at least two possible analyses:

1. First, we may wish to connect the lack of a phonological [tense] specification in the output of phonology with phonetic variability (Keating [1988] with much subsequent literature; for recent discussion, see Strycharczuk [2012]). This makes the testable prediction that the quality of unstressed vowels should exhibit continuous variation, or possibly stochastic choice between categorical variants (Strycharczuk & Simon 2013, Strycharczuk et al. 2014);

2. Another possibility leverages the pattern of contrast: if [tense] is a privative feature distinguishing the first member of a pair such as [e] ∼ [ɛ] from the second, then the deletion of [tense] creates a representation identical to that of the lax vowel, with the prediction that unstressed vowels should be phonetically indistinguishable from stressed lax ones.

As discussed in Iosad (2017b), further empirical study is needed to shed sufficient light on the phonetic realization of unstressed vowels in Welsh. Impressionistic descriptions generally agree that there is significant 'free variation'. It is usually described in terms of alternating categorical variants, although this is probably due to the non-use of instrumental methods. It is generally agreed (e. g. G. E. Jones 1984) that the relevant factors are position vis-à-vis stress (i. e. pretonic vowels behave differently from post-tonic ones) and the presence of a coda (in particular, vowels in open syllables tend to be tense and those in closed syllables tend to be lax).

Some preliminary results reported in Iosad (2017b) indicate that phonological [tense] or similar specifications *may* in fact be present in at least some unstressed syllables, such as those in hiatus or (in some cases) word-final ones. However, even if a [tense] specification can be posited for such vowels, it appears that they are always *not* followed by a consonant, meaning provection would be impossible.

A potential issue for any such account in a privative feature framework is that it forces us to accept differential phonetic implementation of a [tense]-less vowel in stressed and unstressed syllables: such vowels should be lax when stressed but may have a different realization when unstressed. However, this can be accommodated if we accept that the range of permitted phonetic variation in a given position can be (at least partly) determined by the set of contrasts available in this position. Concretely, Welsh stressed syllables may contain both tense and lax vowels, and therefore a [tense]-less representation will be realized with less variation to ensure the implementation of the contrast; in unstressed syllables, where the contrast is absent, there is significantly more leeway for phonetic variability due to surface underspecification; see Dyck (1995) and Dresher (2009) for discussion of similar cases in Romance.[11] This question can only be resolved by further empirical study.

To sum up, I conclude it is plausible that the restriction of provection to the position after stressed vowels is due to the impossibility of a surface [tense] in (relevant) unstressed positions; effectively a type of vowel reduction.

2.5 Non-provection: intervocalic stops

If provection is triggered by the lexical presence of a [tense] feature in the preceding vowel, then an explanation is required for the existence of unprovected forms such as ['pɑːbɪð] *Pabydd* 'Catholic', which have a surface [tense] vowel with a following lenis stop. I suggest that these are best treated as instances of opaque underapplication in a stratal model of morphology-phonology interaction (e. g. Kiparsky 2000, Bermúdez-Otero 2011, 2012).

Specifically, the vowel in these items lacks the [tense] feature underlyingly, so we can represent the word 'Catholic' as /pabɪð/, with a lax /a/ vowel. In a stratal model (Kiparsky 2000, Bermúdez-Otero 2018), phonological computation proceeds in several cycles, with at a distinction between the stem level, the word level, and the postlexical level.

Processes relevant to provection – namely the spreading of [tense] and vowel reduction, i. e. the deletion of [tense] in (some) unstressed syllables – clearly belong to the word level, because the distribution of provection-triggering [tense] can be exceptionlessly derived from the placement of stress. Specifically, neither provection nor stress assignment can belong to the postlexical levels: the former never crosses word boundaries, and stress in Welsh is clearly computed at the level of the inflectionally complete word, falling on the penultimate (or, in certain well-defined circumstances, on the final) syllable of such words. There is also no evidence that either process is restricted to the stem level;[12] hence, the transparency of the interaction follows from the generally transparent nature of within-stratum interactions.

[11] Another possibility is viewing the contrast as a ternary one, so that only the weak (unstressed) position exhibits true surface underspecification, while lax stressed vowels have a representation distinct from both [tense] and unspecified segments. See Iosad (2012a,b) for concrete implementations of such a ternary contrast.

[12] See Iosad (2012b: §6.4.5.3) for discussion of cyclic prosodification in Welsh.

Another process that belongs to the word level is the determination of vowel length on the basis of the quality (and possibly underlying moraicity) of the following consonant, as discussed in section 1.3. Any lengthening or shortening of the vowel must happen at least at the word level, since it is transparently related to stress placement. No such shortening or lengthening takes place at the phrase level, since the restrictions on vowel quantity do not hold across word boundaries: word concatenation counterfeeds length adjustments, so a long vowel in a word like ['tiː] 'house' (*tŷ*) is not shortened before a fortis stop in a word like ['teːg] 'fair' (*teg*): ['tiː 'teːg], not ['ti 'teːg].

Thus, the constraints on vowel quantity identified in 2.2 do not hold at the postlexical level, after word concatenation. On the other hand, the requirement that long stressed vowels are tense and short stressed vowels are lax is exceptionless in the output of the phonology. This could be either because it belongs to the postlexical stratum, or it is enforced at an earlier level and not affected by any postlexical process. I suggest that the former is correct, because it allows us to account for cases of non-provection (as in ['pɑːbɪð] *Pabydd* 'Catholic') in terms of counterfeeding opacity. Specifically, underlying /pabɪð/ undergoes stress placement and vowel lengthening at the word level to produce /'paːbɪð/, where provection does not occur for lack of a [tense] feature on the stressed vowel. This form is the input to phrase-level phonology, where [tense] is inserted to satisfy the requirement that all phonologically long vowels bear this feature: the output is ['pɑːbɪð], with a surface tense vowel but without provection.

Note that the feature [tense] is implicated in two separate interactions with other features. Prior to the output of the word level, it is used to mark provection-triggering vowels. Since ultimately whether an item shows provection is lexically specific, it has to be stored underlyingly, and since I treat provection as assimilation rather than deletion of the feature, it is the triggers rather than the non-undergoers of the process that need to bear the feature underlyingly. However, in surface representations the distribution of [tense] is entirely predictable, because it undergoes two rounds of grammatical computation that enforce this: first, [tense] is restricted to stressed syllables (at the word level), and then it is further restricted to *long* vowels in stressed syllables. In this regard, the analysis is a classic example of how counterbleeding opacity provides a cue for underlying contrasts that are otherwise obscured by the derivation (cf. Vaux 2008).

2.6 Non-provection: monosyllables

One strong generalization regarding the distribution of provection is that it is not found in monosyllables. This is generally true even if provection does occur in polysyllabic forms of the same item, leading to alternations:

(19)	a.	['briːg]	*brig*	'summit' (98)
	b.	['briːka]	*brigau*	'summits' (98)

Provection in the plural confirms that the underlying form is /brig/ with a [tense] vowel, yet provection fails in the singular. This suggests that a phonological mechanism prevents the spread of [tense] to a word-final consonant. I propose that this is due to final-consonant extrametricality, as discussed in section 2.2: a singleton word-final consonant in a Welsh monosyllable lies outside that minimal prosodic domain which includes the nucleus, and spreading cannot cross the

intervening boundary. If spreading of [tense] is constrained not to cross the boundaries of a minimal prosodic word,[13] then provection will generally be blocked in monosyllables but will occur in polysyllables, where the domain of [tense] does not have to cross a minimal word boundary.

One exception from the generalization that provection is disallowed in monosyllables is described by Flowers (1994) for Crynant, where it is found in 2nd person singular imperatives of a few verbs such as ['riːt] 'run!'. I would suggest that in these cases we are dealing not with spreading across a boundary: instead, these verbs are lexicalized with the double link of the [tense] feature already present underlyingly, which prevents the blocking of spreading. I will return to the lexicalization of provection in more detail below in section 4.4.

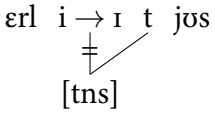
2.7 Neutralizing provection: consonant clusters

This opaque interaction of provection and the assignment of vowel quality on the basis of length also provides an account of provection in consonant clusters. As discussed in section 1.2, provection affects not just intervocalic lenis stops but also stops preceding other consonants (usually sonorants). This can lead to alternations:

(20)	a.	['ɛrlɪd]	*erlid*	'persecute, ill-treat' (351)
	b.	[ɛr'lɪ**t**jʊs]	*erlidiodd*	'(s)he persecuted'[14] (351)
	c.	[ɛr'lɪ**t**jɔg]	*erlidiog*	'prone to persecution' (351)

Here, provection is not visibly triggered by a tense vowel: there is over- rather than underapplication. Nevertheless, the sequence of rules outlined in the previous section correctly predicts this relationship as well. Assume that the underlying form of *erlid* is /ɛrlid/, with a [tense] /i/. In a form like /ɛrlidjʊs/, the [tense] feature spreads rightwards, but the vowel does not lengthen because of the following cluster: the word-level output is /ɛrlitjʊs/ with transparent provection. At the phrase level, the short vowel is stripped of the feature [tense] in line with the general pattern in stressed syllables, producing the overapplication. In autosegmental terms, this involves the delinking of the feature [tense] from the vowel but not from the following consonant.

(21) Provection and vowel laxing



Therefore, the double link of [tense] is impossible in the case of a stressed vowel followed by a consonant cluster: provection results in a representation for the vowel and the following stop that is identical to the representation of a short vowel followed by a *lexical* fortis stop. In this

[13] For instance, using CrispEdge constraints (Itô & Mester 1999) or Contain (Bickmore 2000).

[14] The 3rd person singular preterite suffix in Nantgarw is [ʊs], as in much of South Wales (as opposed to northern and standard *-odd*); this is historically related to the suffix *-wys* attested in Middle Welsh (Simon Evans 1964: §133.(a).4).

Level	Process	/ˈbr**i**ga/	/ˈpabɪð/	/ˈkɔb⋆µɪn/	/ˈɛrl**i**d/	/ɛrˈl**i**djɔg/
Word	Quantity adjustment	ˈbr**iː**ga	ˈpaːbɪð			
	Provection	ˈbr**iː**k**a**				erˈl**it**jɔg
Postlexical	Short ⇒ lax				ˈɛrlɪd	ɛrˈlɪ**t**jɔg
	Long ⇒ tense		ˈp**aː**bɪð			
Output		[ˈbr**iː**k**a**]	[ˈp**aː**bɪð]	[ˈkɔbˑɪn]	[ˈɛrˑlɪd]	[ɛrˈlɪ**t**ˑjɔg]

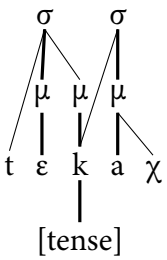
Table 2: Analysis of provection

case, therefore, provection is predicted to be neutralizing: i. e. the [t] should be (pre)aspirated. I am not aware of phonetic data that would allow us to confirm this prediction, but it should be empirically testable.

2.8 Stratal analysis: summary

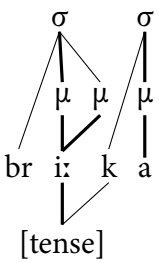
The stratal analysis of provection is summarized in table 2. For ease of reference, boldface shows the domain of the [tense] feature at each stage. The counterfeeding order between tensing (in words like /pabɪð/, surface [ˈpɑːbɪð]) and provection (in words like /briga/, surface [ˈbriːka]) is accounted for via the ordering of strata. The within-stratum mappings, as predicted by Stratal OT, are (largely) transparent, although the interaction between vowel lengthening and provection deserves some comment. As discussed in section 2.2, vowel lengthening in stressed syllables is blocked if the stressed vowel is followed by a fortis stop: [ˈtɛkaχ] 'fairer', *[ˈteːkaχ]. If this is formalized in terms of a constraint ranking that facilitates [tense] stops acquiring a mora (as in Iosad [2012b]), then it is not immediately clear why a stop that becomes [tense] by provection does not show the same behaviour in terms of moraicity. I suggest that the difference lies in the status of the stop within the [tense] domain (e g. Cassimjee & Kisseberth 1998, Jurgec 2010). When the stop is associated with the [tense] feature underlyingly, it is said to *head* the domain of [tense], and it is reasonable that constraints on moraicity (essentially on headship in moraic domains) should refer to the properties of the head of a featural domain. This is shown in (22), where thick lines indicate headship in a domain.

(22)



In (22), the fortis stop heads a [tense] domain, and thus is targeted by pro-moraicity constraints. By contrast, if the stop acquires [tense] by spreading, it does not head the [tense] domain, and moraicity constraints remain unviolated. This is shown in (23).

(23)



Thus, the mapping on the word level envisaged in table 2 can be interpreted as a transparent one.[15] I conclude, therefore, that the view of provection as a type of spreading of [tense] from the stressed vowel to the following stop can be reconciled with the data within a stratal framework.

3 Supporting evidence and discussion

In this section I discuss the evidence for two ancillary hypotheses on which the analysis of provection given above relies, and highlight some further consequences of that analysis for phonological representations and the nature of variation found with provection. In the next two subsections I focus on two propositions that have not received sufficient attention. First, I review the evidence for the greater markedness of fortis stops in Welsh, which is necessary, in a privative framework, to be able to confidently assign a feature to that series of stops. Second, I consider whether it is necessary to assign phonological 'tenseness' specifications to Welsh vowels given that they are largely predictable from factors such as vowel quantity and syllable structure.

3.1 Laryngeal contrast in Welsh

The laryngeal phonology of Welsh is of a type characteristic of languages that the literature on 'laryngeal realism' (e. g. Iverson & Salmons 1995, Honeybone 2005, 2012, Beckman, Jessen & Ringen 2013) identifies as an 'aspirating language'. In such languages, obstruent series traditionally transcribed as /p t k/ vs. /b d g/ are said to be distinguished by the presence of a feature (variously referred to as [(+)spread glottis], [fortis], or, with reference to Element Theory, |H|), which is absent in the latter.

Here, I adduce *phonological* evidence to this effect. The phonological grammar of the language contains rules that must make reference to [spread glottis] specifications, confirming that this feature is required. Lenis stops, on the other hand, demonstrate the traditional hallmarks of unmarkedness: they undergo phonological processes involving laryngeal specifications but do not trigger them. In the framework of 'laryngeal realism', as in other approaches based on privative features, this lack of phonological activity is accounted for via the assumption that unspecified features are completely absent.

[15] Another possibility would be to assume that fortis stops block vowel lengthening because they are underlyingly moraic rather than because of a phonological pattern where they acquire a mora because of their lexical specification (cf. Carlyle [1988] for Breton). However, this approach is poorly equipped to deal with shortening before derived fortis stops, discussed in section 1.3; and see in particular footnote 7.

The activity of [spread glottis] in Welsh phonology can be observed in two separate processes. One of them was briefly discussed above, namely the 'devoicing' of stops in contexts such as adjective inflection and the denominative suffix [(h)ai]:

(24)	a.	['teːg]	*teg*	'fair'
	b.	['tɛkːa]	*tecaf*	'fairest'
(25)	a.	['gwæːg]	*gwag*	'empty' (438)
	b.	[gwa'kai]	*gwacáu*	'to empty' (426)

This type of alternation is straightforward to analyse as the addition of a [spread glottis] feature to the final stop. The alternative analysis, assuming a distinctive [voice] feature, would have to treat this as a type of subtractive morphology, in contrast to the concatenative nature of the [spread glottis] analysis.[16]

The feature [spread glottis] is also active in consonant assimilation. A good example is the abstract-noun suffix *-der*. Underlyingly, it starts with a lenis initial stop:

(26)	a.	['tɛw]	*tew*	'fat' (787)
	b.	['tɛwdar]	*tewder*	'fatness' (787)
(27)	a.	['trʊm]	*trwm*	'heavy' (819)
	b.	['trəmdar]	*trymder*	'heaviness' (vol. 1, part 2, p. 28)

When following a [spread glottis] segment, whether a stop or a fricative, it is the [spread glottis] segment that acts as the trigger of the assimilation, rather than the lenis stop:

(28)	a.	['kɔiθ]	*coeth*	'cheeky' (197)
	b.	['kɔiθtar]	*coethder*	'boldness'[17] (197)
	c.	*['kɔiðdar]		

Within the literature on 'laryngeal realism', it is common to adduce phonetic evidence in favour of classifying languages as 'aspirating' or 'voicing' (Beckman, Jessen & Ringen 2013: see especially). Although the Welsh evidence is also consistent with this approach, it is both in conflict with substance-free theoretical commitments and empirically problematic; for discussion of the latter, see Iosad (2017a: chap. 11), Kirby & Ladd (2019) and Salmons (2020). For this reason, I do not focus on these issues here.

[16] In Iosad (2012b), I analyse this as not simply the addition of a feature but as coalescence with a [h] segment. The strongest support for this analysis is provided by the fact that the denominative suffix, which triggers the alternation in *gwacáu*, actually surfaces with a [h] in other lexical items (*tewháu* 'to fatten; to get fat', *cryfháu* 'to strengthen', *llwfrháu* 'to lose heart'). For a dialect like Nantgarw, the feature may instead have to be analysed as floating, since [h] is essentially impossible in surface representations in many southern varieties.

[17] C. H. Thomas (1993) transcribes forms such as these with the symbols for lenis stops (['kɒiθdar]); however, this is due to the fact that such stops have short VOT, and I assume that in line with similar systems in other languages the stops in such clusters are in fact phonologically [tense]. I return to this issue below in section 4.1. In the present context, the important datum is that /θ/ does not undergo assimilation to [ð], even though both segments are present in the language otherwise.

3.2 The phonological status of tenseness

In the foregoing discussion, especially that dedicated to foot structure in section 2.2, I assumed that vowel quantity is active in the phonological grammar of Welsh. However, as I also discussed in section 2.4, vowel tenseness is predictable from vowel quantity and other factors such as stress and syllable structure (leaving provection aside for the moment). This raises the question of whether, given its predictability (i. e. 'non-contrastive' status), tenseness should at all be represented in the phonology.

As I discuss in more detail in Iosad (2017b), there is significant disagreement in the literature on this point. With a few exceptions, scholars have usually chosen to treat either quantity or quality as phonologically relevant, but only rarely both. For instance, A. R. Thomas (1966), when describing the dialect of Dyffryn Alyn in North Wales, recognizes that long and short vowels normally differ in quality as well: [eː] ≠ [ɛ], [uː] ≠ [ʊ]. However, he argues that length must be represented in the phonology, because there is a phonemic contrast between [a] and [aː], with no quality difference, and thus quality is redundant. Conversely, authors such as G. E. Jones (1984) and C. H. Thomas (1993) only see quality as phonemic, and assume that the length of tense vowels in stressed positions is not phonological.

In many varieties of Welsh, the situation indeed appears to be ambiguous. However, there are also varieties where tense/lax quality must definitely be represented in the phonology. The clearest case is found in south-western dialects, discussed in most detail by Awbery (1986), Wmffre (2003) and Iosad (2017b). Here, long mid vowels occur as both tense [eː oː] and lax [ɛː ɔː]. The former appear in stressed monosyllables and in stressed penultimate syllables before a non-high vowel, while the latter are found in stressed penultimate syllables before a high vowel:

(29)	a.	['kɔːdi]	*codi*	'rise'
	b.	['koːdɔð]	*cododd*	'(s)he rose'
(30)	a.	['treː]	*tref*	'town'
	b.	['trɛːvɪð]	*trefydd*	'towns'

In Iosad (2017b), I show that the distinction between tense and lax vowels in Welsh is both categorical and phonological, and offer a featural analysis of both south-western and 'classical' South Welsh patterns.

The difference between the two patterns demonstrates the potential for differential phonologization paths. In the classical pattern, as discussed at the start of this section, a phonological distinction between tense and lax qualities might exist across all vowels (other than [ə]), but it is relatively inert: its distribution is conditioned solely by metrical structure (specifically length and syllable structure). There is thus little *phonological* evidence for its precise nature. Conversely, in the south-western varieties, at least for the mid vowels, the contrast interacts with the phonological specification for [(±)high] vowels, which strongly suggests that the tense-lax contrast in mid vowels is phonologically interpreted as one of height, so that mid vowels are able to participate in a height dissimilation pattern (Iosad 2017b). Below in section 4 I suggest that provection is a result of another possible phonologization trajectory.

3.3 Vowel quality and schwa

One property of provection that has not been discussed in the analysis above is that it can be triggered by vowels of all qualities. It follows that if provection is triggered by a [tense] feature, then all vowels presumably have distinct [tense] and non-[tense] counterparts. This is relatively unproblematic in the case of peripheral non-low vowels, since it is generally agreed that phonetically, if not necessarily phonologically, Welsh varieties generally distinguish between [i] and [ɪ], [e] and [ɛ] and so on (section 3.2). The two remaining vowels for (South) Welsh are [a] and [ə]. In the case of [a], C. H. Thomas (1993) is quite explicit about the existence of a qualitative distinction between what she writes as [a] and [ɑ], with the latter patterning with other 'tense' ('long') vowels (for other references to this possibility, see Watkins [1967] and G. E. Jones [1984]). Thus, although some varieties appear to neutralize the qualitative distinction between short and long [a] (e. g. Awbery 1986, Mayr & Davies 2011), it is at least possible that the distinction between [tense] and non-[tense] varieties of [a] can be made.

No such distinction is described in the literature for [ə]. This vowel has a number of idiosyncratic properties in Welsh (Hannahs 2013: §4.1). In particular, it appears to lack a phonologically long counterpart: it cannot appear in content words that consist of a single open syllable or in an open stressed penult – both contexts where other vowels are obligatorily long. It also cannot appear in hiatus, where other vowels are normally described as tense (even if they are short).

I suggest, nevertheless, that this does not necessarily disqualify [ə] from supporting a [tense] contrast that would allow some instances of [ə] to trigger provection. In this connection, it might be worth noting that although [ə] does not have a 'long' counterpart analogous to those found for other vowels, there are occasional references in the literature (Awbery 1986, C. H. Thomas 1993) to the possibility of a 'half-long' variety of [ə], possibly with a phonologically conditioned distribution. It is also relevant that if [ə] triggers provection, it still remains short; hence, the following consonant must be a geminate, and the [ə] is still found in a closed syllable (section 2.2). We may thus surmise that 'tense [ə]', even if it exists, should be restricted to closed syllables, explaining why it is not found in hiatus or when long.

All this makes the question of whether there can be two phonologically distinct varieties of [ə] in Welsh worth further study.

3.4 Phonological variation

The account of phonological provection given in section 2 is, of course, considerably neater than the descriptions given in the literature imply. As noted above, it is widely recognized that provection is widely variable across varieties (particularly in its lexical coverage [C. H. Thomas 1975, Thorne 1984, G. E. Jones 2000]) and within varieties, due to social factors (S. E. Thomas 1983, B. Thomas 1990, Flowers 1994, M. C. Jones 1998, Brake 2011).

In the analysis proposed here, provection or lack of it is reduced to the presence of a phonologically [tense] vowel in the relevant lexical item. This approach makes it simple to capture cross-dialectal variation, such as the fact that the word *bugail* 'shepherd' is found with provection (['biːkal]) in Rhigos (near Aberdare) but with no provection (['biːgal]) in Nantgarw (C. H. Thomas 1993: vol. 1, part 1, p. 103): this can be derived if we assume that in Rhigos the

underlying representation is /bigal/ with a [tense] vowel and in Nantgarw it is [bɪgal] with a non-[tense] one.

This approach is less successful with cases where what appears to be the same morpheme shows variants with and without provection. Here, it is useful to distinguish between two cases: variation across lexical items and cross-speaker variation.

3.4.1 Lexical variation

Under widely accepted morphological assumptions, most morphemes have a single underlying representation that is fed into the phonology, with various opinions regarding the extent to which morphological information is accessed by the phonological component. If this is true in all cases, then the analysis of provection offered here faces problems in dealing with data such as the following:

(31)	a.	['ɛrgɪd]	*ergyd*	'shot, blow' (351)
	b.	[ɛr'gɪtjɔn]	*ergydion*	'shots, blows' (351)
	c.	[ɛr'gɪdjo]	*ergydio*	'to strike' (351)

The verb *ergydio* transparently contains the same morpheme as the noun *ergyd*. Provection in the noun plural *ergydion* would indicate that the morpheme is underlyingly /ɛrgid/ with tense [i]. However, provection is absent in the verbal form, which is problematic for this analysis.

In some cases like this, it is possible to appeal to multiple underlying forms, thereby making differences in a provection behaviour a consequence of allomorphy. For instance, the particular case of *ergyd* vs. *ergydio* could be analysed in terms of the dual-route theory of lexical storage proposed by Bermúdez-Otero (2012, 2013). He suggests that (at least in some languages) the minimal unit of lexical storage is not the root but the stem, which comprises the root and morphological material bearing categorial information.[18] Under this régime it is, in principle, viable to analyse the part-of-speech split seen in (31) if the nominal stem is lexically ⟦ɛrgid⟧⋆N and the verbal stem is ⟦ɛrgɪd⟧⋆V. Under this analysis, the lexical items *ergyd* and *ergydio* do *not* contain the same morpheme, and the paradox does not arise. It can, in principle, be ascertained if all such inconsistencies are amenable to this analysis; however, this type of variation must necessarily be supplemented by an account of within-speaker variation.

3.4.2 Within-speaker variation

Another source of variation that is not trivial to accommodate within the phonological account of provection is the concurrent use of provected and non-provected forms of the same word (rather than the same morpheme in different morphological contexts, as above) by the same speaker. In this context, it is unfortunate that many monograph-length dialect descriptions do not dedicate much attention to this type of variation: to the extent it is described, it is normally

[18] Myler (2015) argues that Bermúdez-Otero's (2013) stem-storage analysis of Spanish diphthongization is not necessary (at least if one accepts some specific tenets of Distributed Morphology). Myler's aim, as far as I can see, is not to disprove the existence of stem storage; in any case the important point here is the availability of phonological differences being introduced with reference to part of speech, whether by stem storage or by DM's morphosyntactically sensitive rules. (See also Bermúdez-Otero [2016] for a critique of Myler's argument.)

referred to as 'free', and at least some scholars consciously aim to present an idealized, 'pure' version of the dialect (Wmffre [2012] specifically makes this point about C. H. Thomas [1993]). This makes it difficult to be sure of the extent of variation based on such descriptions. Even so, many of these descriptions do note the widespread existence of 'doublets', i. e. provected and unprovected forms of the same word in the dialect. A particularly explicit description is provided by Flowers (1994), who states directly that provected and unprovected forms may be used by the same speaker within a single register (he discusses particularly the case of free conversation two older speakers).

Most variationist descriptions of provection, however, are more focused on the incidence of provection rather than on any lexical factors. The picture that arises from these studies is one of a recessive feature associated with more local, less standard, and more old-fashioned forms of speech: in fact, in many cases younger and more educated speakers do not use provection at all.

Accommodating these findings within the present account of provection requires an explicit theory of variation, which I cannot hope to provide in this paper. I suggest that at least some of this variation can be accommodated within a competition-based framework. Under such an approach, provected and unprovected forms of the same item exist in the lexicon (e. g. both /brig/ and /brɪg/ for 'peak'). The speaker can then make a stochastic choice at the point of lexical insertion. The factors influencing this choice may be multifarious, both grammatical (e. g. part-of-speech affiliation, as suggested above) and non-grammatical, such as, for instance, indexical information or the social setting (for one model implementing a similar idea, see Adger & Smith [2010]).

With immediate respect to provection we might hypothesize that forms where [tense] is present in the underlying representation (such as /brig/) are associated with those external factors that determine the recessive behaviour of provection: for instance they will be more likely to be chosen in situations where lower-register forms are more acceptable. Once such an underlying representation is used, the underlying [tense] feature will be able to trigger provection.

In the absence of [tense] in underlying representation, the outputs will have the entirely predictable distribution of vowel tenseness driven by vowel quantity and syllable structure that we see in other varieties of Welsh. Such underlying representations may be more likely to be chosen in contexts such as higher registers.

In sum, I suggest that the attested variation in the use of provection can be made compatible with the analysis presented in this paper.

4 Origins of provection

In this section I offer a proposal for the diachronic origins of provection that is consistent with the position taken here that phonological representations are substance-free and emerge from noisy data. I suggest that provection arose due to a stochastic misparsing of lenis stops as fortis stops with short preaspiration. I show that in non-provecting varieties of South Welsh shorter preaspiration duration is associated with relative phonological lowness, and argue that it is this connection that makes the reanalysis plausible. The stochastic nature of the misparsing also

explains the lexically-diffusing rather than Neogrammarian nature of provection as a sound change.

4.1 Vowel-consonant interactions

In section 3.2 I discussed south-western Welsh, where 'tenseness' appears to be phonologized independently of vowel quantity, at least for some vowels. I argued that since it interacts with the phonological distinction between high and non-high vowels, it is represented in the phonology using the same feature as that used to make the high/non-high distinction. Here, I propose that in south-eastern varieties of provection arose as the result of a different process of phonologization: specifically, all 'tense' vowels in this variety are assigned a phonological feature to distinguish them from their 'lax' counterparts, and this feature is identified not with what would traditionally be called [(±)high], but rather with [(±)spread glottis], the feature used for laryngeal contrast in stops. This identification allows 'vocalic' [tense] to interact with the laryngeal specification of a following stop.

Welsh is similar to familiar Germanic languages like English, German, and Icelandic, in that laryngeal contrast in stops is neutralized after (tautosyllabic) fricatives: there is no contrast between tautosyllabic [sp] and [sb], or [ɬt] and [ɬd].[19] As with these familiar examples, the proper phonological interpretation of the outcome of this neutralization is disputed for Welsh. Based on the short VOT shown by stops in such clusters, it is common for phonemic accounts (such as C. H. Thomas [1993]) to use the fortis stop symbols [b d g] for these segments; other sources, notably Awbery (1986), treat them as instances of fortis [p t k].

It is commonly accepted that the short VOT in fricative-stop clusters in Icelandic or English does not stem from the lack of a glottal opening gesture: instead, if the clusters form a single domain for the [spread glottis] feature, then the peak of the opening is timed to the middle of the clusters rather than to the burst as with singleton stops; this produces the short-VOT effect (see e. g. Magnus Pétursson [1978], Kingston [1990] and Iverson & Salmons [1995]; but see also the critical comment by Ahmed, Andersson & Vaux [2020]). It is plausible that in Welsh the same short-VOT effect can be associated with a double link of the [spread glottis]/[tense] feature. In the case of provection, this aspect of the phonetic implementation accounts for the fact, noted among others by S. E. Thomas (1983, 1988) and Spooner (2016), that provected stops, unlike lexically fortis ones, are unaspirated, even though the are also associated with glottal spreading (which is consistent with their lack of voicing).

Utilizing the full possibilities of the autosegmental formalism allows us to account for the phonetic implementation facts. Another, perhaps even more important advantage of this move is that it gives us an account for a slightly unusual asymmetry in the phonology of south-east Welsh dialects. In most positions, for instance in non-branching syllable onsets, Welsh contrasts only two series of stops: 'fortis' and 'lenis'. Just in post-tonic position, however, we have to deal with

[19] I am not aware of Welsh cases similar to potentially contrasting English pairs like *disperse* vs. *disburse* of *Kresge* vs. *pesky* (which are in any case problematic only if we assume the relevant clusters are not separated by some prosodic boundary or do not contain underlying voiced fricatives); thanks to a reviewer for bringing up this important qualification.

a ternary contrast: provected stops are distinct from both lenis and fortis ones.[20] Introducing a different featural specification just to allow for this eventuality appears to be unwarranted. In particular, Spooner (2016) finds that provected stops pattern with lenis unprovected ones in all phonetic respects except the presence of voicing and/or (pre)aspiration: that is, duration of the stop closure, quality of the preceding vowel and duration of the preceding vowel. (Descriptively, Spooner [2016] notes a difference between provected and lenis stops in the duration of the vowel, but he shows that this difference loses statistical significance once a general effect of consonant voicing on preceding vowel duration is accounted for.) All this supports the analysis of provected stops as identical to lenis ones in their suprasegmental parse (nonmoraic and preceded by long, tense vowels) and in some of their featural specification (place and manner features), but distinct from them in terms of laryngeal specification (the presence of the feature [tense]).[21]

4.2 Bottom-up phonologization

The rise of provection as attested in existing descriptions appears to be a non-Neogrammarian type of sound change. The lexical specificity of its application indicates that in many communities it can be the result of lexical diffusion rather than regular sound change. This is a reasonable hypothesis in view of the geographic spread, where lexical coverage of provection dwindles gradually towards peripheral areas (C. H. Thomas 1975, Thorne 1984). In this section I argue that the item-by-item application of provection also sheds light on its origins as a phonological process.

I suggest that provection arose initially (at least in part) as a listener-driven misparsing (e. g. Ohala 1981, Stevens & Harrington 2013) of the stressed vowel-stop nexus occasioned by the absence of preaspiration in fortis stops. The low reliability of this cue predicts the instability of provection, both in lexical terms and with respect to its further diachronic development, namely the frequent neutralization of provected and lexically fortis stops.

The presence of preaspiration in word-medial lexically fortis stops is rarely remarked upon in the literature: for instance, it is not marked even in the phonetically detailed transcription in Hannahs (2013: p. 14), where length (often ignored in the descriptions) and postaspiration are present (['attʰɛb] 'answer'), and only cursorily referred to by Ball & Williams (2001). As noted above, however, more recently the use of preaspiration has been noted by Morris (2010) and Morris & Hejná (2020), who focus on Northern Welsh. In this section, I examine the conditioning of preaspiration in a non-provecting variety of South Welsh, in order to examine the potential role of preaspiration as a precursor to provection. I use the dataset referred to above (Iosad 2019), which contains 2767 tokens of disyllabic words with penultimate stress from

[20] Note that this kind of ternarity is distinct from the kind of underlyingly ternary contrast that the computation collapses into a binary one in surface representations, like the laryngeal contrast of Turkish as analysed by Inkelas (1994) and Reiss (2003).

[21] A remaining question here is whether the double link of [tense] is in fact important: one could perhaps assume simply that the unaspirated [tense] stops in provected ['keːkɪn] *cegin* 'kitchen' are distinct from aspirated [tense] stops in fortis ['tʰɛkʰːa] *tecaf* 'fairest' only because the latter is moraic. Verifying this requires a better understanding of phonetic variation in the realization of fortis stops both in provected contexts and in the handful of cases where a nonmoraic fortis stop is not due to provection (i. e. in borrowings like *pocer* 'poker').

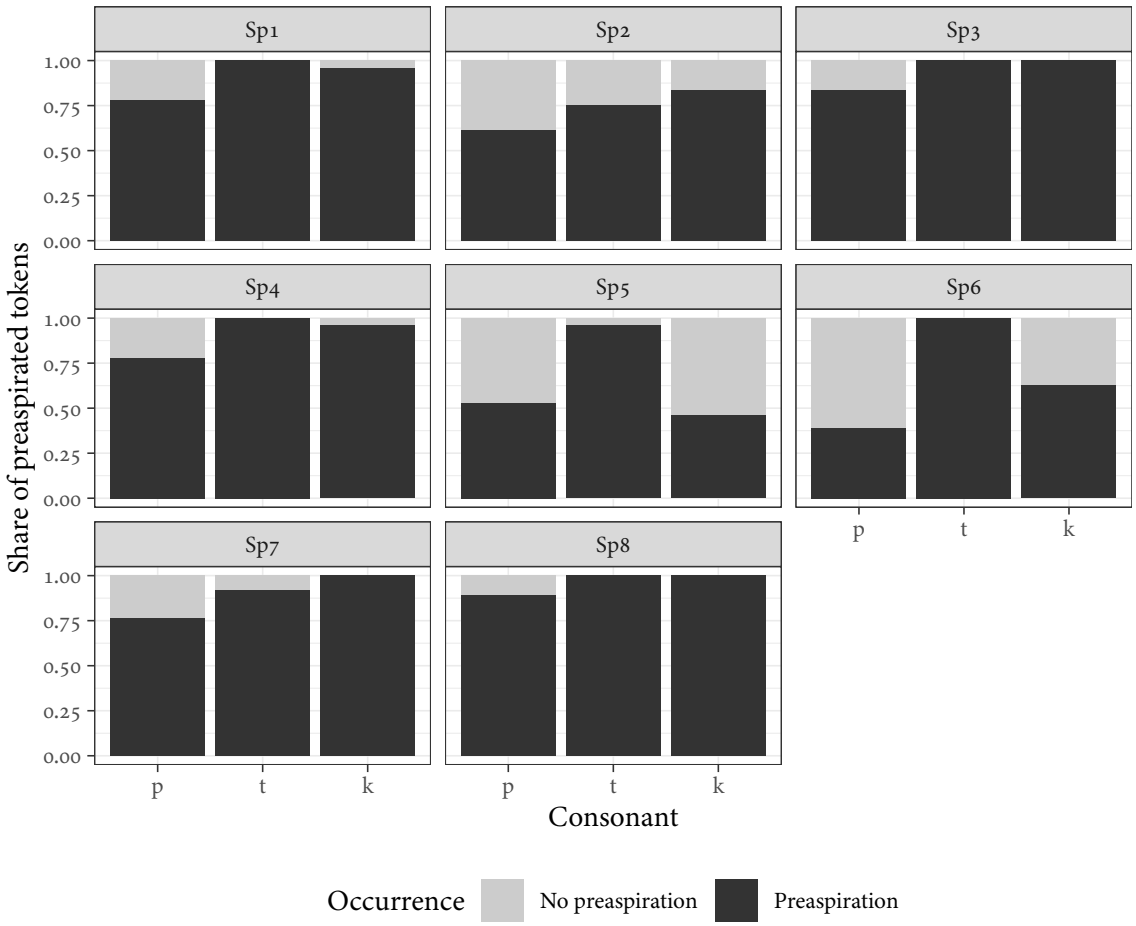


Figure 2: Occurrence of preaspiration in stops by speaker

8 speakers. The test words includes all stressed vowels of South Welsh except [a], controlled for phonological context, including manner, place, laryngeal specification, and moraicity of the following consonant (and hence, by implication, the length of the stressed vowels), and high vs. non-high specification of the vowel in the unstressed final syllable. I refer to the online supplementary material to Iosad (2017b) for demographic and technical information.

Figure 2 shows the rates of preaspiration *occurrence* before fortis stops in these data. A token was coded as having preaspiration if it showed *voiceless* preaspiration. As Hejná (2015) discusses, voiceless preaspiration is frequently accompanied by a preceding period of breathy voice (which is frequently included into the duration of the preaspiration by phoneticians), and in some cases this breathiness may cue the fortis status of a following consonant even in the absence of voiceless preaspiration. I cannot consider the question of the relationship between breathiness and voiceless preaspiration here in detail, in particular since so little is known at this stage about the status of breathiness in provection. The consideration of voiceless preaspiration (henceforth simply 'preaspiration') is nevertheless instructive.

It is readily apparent that preaspiration is pervasive in these varieties: it is much more common for a fortis stop to have than not to have it, to the extent where for some speakers it is all but categorically present at least in some stops. There are also clear trends with respect to place asymmetries: for all but one speaker, preaspiration is most frequent for [t], less frequent for [k], and least frequent for [p].

Given the high rate of occurrence, it is a reasonable supposition that preaspiration might be an important cue for the fortis/lenis contrast. Moreover, given that provected stops are described as unaspirated, its absence may be the hallmark of provection. Provection then may be described as a manipulation of preaspiration sensitive to the phonological specification of the preceding vowel: a provected stop is one that is phonologically fortis (it is associated with a [tense] feature) but does not have (voiceless) preaspiration. Thus, if provection, as I hypothesize, is driven by a preceding phonologically [tense] vowel, then [tense] vowels must provide a plausible context for a shortening or disappearance of preaspiration.

To test whether phonological factors influence the *occurrence* of preaspiration, we fit a variety of generalized additive mixed models with a binomial response function; model fitting was conducted in R (R Core Team 2017) using the `mgcv` package (Wood 2006). The data is limited to the six speakers in the study who show the vowel system specific to south-western varieties of Welsh described in section 3.2.

First, we attempt to model the *frequency* of preaspiration. Speaker and lexical item were treated as random intercepts. Table 3 shows the estimated values of model that includes closure duration, vowel category, and stop place of articulation as predictors. The table shows estimated effect sizes and 95% confidence intervals.[22] The numbers confirm the impression that can be drawn from fig. 2 that coronals are much more strongly associated with preaspiration than non-coronals (whereas the difference between labials and dorsals does not reach statistical significance). However, there are no significant differences among the vowels in predicting whether preaspiration occurs in the following stop. In fact, including vowel quality as a predictor does not appear to significantly improve the model fit: a likelihood ratio test for the second model in table 3 produces $\chi^2(4, 2) = 1.78, p = 0.78$ for vowel quality as a predictor. The phonetic properties of the vowel also do not seem to matter: adding predictors such as vowel duration and F1 and F2 values (or smoothing them together) does not improve model fit. The occurrence of preaspiration as such is thus not sensitive to phonetic or phonological properties of preceding vowels. Note, however, that phonological quantity could not be controlled for, as fortis stops (the only ones that can be preaspirated) are preceded by phonologically short vowels except in borrowings, which were excluded from the sample.

A different picture emerges from the modelling of the *duration* of preaspiration. We exclude tokens with no voiceless preaspiration at all (with a voice offset time of 0; for relevant discussion, see Hejná [2019]) and fit a variety of generalized additive mixed models. Initial modelling with duration of preaspiration as dependent variable showed a skew in the normal quantile-quantile plots, which was rectified by using $\log_2$ of the duration to account for the fact that duration is frequently log-normally distributed (Rosen 2005). Dependent parameters were smoothed

[22] The confidence intervals given for the smooth effects (closure duration) and random effects are less meaningful than p-values here; since our focus is on parametric effects like vowel quality, I do not dwell on these effects in detail.

	Model without vowel effect	Model with vowel effect
Intercept	1.03	0.49
	[−0.70; 2.76]	[−2.11; 3.10]
Coronal stop	2.37*	2.53*
	[1.51; 3.23]	[0.84; 4.23]
Dorsal stop	0.86*	0.77
	[0.19; 1.54]	[−0.74; 2.28]
Closure duration smooth	2.00	2.00
	[−1.92; 5.92]	[−1.92; 5.92]
Speaker random effect	4.68	4.77
	[−5.12; 14.48]	[−5.03; 14.57]
Preceding [e]		0.55
		[−1.91; 3.01]
Preceding [i]		0.85
		[−1.25; 2.95]
Preceding [o]		1.19
		[−0.60; 2.98]
Preceding [u]		0.79
		[−0.95; 2.53]
Word random effect		9.46
		[−19.94; 38.86]
AIC	293.84	282.03
BIC	332.30	374.31
Log Likelihood	−137.24	−117.78
Deviance	274.47	235.55
Deviance explained	0.27	0.37
R^2	0.27	0.36
GCV score	−0.25	−0.28
Num. obs.	392	392

* Null hypothesis value outside the confidence interval.

Table 3: Models for preaspiration occurrence, showing 95% confidence intervals

	No vowel effect	Fixed vowel effect	Random vowel effect
Intercept	-5.06^*	-5.43^*	-5.06^*
	$[-5.45; -4.67]$	$[-5.91; -4.94]$	$[-5.45; -4.68]$
Closure duration	-2.37^*	-2.12^*	-2.22^*
	$[-3.77; -0.98]$	$[-3.48; -0.77]$	$[-3.53; -0.91]$
Coronal stop	0.51^*	0.49^*	0.48^*
	$[0.09; 0.93]$	$[0.19; 0.78]$	$[0.06; 0.90]$
Dorsal stop	0.30	0.27	0.28
	$[-0.12; 0.72]$	$[-0.02; 0.57]$	$[-0.15; 0.70]$
Preceding /i/		0.05	
		$[-0.41; 0.50]$	
Preceding /e/		0.47^*	
		$[0.05; 0.90]$	
Preceding /o/		0.77^*	
		$[0.36; 1.18]$	
Preceding /u/		0.15	
		$[-0.26; 0.56]$	
AIC	573.59	568.80	580.46
BIC	599.95	610.21	659.53
Log Likelihood	-279.80	-273.40	-269.23
Number of observations	319	319	319
Number of groups: word	22	22	22
Number of groups: speaker	6	6	6

* Null hypothesis value outside the confidence interval.

Table 4: Models for preaspiration duration, showing 95% confidence intervals

closure duration (with a thin plate regression spline basis), consonant place, and category of preceding vowel. In these models, smoothed closure duration was significant, with effective degrees of freedom equalling 1 – in other words, the model was equivalent to a linear model. Hence, further models were fit using the package `lme4` (Bates et al. 2015).

The results show that the category of the preceding vowel exerts some influence on the duration of preaspiration (along with closure duration and consonant place, just as with preaspiration occurrence): a likelihood ratio test for the two models produces $\chi^2() = 24.14, p < 0.001$ (as table 4 shows, the model including the vowel effects also has a smaller AIC value). However, to exclude the possibility that this influence is an artefact of within-speaker variation, we fit models with vowel category as random slope by speaker; table 4 shows such a model with no random intercept by speaker, which is not superior to the model with vowel as fixed effect. Thus, I conclude that there is indeed a fixed effect of preceding vowel category on the duration of the voiceless preaspiration.

The model was run using treatment contrasts, with the effect of a preceding [ə] vowel as zero, and effects showing the estimated difference from the effect of [ə]. The results show that the effect of the high vowels [ɪ ʊ] is not significantly different from the effect of [ə], with small estimates of the effect and the bounds of the confidence intervals well either side of zero. With the mid vowels, [ɔ] clearly has a significant effect: on average, preaspiration following [ɔ] is 1.71 times longer than following [ə], and the 95% confidence interval excludes zero. As for [ɛ], while the effect is smaller (preaspiration after [ɛ] is on average 1.39 times longer than after [ə]), it is still larger than that of [ɪ] and [ʊ], and again zero falls outside the confidence interval. Hence, it is fairly clear from the effect sizes that the high vowels [ɪ ʊ] and the mid vowels [ɛ ɔ] pattern differently in terms of their influence on the duration of preaspiration in a following consonant. This is illustrated in fig. 3, which shows the estimated effects and 95% confidence intervals as durations of preaspiration after the respective vowels compared to [ə], with the dashed line indicating no effect.

I interpret these results as indicating that the vowel effect on preaspiration duration is *phonological* rather than *phonetic*: it is both categorical and sensitive to the phonological status of [ɛ ɔ] as mid vowels rather than to their phonetic quality. This suggestion is further reinforced by the fact that [ə], phonetically a mid vowel, patterns with the high vowels [ɪ ʊ]. This is not an entirely unexpected result. Typologically, the vowel [ə] often shows behaviour more characteristic of high vowels than of mid ones, notably in sonority-sensitive stress and tone patterns (Kenstowicz 1997). In Welsh itself, [ə] participates in alternations with high vowels ([i] in south Welsh and [ɨ] in North Welsh, as well as [u]; see Hannahs [2013] for details and relevant references), while there is little if any evidence for its patterning with mid vowels in the phonological grammar; see also Iosad (2017b) for more evidence that [ə] can pattern together with [ɪ ʊ] in Welsh.

If the driving factor in the effect on preaspiration duration is phonological, then phonologically *lower* vowels [ɛ ɔ] are associated with *longer* preaspiration durations. Although in the variety studied here there is no provection effect, I take these results to confirm that (relative) *phonological* lowness may be associated with longer preaspiration times.[23]

[23] Notably, Hejná (2015) identifies a similar effect in Aberystwyth English.

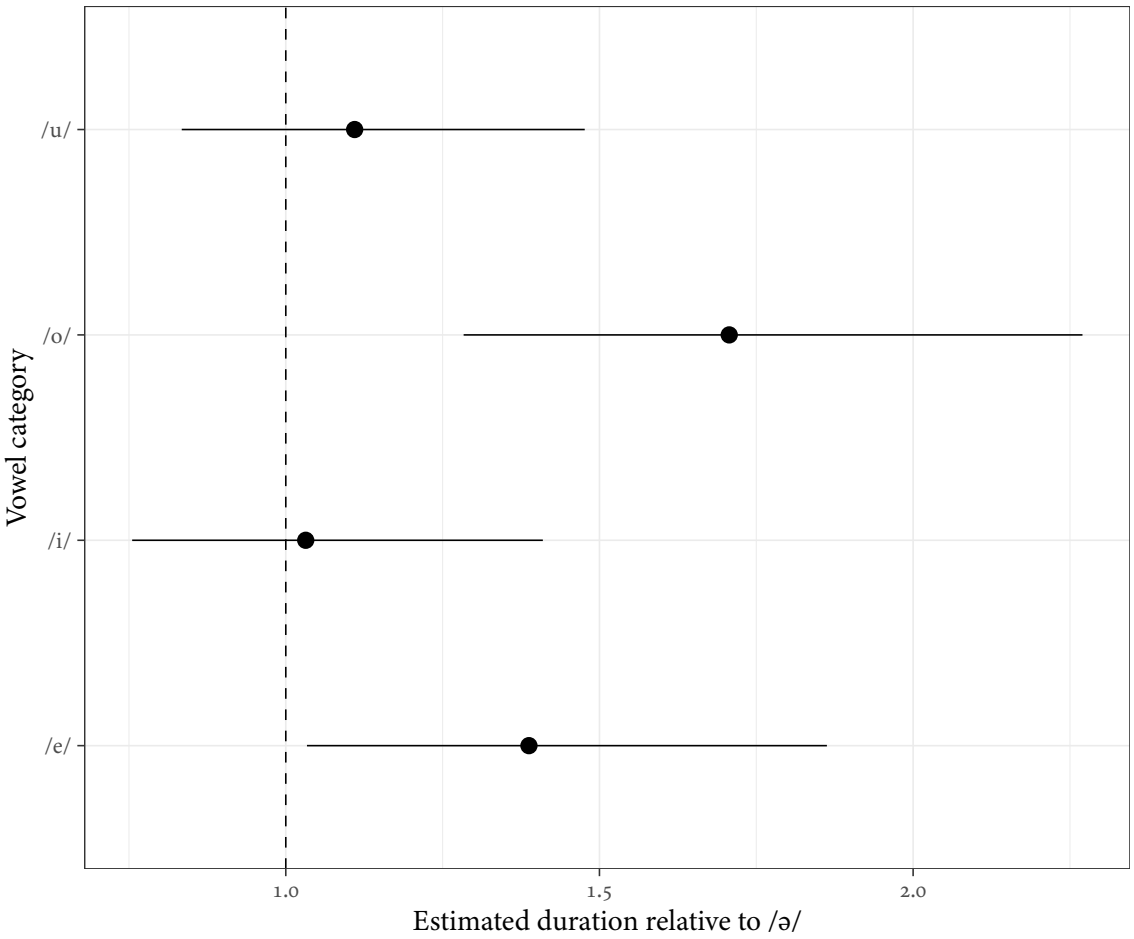


Figure 3: Estimated effects of vowel category on preaspiration duration

This, I suggest, is the key to the origin of provection: once speakers treat the tense-lax distinction within vowel pairs as phonological, they may interpret the amount of preaspiration following a stressed vowel as being conditioned by the [tense] specification of the vowel. This, of course, crucially relies on the possibility that 'height' (i. e. the distinction between [i u] and [e o] classes of vowels, and potentially also [a]) can also distinguish between the 'tense'/'lax' vowel pairs within a height category. I suggested in section 3.2 above that this would not be unprecedented in Welsh.[24] It would also be desirable to know what the behaviour of preaspiration would be after low vowels: logically, we could expect either a three-way distinction (such that preaspiration after [a] would be significantly longer than after mid vowels) or a two-way distinction between high vowels (including, phonologically, [ə]) and non-high ones. However, there are no low stressed vowels in the dataset used here, so this must be left to future work.[25] In any case, the precise import of the two different outcomes here would depend in very large measure on the featural specifications assumed for the vowels (see Iosad [2017b] for more explicit discussion), so I cannot discuss it here in too much detail.

Nevertheless, if the assumptions laid out above are accepted, then the following scenario can be envisaged. An unprovected form like ['briːga] 'peaks' contains a lenis stop, which has no preaspiration, but, like other Welsh lenis stops, may not necessarily be fully voiced. The lack of voiceless preaspiration is, of course, originally conditioned by the lenis status of the stop. However, given the [tense] specification of the preceding [iː], a listener may reinterpret the lack of preaspiration as a special case of *shorter* preaspiration conditioned by the relative highness of the tense member of the [iː]/[ɪ] contrast. The short VOT is also consistent with a doubly linked [tense] feature, similar to fricative-stop clusters. Hence, the learner may hypercorrectly assume that the speaker intended the [g] in ['briːga] as a *fortis* stop.

This process is by necessity stochastic. In particular, it depends on the amount of postvocalic voicing in lenis stops to be plausibly small for the listener to be able to categorize the tokens as sufficiently ambiguous between the fortis and lenis categories. Since the amount of closure voicing itself is also variable, we do not necessarily expect the reinterpretation to happen across the board. Instead, it depends on assigning the feature [tense] to vowels in those items that happen to demonstrate sufficient ambiguity. This categorization establishes the link to individual lexical items, since a particular morpheme ends up containing either a [tense] vowel (which therefore triggers provection) or a non-[tense] vowel.

If this reconstruction of the genesis of provection is correct, it can be read as an argument for the emergent nature of phonological features. Phonologization proceeds 'from below', as listeners identify categorical distributions in the ambient data and form hypotheses regarding the number and extent of phonological categories. The precise labelling of these categories is determined not by their phonetic substance (as in classic innate-feature theory) but by the overall structure of the distinctions and by the dynamic patterns found in the data. For instance, in section 3.2 I argued that the 'tenseness' distinction in south-western varieties of Welsh is conceptualized in the same terms as vowel aperture (see Iosad [2017b] for details), while in

[24] To make the argument fully convincing, it would be desirable to show a similar *direct* effect of 'tenseness' on preaspiration, without the mediation of tenseness by quantity. However, given that long vowels before fortis stops are absent from the Welsh native lexicon, being found only in a small number of borrowings, this is difficult to do.

[25] Hejná (2015) also does not include low vowels in comparing preaspiration duration after vowels of different heights.

provecting dialects essentially the same vowel contrast has come to be associated with the laryngeal feature used in stops. Thus, the precise featural specification ends up being emergent and dependent on language-specific phonological patterning.

4.3 The life cycle of provection

The genesis of provection as analysed here presents a case study in the life cycle of phonological processes (e. g. Kiparsky 1995, Bermúdez-Otero 2007, 2015, Bermúdez-Otero & Trousdale 2012, Ramsammy 2015). The phonologization of [tense] proceeds as expected, from a cognitively controlled gradient ('phonetic') rule, plausibly motivated by enhancement (Stevens & Keyser 2010, Keyser & Stevens 2006, D. C. Hall 2011) of the original length distinction to a phonological process. The phonological pattern enforcing [tense] specifications dependent on vowel quantity must be phrase-level, as discussed in section 2.5. Provection represents the ascent of [tense]-manipulating rules to the word level, and this ascent may potentially demonstrate a type of phonological change that we may call 'relabelling'.

As discussed above, numerous Welsh varieties show evidence of categorical, quantity-driven distribution of vowel quality that we have identified as potentially diagnostic of phonologization. The distinction often remains inert: the data put no significant constraints on how the categories should be labelled. As we saw in section 3.2, there are several responses to this indeterminacy attested across Welsh dialects. In south-western dialects, the qualitative distinction is clearly phonologized for the mid vowels, and interacts with the height specification of other vowels. This suggests that the distinction is formalized using the same feature as that used for encoding height. In south-eastern Welsh, by contrast, the qualitative distinction covers all vowels and is formalized using what is ordinarily a laryngeal feature. Potentially, therefore, the rule's ascent to the word level involves a change in how a distinction is interpreted phonologically, without necessarily altering its phonetic expression.

This possibility is inherent in a framework where the construction of the representational system proceeds bottom-up. The learner identifies relatively sharp distinctions in the data and operates with the hypothesis that they may correspond to distinctions in symbolic phonological representation. With this assumption in place, these phonological categories will be labelled in a way that is connected to their patterning in the grammar. This process may be sensitive to complementarity of distributions (see Peperkamp et al. [2006] for an explicit proposal and K. C. Hall [2013] for an overview of statistically gradient 'contrastiveness'), but the complementarity criterion is not necessarily decisive. For instance, Dillon, Dunbar & Idsardi (2012) implement a bottom-up learning model that is able to take into account phonological information during the construction of categories. In this model, there is a bias towards collapsing predictably distributed phonetic distinctions into single 'phonemic' categories, but it also allows for such distinctions to be assigned different phonemic categorization (see also Dunbar, Dillon & Idsardi 2013).

If correct, the scenario proposed here demonstrates the malleability of phonetic conditioning for phonological change. It would be extremely unwise to deny that 'functional', phonetic factors *can* play a significant, perhaps sometimes decisive, role in phonological change (e. g. Garrett & Johnson 2013). It does not automatically follow, however, that *all* phonological change should be

of this nature, or that all phonological change should necessarily be triggered by some phonetic phenomenon.

4.4 Provection and lexicalization

As discussed above, the proposed phonologization scenario involved in provection involves statistically unreliable, item-by-item extraction of phonological patterns from the data. The process is therefore prone to variability (understood at least as lack of Neogrammarian, across-the-board regularity) and potentially to misanalysis. This potential arises from the ambiguity of the signal and from the need to recover underlying representations along with surface ones. In particular, as suggested in section 2.6 above, it is conceivable that provected lexical items may be lexicalized not with an underlying [tense] feature on the vowel that undergoes rightward spread but with the double link of the [tense] feature already in place.

It is predicted that such underlying representations can show provection in contexts where it is normally not licensed, which is indeed attested in south-eastern Welsh, in particular after unstressed vowels:

(32) a. [ˈkoːtʊm] *codwm* ‘fall’ (p. 206)
b. [kɔˈtəma] *codymau* ‘falls’ (p. 206)

Assume this item is stored underlyingly as /k{ot}⋆[tense]ʊm/. In the unsuffixed form, the doubly linked [tense] structure surfaces faithfully (modulo vowel lengthening). When the vowel is unstressed, [tense] is delinked from the vowel but preserved by the consonant, as in (21).

This account, however, is more plausible in cases such as (32), where there is explicit evidence for the possibility of a [tense] specification on the vowel, than in cases such as the following:

(9, repeated) a. [ˈmɔtrɪb] *modryb* ‘aunt’
b. [məˈtrəpɔð] *modrybedd* ‘aunts’
c. *[mədˈrəpɔð]

Here, the vowel is never tense on the surface: it always precedes a consonant cluster, and thus does not tense even when stressed. Provection in such items is always of the neutralizing variety, and it does not create alternations. In that sense, such cases are unproblematic for a synchronic account: ‘aunt’ is simply underlyingly /mɔtrɪb/ rather than the historical /mɔdrɪb/.

5 The argument for emergent features

In this section I argue that the account of provection given here does in fact provide an argument in favour of an emergent theory of features as opposed to an approach where features are universal (and probably innate). I consider two questions here. First, is the association between ‘tenseness’ in vowels and in ‘fortis’ stops entirely phonetically arbitrary, or is there a phonetic rationale behind it? Second, to what extent does the patterning of [tense] in south-eastern Welsh support an approach to phonology that relies on emergent, substance-free mappings

between phonetics and phonology, or should we aim to keep the proposition that features are phonetically grounded as the null hypothesis, as argued by Hall (this volume)?[26]

5.1 How arbitrary is [tense]?

So far, I have been referring to the feature [tense] used to implement the 'tenseness' distinction in vowels and laryngeal contrast in consonants as an 'emergent', substance-free one. This might seem to imply that the relationship between these segment classes is not phonetically motivated, and thus perhaps unexpected, but this assumption has so far not been examined.

Some typological facts have been analysed in previous literature do show that 'tense' quality in vowels – more usually analysed as [advanced tongue root] – seems to be preferentially associated with *voiced* rather than voiceless obstruents (Trigo 1991, Vaux 1996, 2009, Ahn 2018). Phonetically, the connection seems to lie in the fact that tongue root advancement leads to pharyngeal expansion, which is a possible strategy for the promotion of voicing. Phonologically, a number of patterns indicate that an [±ATR] feature can spread from voiced consonants to vowels.

As Vaux (1996) in particular discusses, however, the status of [tense] in such patterns is less clear. Traditionally, it has been assumed that 'voiceless' obstruents are [+tense], whilst 'lenis' obstruents are [−tense] (cf. the usage of *tenuis* as a synonym for *fortis* in some traditional philological literature). Roman Jakobson and Morris Halle, in an appendix to Jakobson, Fant & Halle (1951), discuss this matter explicitly, arguing that 'tense' vowels and fortis obstruents do share a phonetic property, which they identify as relatively great build-up of air pressure. Similar ideas have recurred in later literature, as documented by Kim & Clements (2015) with ample references, who themselves endorse the idea of a single [tense] feature, manifested in a number of distinct ways (cf. the idea of 'power' in Kohler [1984]).

However, despite these phonetic similarities *phonological* patterns in which tense vowels and fortis obstruents would pattern together are rare. The specification of consonants as [±tense] in the phonology is rarely required (with the notable exception of Korean, where it has frequently been proposed to form one of the poles of the laryngeal contrast). Cross-linguistically, however, as Kim & Clements (2015) point out, tenseness is very frequently used to *enhance* consonantal properties such as voicelessness or duration. Nevertheless, normally vocalic and consonantal tenseness are primarily expressed by different features.[27] The reason for this, as Vaux (1996) again points out, may be that where interactions between vowel tenseness and consonant voicing are found, they tend to belong to the type discussed above, where vocalic tenseness is patterns with 'voicing' in consonants rather than voicelessness: since voiced consonants are assumed to be [−tense], all these patterns would have to be treated as dissimilatory, with no corresponding assimilations apparently unattested.

In this sense, if the analysis of provection here is correct, Welsh provides important evidence for the possibility of [+tense] vowels and consonants patterning together in the phonological grammar. As such, the analysis given here can probably be reproduced without an explicit

[26] I am particularly grateful to Bert Vaux for raising many of the issues discussed in this section.

[27] One exception identified by Jakobson, Fant & Halle (1951) is de Groot (1929), who suggests that in Dutch „der wesentliche Unterschied zwischen den sog. langen und kurzen Vokale [ist] genau derselbe wie der zwischen den sog. stimmlosen und stimmhaften Explosiven und Spiranten".

commitment to a substance-free, emergent theory of featural specification, assuming one accepts that the feature [±tense] exists in both consonants and vowels as proposed in early feature theory, rather than being entirely replaced by laryngeal-state features (e. g. Halle & Stevens 1971, Avery & Idsardi 2001) in the former and [±ATR] in the latter. The question, however, is to what extent this grounding is a *post hoc* rationalization of the pattern: a theory that 'requires features to be phonetically interpretable' (Hall, this volume) only has heuristic (not to mention predictive) power only insofar as we can define 'phonetic interpretability' in advance, on principled grounds. To what extent patterns such as the one discussed here should be included in the restricted space of possibilities is the key question here. Answering it requires a better understanding of the phonological and phonetic dynamics of 'tenseness' within and across languages.

5.2 Language-specific representations

The proposal in this paper relies heavily not just on the specification of vowels with *some* phonological feature but also on the assumption that lenis and fortis stops are featurally identical save for *that same* feature [tense]. The analysis receives support from other phenomena in the language. In particular, the 'paired' status of fortis and lenis stops is confirmed by the existence of fortition processes such as those discussed in section 3.1, analysable using the addition of [spread glottis] to a lenis stop.

The association of provection with other processes in Welsh that involve the laryngeal contrast in stops is further confirmed by the fact that none of these patterns involve lenis fricatives [v ð], even though their fortis counterparts [f θ] do exist in the language. For instance, there is no 'devoicing' of [v ð] in degree inflection in adjectives:[28]

(33)	a.	['kriːv]	*cryf*	'strong' (218)
	b.	['krəvaχ]	*cryfach*	'stonger' (218)
	c.	*['krəfaχ]		

Similarly, [v ð] do not act as targets of [spread glottis] assimilation:

(34)	a.	['nau]	*naw*	'nine' (584)
	b.	['nauvad]	*nawfed*	'ninth' (585)
	c.	['saiθ]	*saith*	'seven' (703)
	d.	['səiθvad]	*seithfed*	'seventh' (757)
	e.	*['səiθfad]		

This inertness of [v ð] in laryngeal-feature processes suggests that they are not, in fact, the non-[tense] counterpart of [f θ], but are instead unspecified for laryngeal features, much like sonorants (cf. Botma & van 't Veer 2013). Crucially, this also accounts for the inertness of [v ð] in provection.[29]

[28] Such devoicing is possible with the verbalizing suffix *-háu*, but it is lexically irregular: *cofáu* 'to remember' from *cof* 'memory', but *cryfháu* 'to strengthen' rather than **cryffáu*.

[29] As noted above, there are some examples of what looks like provection of sonorants in inflected forms of the preposition *heb* 'without': [ɛpθo] *hebddo* 'without him'. However, the paradigm for *heb* given by C. H. Thomas

This asymmetry in the laryngeal specification of obstruents underscores the importance of internal phonological evidence for determining featural specification. In the kind of analysis envisaged here, the inertness of lenis fricatives [v ð] stems from the fact that phonologically they are not simply non-[tense] (or [−tense]) counterparts of the corresponding fortis fricatives [f θ]. This is a language-specific circumstance that can only be discovered through analysis of the phonological pattern; if featural specification relied on the phonetic properties of the relevant segments alone, [v ð] would very likely be specified as the [−tense] counterparts of [f θ].

Resolving this asymmetry does not, of course, require the adoption of a theory based on emergent features: for instance, the lack of a [±tense] specification in [v ð] can be achieved via a properly constructed contrastive hierarchy, as advocated among others by Hall (this volume) (see Iosad [2012b] for a concrete implementation of this idea for Welsh). Its existence does emphasize, however, the conceptually arbitrary nature of both phonological classes and the labels we use to refer to them.

In this context, it is instructive to reconsider the relationship of the Welsh system of laryngeal contrast and 'laryngeal realism'. As noted in section 3.1, although Welsh seems to be a relatively 'well-behaved' language in terms of laryngeal realism, in that the phonologically marked pole of the laryngeal contrast (the fortis stops) is generally marked with post- and preaspiration, the extremely tight coupling between the relative phonological markedness of segments and the precise way in which contrast is realized proposed by laryngeal realism, cannot be sustained, either empirically (Vaux & Samuels 2005, Iosad 2017a, Kirby & Ladd 2019, Salmons 2020) or conceptually within the framework adopted here.

I suggest, however, that the *phonological* insights behind many laryngeal realism analyses may be sound: asymmetries in phonological behaviour do exist in some languages, and can be insightfully captured via structural asymmetries in featural specification, and languages differ as to whether it is the 'fortis' or the 'lenis' series that is more marked.[30] The phonetic implementation of such asymmetries, however, is more complex than a simple binary distinction.

It is here that the explicitly arbitrary nature of feature labels as cover descriptions of phonological behaviour provides an important insight. The diversity of phonetic realizations, especially across different positions, should not detract from the fact that in phonological terms the distinction is still basically (in Welsh or English) a binary one. In non-neutralizing positions, for instance in the onset of a stressed syllable (ignoring for now the potential presence of a fricative), Welsh only offers a two-term contrast between 'fortis' and 'lenis' stops. More distinctions are available in post-tonic position, since the featural contrast is cross-cut by a quantity distinction (cf. ['pɑːbɪð] vs. ['kɔbːɪn]). Crucially, there is no independent evidence for a third term of the laryngeal contrast that should be expressed featurally. The distinction between provected [tense] stops (which do not head a [tense] domain) and 'ordinary' ones (that do) is expressed by autosegmental structure, but there is no phonological evidence that requires reference to any

(1993: vol. 2, p. 92) shows that *all* inflected forms contain [θ]. There is thus no synchronic alternation betweeh [ð] and [θ] in these cases. (There is one other preposition that has [ð] in the stem of inflected forms: [tru] *trwy* 'through', stem [trʊjð-]. Generally the morphology of the inflected prepositions is quite irregular.)

[30]The question of whether laryngeal features can be privative or whether binary specification is required is clearly related, but goes beyond the scope of this paper; again see the literature cited in footnote 8 above for more discussion.

other features. Any such features are redundant for the purposes of lexical contrast, and since they do not also appear to be active in alternations, the phonological evidence for them is slim.

This analysis contrasts with a possible approach that still relies on a universal SPE-style feature set. If we accept a universal feature [±tense], it is possible to assume that both vowels and stops are underlyingly specified for it, and [±voice] and [±spread glottis] values for the latter are introduced later in the derivation. In that case, the analysis of provection offered here can be reproduced by assuming [+tense] assimilation of consonants to preceding vowels, along the lines of the following rule (or its autosegmental analogue):

(35) $\begin{bmatrix} -\text{son} \\ -\text{cont} \end{bmatrix} \rightarrow \begin{bmatrix} +\text{tense} \end{bmatrix} / \begin{bmatrix} +\text{syl} \\ +\text{tense} \end{bmatrix} -$

The output of this rule (as well as any other rules involving [±tense], such as the 'provection' induced by /h/-initial suffixes) is the subjected to redundancy rules of the following type:

(36) a. Fortis stops

$$\begin{bmatrix} -\text{son} \\ -\text{cont} \\ +\text{tense} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{voi} \\ +\text{spr gl} \end{bmatrix}$$

b. Provected stops

$$\begin{bmatrix} -\text{son} \\ -\text{cont} \\ +\text{tense} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{voi} \\ -\text{spr gl} \end{bmatrix} / \begin{bmatrix} +\text{syl} \\ +\text{tense} \end{bmatrix} -$$

This succeeds in deriving the apparent ternary distinction on the surface (although, depending on the mechanism used, the parallelism between two types of doubly linked [tense] structures may be lost), at least if we grant that provected stops are [−spread glottis] (which is not quite assured, given that some degree of glottal spreading is necessary to block closure voicing). Crucially, however, this ternary *featural* distinction is introduced not on any independently verifiable grounds, but because universal featural structure does not allow any means of expressing categorical phonetic distinctions other than featural differences. The analysis thus has to commit to structural reflection of phonetic differences that otherwise appear to be invisible to the phonology. At the same time it remains vulnerable to the criticism that the featural structures are still insufficient to determine the phonetic realization of the contrast to the necessary extent, which, after all, is largely the point of having a substance-based representation in the first place.

I do not wish to claim here that the analysis of provection offered above, if correct, provides a knock-down argument in favour of substance-free theories relying on emergent features. In fact, as a reviewer points out, even if we reject the substance-free approach and accept an analysis using a universal [±tense] feature along the lines sketched in the preceding paragraphs, Welsh provection still provides a valuable example of a morphophonological process in which 'fortis' stops are [+tense]: the kind of phenomenon that traditional approaches (where fortisness is associated with tenseness) lead us to expect but that nevertheless has proven to be relatively elusive, unlike cases of [ATR]-voicing interaction that have to be treated as dissimilatory for

[±tense] (Vaux 1996). Therefore, this paper provides important evidence for feature theory even if the case for emergent features made here is not seen as sufficiently convincing.

Nevertheless, the substance-free approach has the advantage of foregrounding the idea that phonological computation is, in principle, 'unnatural' (Anderson 1981). This is reflected in the analysis given here in two ways. First, [tense] acts as a cover feature for a variety of phonetic realizations across prosodic contexts, reflecting the essentially two-term nature of the distinction as observed in the phonology. Second, the phenomenon is 'unnatural' in that it does not have an obvious phonetic motivation, and yet the phonological computation is able to derive the pattern using perfectly ordinary mechanisms such as autosegmental spreading and delinking, and thus the grammar allows it to exist. I conclude, therefore, that a detailed examination of an 'unnatural' sound pattern can support the proposition that a substance-free account is conceptually superior to one where phonological features are closely tied to phonetic substance. Moreover, it shows that a substance-free theory reliant on emergent features is a viable framework not just for the sort of large-scale comparisons discussed by Mielke (2008, 2013) and Mielke, Magloughlin & Hume (2010) and aptly criticized by Hall (2010 and this volume), but also for in-depth synchronic (and potentially diachronic) analysis of individual languages.

5.3 Why features should be emergent

The analysis of provection offered above shows a concrete example of emergent featural specifications within a substance-free phonological framework. Under the view of phonologization espoused here, phonological representation is substance-free precisely *because* it is emergent. More specifically, a representational system requires a set of categories that participate in phonological computation and a set of labels assigned to these features. The set of categories, in an emergentist approach, can be induced from discrete distributions in the data (cf. Seinhorst, Boersma & Hamann 2019). As for labelling, in principle a simple system might operate with holistic categorization whereby every category is assigned a unique label. This is commonly assumed not to happen because of the existence of class behaviour (although see Nazarov [2014] for a qualification). However, class behaviour only allows the learner to pinpoint the extent of the class of segments a label should encompass: the choice of the labels themselves is logically an independent operation. This operation is (one possible) source of 'substance-free' representations, if we assume that labelling proceeds on the basis of not only phonetic properties of the segments but also their phonological behaviour such as patterns of contrasts and alternations. This latter assumption is of course shared with various theories of contrastive specification (Dresher 2009).

Of course, the scale of this 'freedom from substance' should not be overstated: in most cases, featural classes end up being phonetically coherent, presumably because of the greater learnability of a more systematic mapping between phonological representation and phonetic interpretation (see, for instance, Baer-Henney & van de Vijver [2012] and van de Vijver & Baer-Henney [2014]). Accepting 'tenseness' as a phonetic common denominator for fortis stops and tense vowels is not at all incompatible with an emergent-feature theory. The biases responsible for this, however, should not be seen as a hard constraint. The substance-free approach allows us to posit 'cover' features that lack a phonetic invariant found in all contexts: Welsh [tense]

may be such a case. Admittedly, as noted above, we may consider 'tenseness' as described by Kim & Clements (2015) to be a candidate for such an invariant; however, their description is also so broad as to make any claims to universal status for this distinction rather shaky. This breadth of possible specific correlates highlights the fact that *some* degree of latitude in the phonology-phonetics mapping should be allowed for.[31] And yet, if phonological representation is not completely determined by the phonetic substance, then the *null* hypothesis is that it is fully substance-free (*pace* Hall, the volume).

6 Conclusion

In this paper I proposed a new interpretation of the phenomenon of 'provection' in south-eastern Welsh. Provection is traditionally described as a process of 'devoicing' targeting 'voiced' stops after a stressed vowel. I argued that it can be understood as an unusual type of consonant-vowel interaction. Although typologically vowel tenseness is associated with voicing rather than with devoicing, I suggested that in south-eastern Welsh tense vowels and fortis stops share a feature to the exclusion of lenis stops.

If the account of provection offered in the paper is correct, it makes several predictions. First, and most obviously, it relies crucially on the assumption that provected stops do not show significant volumes of preaspiration. Although the pilot study by Spooner (2016) gives reason to give some credence to this assumption, further empirical work is clearly required. Second, it predicts that provection in consonant clusters (as in [go'vɪ**t**jo]) should be neutralizing, in contrast to provection of singletons (as in ['keːkɪn]). Again, this appears to follow from existing descriptions, but has not been confirmed empirically. The verification of these predictions is left to future work.

Apart from providing an example of a robustly phonological alternation involving the [tense] specification of voiceless obstruents, I have also argued that this kind of vowel-consonant interaction highlights the nature of features as cover labels for categories involved in phonological patterns, rather than categories having intrinsic phonetic content (cf. in particular Ladd 2011, Odden 2013). I suggest that under such an approach the null hypothesis is that the mapping between phonological and phonetic categories is *in principle* arbitrary, and hence that phonological features are fully emergent. Of course, this does not mean that the null hypothesis must be accepted. Principled restrictions on the arbitrariness of this mapping can exist, arising from some properties of the faculty of language (as in the substance-free framework of Hale & Reiss [2008], Volenec & Reiss [2017] and Reiss [2018]), or perhaps from more general cognitive mechanisms (as in Samuels [2011]). I suggest, however, that pursuing phonological analysis whilst accepting the emergentist null hypothesis is also a productive enterprise that can bring new insights into the structure of phonological theory.

[31] Much the same point can be made with respect to the element |H| of Element Theory (e. g. Harris 1994, Backley 2011), which could otherwise supplant [tense] as used here.

References

Adger, David & Jennifer Smith. 2010. Variation in agreement: A lexical feature-based approach. *Lingua* 120(5). 1109–1134.

Ahmed, Samuel K., Samuel Andersson & Bert Vaux. 2020. English phonology and morphology. In Bas Aarts, April McMahon & Lars Hinrichs (eds.), *The handbook of English linguistics*, 2nd edn., 345–364. Oxford: Wiley-Blackwell.

Ahn, Suzy. 2018. The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazilian Portuguese. *Journal of Phonetics* 71. 451–467.

Anderson, Stephen R. 1981. Why phonology isn't 'natural'. *Linguistic Inquiry* 12(4). 493–539.

Avery, Peter. 1996. *The representation of voicing contrasts*. Toronto: University of Toronto dissertation.

Avery, Peter & William J. Idsardi. 2001. Laryngeal dimensions, completion, and enhancement. In T. Alan Hall (ed.), *Distinctive feature theory* (Phonetics and Phonology 2), 41–71. Berlin: Mouton de Gruyter.

Awbery, Gwenllian M. 1984. Phonotactic constraints in Welsh. In Martin J. Ball & Glyn E. Jones (eds.), *Welsh phonology: Selected readings*, 65–104. Cardiff: University of Wales Press.

Awbery, Gwenllian M. 1986. *Pembrokeshire Welsh: A phonological study*. Llandysul: Welsh Folk Museum.

Backley, Phillip. 2011. *An introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Baer-Henney, Dinah & Ruben van de Vijver. 2012. On the role of substance, locality, and amount of exposure in the acquisition of morphophonemic alternations. *Laboratory Phonology* 3(2). 221–256.

Ball, Martin J. 1984. Phonetics for phonology. In Martin J. Ball & Glyn E. Jones (eds.), *Welsh phonology: Selected readings*, 5–39. Cardiff: University of Wales Press.

Ball, Martin J. & Briony Williams. 2001. *Welsh phonetics* (Welsh Studies 17). Lewiston, Queenston & Lampeter: Edwin Mellen Press.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Beckman, Jill, Michael Jessen & Catherine Ringen. 2013. Empirical evidence for laryngeal features: Aspirating vs. true voice languages. *Journal of Linguistics* 49(2). 259–284.

Bermúdez-Otero, Ricardo. 2007. Diachronic phonology. In Paul de Lacy (ed.), *The Cambridge handbook of phonology*, 497–518. Cambridge: Cambridge University Press.

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*. Vol. 4, 2019–2048. Oxford: Blackwell Publishing.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed.), *The phonology and morphology of exponence: The state of the art* (Oxford Studies in Theoretical Linguistics 41), 8–83. Oxford: Oxford University Press.

Bermúdez-Otero, Ricardo. 2013. The Spanish lexicon stores stems with stem vowels, not roots with inflectional class features. *Probus* 25(1). 3–103.

Bermúdez-Otero, Ricardo. 2015. Amphichronic explanation and the life cycle of phonological processes. In Patrick Honeybone & Joseph C. Salmons (eds.), *The Oxford handbook of historical phonology*, 374–399. Oxford: Oxford University Press.
Bermúdez-Otero, Ricardo. 2016. We do not need structuralist morphemes, but we do need constituent structure. In Daniel Siddiqi & Heidi Harley (eds.), *Morphological metatheory* (Linguistik Aktuell 229), 387–430. Amsterdam & Philadelphia: John Benjamins.
Bermúdez-Otero, Ricardo. 2018. Stratal phonology. In S. J. Hannahs & Anna R. K. Bosch (eds.), *The Routledge handbook of phonological theory*, 100–134. London, New York: Routledge.
Bermúdez-Otero, Ricardo & Graeme Trousdale. 2012. Cycles and continua: On unidirectionality and gradualness in language change. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *Handbook on the history of English: Rethinking approaches to the history of English*, 691–720. Oxford: Oxford University Press.
Bickmore, Lee S. 2000. Downstep and fusion in Namwanga. *Phonology* 17(3). 297–331.
Botma, Bert & Marijn van 't Veer. 2013. A fraction too much friction: The phonological status of voiced fricatives. *Linguistics in the Netherlands* 30. 46–60.
Brake, Phylip. 2011. Amrywio ieithyddol ymhlith siaradwyr Cymraeg Treorci ar ddiwedd y saithdegau. *Gwerddon* 7. 9–44.
Carlyle, Karen Ann. 1988. *A syllabic phonology of Breton*. Toronto: University of Toronto dissertation.
Cassimjee, Farida & Charles W. Kisseberth. 1998. Optimal domains theory and Bantu tonology: A case study from Isixhosa and Shingazidja. In Larry M. Hyman & Charles W. Kisseberth (eds.), *Theoretical aspects of Bantu tone* (CSLI Lecture Notes), 265–314. Stanford, CA: CSLI.
de Groot, Albert Willem. 1929. Zum phonologischen System des Nordniederländischen. In *Donum natalicium Schrijnen: Verzameling van opstellen door oud-leerlingen en bevriende vakgenooten opgedragen aan Mgr. Prof. Dr. Jos Schrijnen bij gelegenheid van zijn zestigsten verjaardag*, 549–550. Nijmegen & Utrecht: N. V. Dekker & van de Vegt.
Dillon, Brian, Ewan Dunbar & William J. Idsardi. 2012. A single-stage approach to learning phonological categories: Insights from Inuktitut. *Cognitive Science* 37(2). 344–377.
Dresher, B. Elan. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.
Dresher, B. Elan. 2014. The arch not the stones: Universal feature theory without universal features. *Nordlyd* 41(2). 165–181.
Dunbar, Ewan, Brian Dillon & William J. Idsardi. 2013. A Bayesian evaluation of the cost of abstractness. In Montserrat Sanz, Itziar Laka & Michael K. Tanenhaus (eds.), *Language down the garden path: The cognitive and biological basis for linguistic structures*, 360–384. Oxford: Oxford University Press.
Dyck, Carrie. 1995. *Constraining the phonology–phonetics interface, with exemplification from Spanish and Italian dialects*. Toronto: University of Toronto dissertation.
Flowers, Ness. 1994. Astudiaeth o galediad yn llafar pobl Y Creunant, Cwm Dulais. In Hywel Teifi Edwards (ed.), *Cyfres y Cymoedd: Nedd a Dulais*, 56–94. Llandysul: Gwasg Gomer.
Fynes-Clinton, Osbert Henry. 1913. *The Welsh vocabulary of the Bangor district*. Oxford: Oxford University Press.

Garrett, Andrew & Keith Johnson. 2013. Phonetic bias in sound change. In Alan C. L. Yu (ed.), *Origins of sound change: Approaches to phonologization*, 51–97. Oxford: Oxford University Press.
Ghini, Mirco. 2001. Place of articulation first. In T. Alan Hall (ed.) (Phonetics and Phonology 2), 147–176. Berlin: Mouton de Gruyter.
Grawunder, Sven, Sabine Asmus & Cormac Anderson. 2015. On the correlation of acoustic vowel and coda duration in modern Welsh C(C)VC monosyllables. In The Scottish Consortium for ICPhS 2015 (ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: The University of Glasgow.
Greene, David. 1967. Varia: provection and *caledíad*. *Studia Celtica* 2. 101–104.
Hale, Mark & Charles Reiss. 2008. *The phonological enterprise*. Oxford: Oxford University Press.
Hall, Daniel Currie. 2010. Probing the unnatural. *Linguistics in the Netherlands* 27. 73–85.
Hall, Daniel Currie. 2011. Phonological contrast and its phonetic enhancement: Dispersedness without dispersion. *Phonology* 28(1). 1–54.
Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2). 215–275.
Halle, Morris & Kenneth Stevens. 1971. A note on laryngeal features. In *Quarterly progress report, Research Laboratory of Electronics*, vol. 101, 198–211. Cambridge, MA.
Hannahs, S. J. 2009. Welsh svarabhakti: Sonority sequencing and foot structure. *Journal of Celtic Linguistics* 13. 21–44.
Hannahs, S. J. 2011. Unity in diversity in Welsh: The avoidance of sonority sequencing violations. In Andrew Carnie (ed.), *Formal approaches to Celtic linguistics*, 359–374. Newcastle upon Tyne: Cambridge Scholars Publishing.
Hannahs, S. J. 2013. *The phonology of Welsh*. Oxford: Oxford University Press.
Harris, John. 1994. *English sound structure*. Oxford: Blackwell.
Hejná, Michaela. 2015. *Pre-aspiration in Welsh English: A case study of Aberystwyth*. Manchester: University of Manchester dissertation.
Hejná, Michaela. 2019. Pre-aspiration and the problem of zeroes: Phonological rules can be variable. In Ken Ramshøj Christensen, Johanna Wood & Henrik Jørgensen (eds.), *The sign of the V: Papers in honour of Sten Vikner*, 227–242. Aarhus: AU Scholarly Publishing Services.
Hirose, Aki & Michael Ashby. 2007. An acoustic study of devoicing of the geminate obstruents in Japanese. In *Proceedings of ICPhS XVI*, 909–912. Saarbrücken. http://www.icphs2007.de/conference/Papers/1425/1425.pdf.
Honeybone, Patrick. 2005. Diachronic evidence in segmental phonology: The case of obstruent laryngeal specification. In Marc van Oostendorp & Jeroen van de Weijer (eds.), *The internal organization of phonological segments* (Studies in Generative Grammar 77), 319–354. Berlin: Mouton de Gruyter.
Honeybone, Patrick. 2008. Lenition, weakening and consonantal strength: Tracing concepts through the history of phonology. In Joaquim Brandão de Carvalho, Tobias Scheer & Philippe Ségéral (eds.), *Lenition and fortition* (Studies in Generative Grammar 99), 9–93. Berlin: Mouton de Gruyter.
Honeybone, Patrick. 2012. Lenition in English. In Terttu Nevalainen & Elizabeth Closs Traugott (eds.), *Handbook on the history of English: Rethinking approaches to the history of English*, 773–787. Oxford: Oxford University Press.

Inkelas, Sharon. 1994. The consequences of optimization for underspecification. *North East Linguistic Society (NELS)* 27. 287–302.
Iosad, Pavel. 2012a. Final devoicing and vowel lengthening in Friulian: A representational approach. *Lingua* 122(8). 922–951.
Iosad, Pavel. 2012b. *Representation and variation in substance-free phonology: A case study in Celtic*. Tromsø: University of Tromsø dissertation.
Iosad, Pavel. 2017a. *A substance-free framework for phonology: An analysis of the Breton dialect of Bothoa* (Edinburgh Studies in Theoretical Linguistics 2). Edinburgh: Edinburgh University Press.
Iosad, Pavel. 2017b. The phonologization of redundancy: Length and quality in Welsh vowels. *Phonology* 34(1). 121–162.
Iosad, Pavel. 2017c. Welsh svarabhakti as stem allomorphy. *Transactions of the Philological Society* 115. 141–175.
Iosad, Pavel. 2019. Welsh vowels, 2014. Dataset, University of Edinburgh.
Itô, Junko & Armin Mester. 1999. Realignment. In René Kager, Harry van der Hulst & Wim Zonneveld (eds.), *The prosody–morphology interface*, 188–217. Cambridge: Cambridge University Press.
Iverson, Gregory K. & Joseph C. Salmons. 1995. Aspiration and laryngeal representation in Germanic. *Phonology* 12(3). 369–396.
Jakobson, Roman, Gunnar Fant & Morris Halle. 1951. *Preliminaries to speech analysis*. Cambridge, MA: MIT Press.
Jones, Glyn E. 1984. The distinctive vowels and consonants of Welsh. In Martin J. Ball & Glyn E. Jones (eds.), *Welsh phonology: Selected readings*, 40–64. Cardiff: University of Wales Press.
Jones, Glyn E. 1988. Some features of the Welsh of Breconshire. In Martin J. Ball (ed.), *The use of Welsh: A contribution to sociolinguistics* (Multilingual Matters 36), 97–103. Clevedon, Philadelphia: Multilingual Matters.
Jones, Glyn E. 2000. *Iaith lafar Brycheiniog: Astudiaeth o'i ffonoleg a'i morffoleg*. Caerdydd: Gwasg Prifysgol Cymru.
Jones, Mari C. 1998. *Language obsolescence and revitalization: Linguistic change in two sociolinguistically contrasting Welsh communities* (Oxford Studies in Language Contact). Oxford: Oxford University Press.
Jurgec, Peter. 2010. *Feature spreading 2.0: A unified theory of assimilation*. Tromsø: University of Tromsø dissertation.
Keating, Patricia. 1988. Underspecification in phonetics. *Phonology* 5(2). 275–292.
Kenstowicz, Michael. 1997. Quality-sensitive stress. *Rivista di linguistica* 9. 157–187.
Keyser, Samuel Jay & Kenneth N. Stevens. 2006. Enhancement and overlap in the speech chain. *Language* 82(1). 33–63.
Kim, Hyunsoon & G. Nick Clements. 2015. The feature [tense]. In Annie Rialland, Rachid Ridouane & Harry van der Hulst (eds.), *Features in phonology and phonetics: Posthumuous writings by Nick Clements*, 159–178. Berlin: De Gruyter.
Kim, Yuni. 2002. Phonological features: privative or equipollent? A. B. thesis, Harvard University.
Kingston, John. 1990. Articulatory binding. In John Kingston & Mary Beckman (eds.), *Papers in Laboratory Phonology I*, 406–434. Cambridge.

Kiparsky, Paul. 1995. The phonological basis of sound change. In John Goldsmith (ed.), *The handbook of phonological theory*, 640–670. Oxford: Blackwell.
Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17(2–4). 351–367.
Kirby, James & D. Robert Ladd. 2019. Effects of obstruent voicing on vowel F_0: Implications for laryngeal realism. *Yearbook of the Poznań Linguistic Meeting* 4. 213–235.
Kohler, Klaus J. 1984. Phonetic explanation in phonology: The feature fortis/lenis. *Phonetica* 41(3). 150–174.
Ladd, D. Robert. 2011. Phonetics in phonology. In John Goldsmith, Jason Riggle & Alan Yu (eds.), *The handbook of phonological theory*, 2nd edn., 348–373. Oxford: Wiley-Blackwell.
Magnus Pétursson. 1978. Jointure au niveau glottal. *Phonetica* 35(2). 65–85.
Mayr, Robert & Hannah Davies. 2011. A cross-dialectal acoustic study of the monophthongs and diphthongs of Welsh. *Journal of the International Phonetic Association* 41(1). 1–25.
Mielke, Jeff. 2008. *The emergence of distinctive features*. Oxford: Oxford University Press.
Mielke, Jeff. 2013. Phonologization and the typology of feature behavior. In Alan C. L. Yu (ed.), *Origins of sound change: Approaches to phonologization*, 165–180. Oxford: Oxford University Press.
Mielke, Jeff, Lyra Magloughlin & Elizabeth Hume. 2010. Evaluating the effectiveness of Unified Feature Theory and three other feature systems. In John A. Goldsmith, Elizabeth Hume & W. Leo Wetzels (eds.), *Tones and features: Phonetic and phonological perspectives* (Studies in Generative Grammar 107), 223–263. Berlin: De Gruyter.
Morén, Bruce. 2001. *Distinctiveness, coercion, and sonority: A unified theory of weight*. London, New York: Routledge.
Morén, Bruce. 2006. Consonant–vowel interactions in Serbian: Features, representations and constraint interactions. *Lingua* 116(8). 1198–1244.
Morén, Bruce. 2007. The division of labour between segment-internal structure and violable constraints. In Sylvia Blaho, Patrik Bye & Martin Krämer (eds.), *Freedom of analysis?* (Studies in Generative Grammar 95), 313–344. Berlin: Mouton de Gruyter.
Morris, Jonathan. 2010. Phonetic variation in North Wales: preaspiration. In *Proceedings of the Second Summer School on Sociolinguistics, University of Edinburgh*. http://www.lel.ed.ac.uk/sssocio/proceedings/Jon.pdf.
Morris, Jonathan & Míša Hejná. 2020. Pre-aspiration in Bethesda Welsh: A sociophonetic analysis. *Journal of the International Phonetic Association* 50(2). 168–192.
Myler, Neil. 2015. Stem storage? Not proven: A reply to Bermúdez-Otero (2013). *Linguistic Inquiry* 46(1). 173–186.
Nazarov, Aleksei. 2014. A radically emergentist approach to phonological features: Implications for grammars. *Nordlyd* 41(1). 21–58.
Ní Chasaide, Ailbhe. 1986. *Preaspiration in phonological stop contrasts: An instrumental phonetic study*. Bangor: University College of North Wales dissertation.
Odden, David. 2013. Formal Phonology. *Nordlyd* 40(1). 24–43.
Ohala, John J. 1981. The listener as the source of sound change. *CLS* 17. 178–203.
Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter F. MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.

Peperkamp, Sharon, Rozenn Le Calvez, Jean-Pierre Nadal & Emmanuel Dupoux. 2006. The acquisition of allophonic rules: Statistical learning with linguistic constraints. *Cognition* 101(3). B31–B41.
R Core Team. 2017. R: A Language and Environment for Statistical Computing. Version 3.4.2. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/.
Ramsammy, Michael. 2015. The life cycle of phonological processes: Accounting for dialectal microtypologies. *Language and Linguistics Compass* 9(1). 33–54.
Rees, Iwan Wyn. 2015. Phonological variation in Mid Wales. *Studia Celtica* 49(1). 149–174.
Reiss, Charles. 2003. Deriving the feature-filling/feature-changing contrast: An application to Hungarian vowel harmony. *Linguistic Inquiry* 34(2). 199–224.
Reiss, Charles. 2018. Substance-free phonology. In S. J. Hannahs & Anna R. K. Bosch (eds.), *The Routledge handbook of phonological theory*, 425–452. London, New York: Routledge.
Rosen, Kristin M. 2005. Analysis of speech segment duration with the lognormal distribution: A basis for unification and comparison. *Journal of Phonetics* 33(4). 411–426.
Salmons, Joseph C. 2020. Germanic laryngeal phonetics and phonology. In Richard B. Page & Michael T. Putnam (eds.), *The Cambridge handbook of Germanic linguistics*, 119–142. Cambridge: Cambridge University Press.
Samuels, Bridget. 2011. *Phonological architecture: A biolinguistic perspective* (Oxford Studies in Biolinguistics 2). Oxford: Oxford University Press.
Sandstedt, Jade J. 2018. *Feature specifications and contrast in vowel harmony: The orthography and phonology of Old Norwegian height harmony*. Edinburgh: The University of Edinburgh dissertation.
Seinhorst, Klaas, Paul Boersma & Silke Hamann. 2019. Iterated distributional and lexicon-driven learning in a symmetric neural network explains the emergence of features and dispersion. In *Proceedings of the 19th International Congress of Phonetic Sciences*, 1134–138. Canberra: Australasian Speech Science & Technology Association.
Simon Evans, D. 1964. *A grammar of Middle Welsh*. Dublin: Dublin Institute for Advanced Studies.
Sommerfelt, Alf. 1925. *Studies in Cyfeiliog Welsh: A contribution to Welsh dialectology*. Avhandlinger utgitt av Det Norske Videnskaps-Akademi i Oslo. II. Hist.-Filos. Klasse. 192, No. 3. Oslo: I kommission hos Jacob Dybwad.
Spooner, Ciarán. 2016. *Provection in Cwmtawe Welsh*. Edinburgh: The University of Edinburgh MA (Hons) dissertation.
Stevens, Kenneth N. & Samuel Jay Keyser. 2010. Quantal theory, enhancement and overlap. *Journal of Phonetics* 38(1). 10–19.
Stevens, Mary & Jonathan Harrington. 2013. The individual and the actuation of sound change. *Loquens* 1(1). e003.
Strycharczuk, Patrycja. 2012. *Phonetics-phonology interaction in pre-sonorant voicing*. Manchester: University of Manchester dissertation.
Strycharczuk, Patrycja & Ellen Simon. 2013. Obstruent voicing before sonorants: The case of West-Flemish. *Natural Language and Linguistic Theory* 31(2). 563–588.
Strycharczuk, Patrycja, Marijn van 't Veer, Martine Bruil & Kathrin Linke. 2014. Phonetic evidence on phonology–morphosyntax interactions: Sibilant voicing in Quito Spanish. *Journal of Linguistics* 50(2). 403–453.

Thomas, Alan R. 1966. Systems in Welsh phonology. *Studia Celtica* 1. 93–127.
Thomas, Beth. 1990. Amrywio sosioiethyddol yn nhafodiaith Pont-rhyd-y-fen. In Martin J. Ball, James Fife, Erich Poppe & Jenny Rowland (eds.), *Celtic linguistics/Ieithyddiaeth Geltaidd: Readings in the Brythonic languages. Festschrift for Arwyn T. Watkins*, 41–52. Amsterdam: John Benjamins.
Thomas, Ceinwen H. 1975. Some phonological features of dialects in South-East Wales. *Studia Celtica* 10–11. 345–366.
Thomas, Ceinwen H. 1993. *Tafodiaith Nantgarw: Astudiaeth o Gymraeg llafar Nantgarw yng Nghwm Taf, Morgannwg*. Caerdydd: Gwasg Prifysgol Cymru.
Thomas, Siân Elizabeth. 1983. *Astudiaeth o galediad yn Ystalyfera*. Swansea: University College of Swansea MA thesis.
Thomas, Siân Elizabeth. 1988. A study of *caledíad* in the Upper Swansea Valley. In Martin J. Ball (ed.), *The use of Welsh: A contribution to sociolinguistics* (Multilingual Matters 36), 85–96. Clevedon, Philadelphia: Multilingual Matters.
Thorne, David. 1984. The correlation of dialect and administrative boundaries in Welsh: A review. In Martin J. Ball & Glyn E. Jones (eds.), *Welsh phonology: Selected readings*, 176–188. Cardiff: University of Wales Press.
Trigo, Loren. 1991. On pharynx–larynx interactions. *Phonology* 8(1). 113–136.
Trommer, Jochen & Eva Zimmermann. 2014. Generalised mora affixation and quantity-manipulating morphology. *Phonology* 31(3). 463–510.
van de Vijver, Ruben & Dinah Baer-Henney. 2014. Developing biases. *Frontiers in Psychology* 5.
Vaux, Bert. 1996. The status of ATR in feature geometry. *Linguistic Inquiry* 27(1). 175–182.
Vaux, Bert. 2008. Why the phonological component must be serial and rule-based. In Bert Vaux & Andrew Nevins (eds.), *Rules, constraints and phonological phenomena*, 20–61. Oxford: Oxford University Press.
Vaux, Bert. 2009. [ATR] and [back] harmony in the Altaic languages. In Sergei Tatevosov (ed.), *Investigations into formal Altaic linguistics: Proceedings of WAFL 3*, 50–67. Moscow: MAKS-Press.
Vaux, Bert & Bridget Samuels. 2005. Laryngeal markedness and aspiration. *Phonology* 22(3). 395–436.
Vaux, Bert & Andrew Wolfe. 2009. The appendix. In Eric Raimy & Charles Cairns (eds.), *Contemporary views on architecture and representations in phonology* (Current Studies in Linguistics 48), 101–144. Cambridge, MA: MIT Press.
Volenec, Veno & Charles Reiss. 2017. Cognitive Phonetics: The transduction of distinctive features at the phonology-phonetics interface. *Biolinguistics* 11.SI. 251–294.
Watkins, Arwyn T. 1967. Some phonological features of the Welsh dialect of Llansamlet (West Glamorgan). In Wolfgang Meid (ed.), *Beiträge zur Indogermanistik und Keltologie: Julius Pokorny zum 80. Geburtstag gewidmet*, 315–322. Innsbruck: Sprachwissenschaftliches Institut der Universität Innsbruck.
Wells, John C. 1979. Final voicing and vowel length in Welsh. *Phonetica* 36(4–5). 344–360.
Williams, Briony. 1985. Pitch and duration in Welsh stress perception: The implications for intonation. *Journal of Phonetics* 13(4). 381–406.
Wmffre, Iwan. 2003. *Language and place-names in Wales: The evidence of toponymy in Cardiganshire*. Cardiff: University of Wales Press.

Wmffre, Iwan. 2012. Review of Brian Ó Curnáin. 2007. *The Irish of Iorras Aithneach, County Galway*. Dublin: Dublin Institute for Advanced Studies. *Journal of Celtic Linguistics* 14. 130–151.

Wmffre, Iwan. 2013. *The qualities and origins of the Welsh vowel [ɨ]*. Berlin: Curach Bhán Publications.

Wood, Simon N. 2006. *Generalized additive models: An introduction with R*. Chapman & Hall/CRC.

Youssef, Islam. 2010. Laryngeal assimilation in Buchan Scots. *English Language and Linguistics* 14(3). 321–345.