Acta Linguistica Academica Vol. **67** (2020) 1, 53–71
DOI: 10.1556/2062.2020.00005

Tense? (Re)lax!

A new formalisation for a controversial contrast

Markus A. Pöchtrager
University of Vienna
markus.poechtrager@univie.ac.at

Abstract: This article looks at what is referred to as the tense/lax contrast in English and proposes that members of the two sets of vowel have the same basic structure but differ in how part of that structure is made use of by its neighbours. The proposal forms part of a general theory of the representation of vowel height within the framework of Government Phonology 2.0.

Keywords: English; tense; lax; height; vowel system; Government Phonology 2.0

1. Introduction

The representation of vowel height, as Pulleyblank (2011) points out, is not trivial: in a five-vowel system (say, Greek), *i* would typically be categorised as high, *e* as mid, *a* as low. A seven-vowel system (Brazilian Portuguese, Hindi) with two series of mid vowel (open-mid [ɛ]/[ɔ] along close-mid [e]/[o]) might already be subject to disagreement. Are we to treat the difference between these vowels as one of openness only, as part of a series of equal members (high vs. mid-close vs. mid-open vs. low), where the type of relation between each pair of neighbours is constant? Or do we assume a separate category (for example, tense/lax) intersecting with a three-step openness scale (high/mid/low)? Clearly the symbols chosen in a transcription will not tell: for the DRESS vowel Wells (1982) uses [e] for RP and [ɛ] for General American, but points out that the phonetic difference is in fact miniscule (one of [e̞] versus [ɛ̝]). Yet, crucially, the vowel behaves identically in both varieties (as a checked one, in Wells's system), so a difference in symbol is no guarantee we are dealing with different phonological objects. Similar concerns hold for articulation: what is commonly transcribed as [u] is nowhere near as high as [i], yet both are classified as high (Ladefoged & Johnson 2010, 21). Articulation will not give a definite answer how to carve up the vowel space (Ladefoged & Maddieson 1996, 282ff). The answer must come from a theory on the internal structure of vowels, which this article forms part of. It builds on and extends the proposals in Pöchtrager (2018), itself couched within the framework of Government Phonology (GP) 2.0.[1] The current article will concern itself with the so-called tense/lax contrast, and data will be drawn from one variety of English, Received Pronunciation (RP) following Wells (1982), with occasional digressions into other varieties. RP, as any standard, has its problems, but many of the proposals made here will also be relevant to other varieties and languages where tense/lax is an issue (such as German or Québec French).

2. English: tense/lax?

The stressed vowels of English are usually divided into two disjoint sets which go by various names: tense/lax (with a focus on articulation, though see below, and/or behaviour), long/short (focus on quantity), or free/checked (focus on behaviour). (1) gives the stressed vowels of RP (cf. Wells 1982, 119 for the transcription). In addition to traditional labels the chart also refers to "L-type/T-type", which will be explained anon.

(1)	a.	Lax/short/checked/ L-type	b.	Tense/long/free T-type	
				[ɪə] *fear*	[ʊə] *sure*
		[bɪt] *bit*		[iː] *beat*	[uː] *boot*
		[bet] *bet*		[eɪ] *bate*	[əʊ] *boat*
		[bæt] *bat*		[ɛə/ɛː] *bare*	[ɔː] *bought*
		[pʊt] *put*		[ɜː] *dirt*	[ɑː] *part*
		[bʌt] *but*			[aɪ] *bite*
		[pɒt] *pot*			[aʊ] *bout*
					[ɔɪ] *void*

While (1a) contains short monophthongs only, (1b) is more varied. We have long monophthongs, but also diphthongs such as [eɪ] and [əʊ], which are typical of certain varieties; in their stead we find the long monophthongs [eː/oː] in other areas, e.g., in the north of England. In some diphthongs a property is shared between both members (the element **I** in [eɪ]), in others not (as in [aɪ], [aʊ], [ɔɪ]).

The phonological property underlying the contrast (and thus the name given to each group), or whether there even is one, has been the topic of much debate, cf. Bauer (1980) and Durand (2005) for overviews.

[1] References will be given throughout the text.

Long/short is often deemed problematic because the vowel in *bid* (short) tends to be much longer phonetically than that of *beat* (long); a criticism raised by Wells (1982, 119). The problem here is that the terms long/short confuse and collapse two independent axes; that of the lexical length of the vowel (corresponding to the division in (1)), and that of the additional influence of the following consonant, usually referred to as pre-fortis clipping; cf. Pöchtrager (2006) for extensive discussion and analysis. Wells (*ibid.*) prefers the (Trubetzkoyan) dichotomy free/checked: in a checked vowel (e.g., *bit*) the "final consonant [...] check[s] the pulse of air for the syllable and its vowel" (as opposed to *beat*). It is unclear what exactly that means, though, and whether the pulse of air is also checked (as it should be) in the rather long vowel of *bid*. In other words, it is not obvious that the pair free/checked is more precise than the terms long/short.

Similar problems of definition arise with tense/lax; terms which are meant to refer to muscular tension. But that alleged difference in muscular tension has been called into question time and again (Raphael & Bell-Berti 1975; Lass 1976, 47–49; Durand 2005). Ladefoged and Johnson (2010), while acknowledging phonetic differences between the two sets, also reject muscular tension and treat tense/lax as "just labels" (*op.cit.*, 98) for something that is essentially a difference in terms of **behaviour**. This lack of a clear definition of tense/lax also shows in the choice of symbols, as the IPA does not provide unique symbols associated with either kind; [ɛ] is of course not inherently "lax", but simply a cover for anything close enough to cardinal vowel 3. Assigning a given quality to an ideal cardinal vowel is not always easy, either, and the differences in transcription between authors for what should be one and the same variety (RP) bears witness to the problem: compare Wells's (1982) *bet* [bet] to Ladefoged & Johnson's (2010) [bɛt]. That [ɪ], a member of one group, comes back as part of a member of the other group (in [eɪ]) illustrates the same problem (and could be shown by other pairs, too). Replacing tense/lax by ±ATR will not do either, as Durand (2005) points out: [æ] and [iː] would both be +ATR, though they are not in the same group in (1), and neither would their −ATR counterparts [ɑː] and [ɪ] be.[2] Finally, the notion of "deviation from the neutral

[2] In addition, in languages with "true" ATR (Igbo, Akan) the tongue root position is not correlated with vowel height, while English tense/lax pairs often do vary in height (Ladefoged & Maddieson 1996, 302ff). Also, true ATR does not seem to interact with constituent structure in the way that English does (i.e., distributionally). For further arguments that the equation of tense/lax (however defined) with ATR is problematic and presumably misguided, cf. Bauer (1980); Durand (2005); Ladefoged (1968); Ladefoged & Johnson (2010); Lindau (1979). The two concepts (so-called tense/lax and true ATR) are kept strictly separate in Pöchtrager & Kaye (in preparation).

position", which Laver (1994) assumes, is problematic in that what should count as the neutral position is debatable (Bauer 1980).

Since using terms with an unclear or controversial definition runs the risk of misunderstandings, it would be wiser to abstain from traditional labels and just refer to the two sets by neutral, arbitrary names.[3] For the sake of an easy mnemonic I will use "L-type" and "T-type" for the two sets in (1); on the understanding that I do not therefore give preferential treatment to the terms tense/lax, even though it is of course clear where the letters T and L come from. Obviously, when using those terms in discussing other people's work, I also do not wish to imply that they would use those labels or agree with them. Note furthermore that by T-type/L-type I will be referring to the entire *nucleus*, independently of whether it is a monophthong or diphthong: [aɪ] as a whole is of the T-type, not its component parts. This is of course also how Wells uses free/checked. "L-type/T-type" are not intended to downplay the importance of the other characteristics that the members of each set have, and we will come back to that as we proceed.

Given the difficulty in isolating a responsible factor setting apart the two sets in (1), one can hardly blame Ladefoged & Maddieson (1996, 302ff) for rejecting the idea that there could be a common factor altogether. (In their view, nothing beyond height, backness, and rounding is needed to characterise the vowels of English.) But that will not do, either, as it leaves out a crucial piece of the puzzle, namely the fact that it is **behaviour** that is the clearest dividing line between the two sets. L-type nuclei are disallowed finally (2a) and pre-hiatus (2b), but occur freely before coda–onset clusters (2c).[4] The opposite is true for the T-type.

(2) a. *[bɪ], *[zʊ], but *bee* [biː], *zoo* [zuː]
    b. **l*[ɪ]*o*, **rod*[ɛ]*o*, but *l*[iː]*o*, *rod*[eɪ]*o*, etc.
    c. *limp* [lɪmp], *[liːmp]; *whimper* [wɪmpə], *[wiːmpə]

How can those differences in behaviour be explained? As we have seen, Wells (1982, 119) treats L-type nuclei as "checked" by a following consonant. This accounts for (2a–b), where there is no consonant following the

[3] At this point I need to thank my reviewers for their patience – I hope that the several revisions of this text have made me a wiser man and that by now I manage to avoid confusion caused by terminology.

[4] For arguments that coda–onset clusters also occur finally cf. Kaye (1990); Harris (1994); Harris & Gussmann (1998).

nucleus in question, i.e., there is nothing that could check it, hence only a free/unchecked/T-type nucleus occurs. However, this does not yet explain why L-type nuclei need checking in the first place or why they differ in quality from the T-type ([ɪ] vs. [i], [ɒ] vs. [ɔ] etc.). It also does not explain why a single final consonant can check a preceding nucleus, while a final cluster must check it (2c).

Similarly, earlier versions of GP equal T-type with headed and L-type with unheaded (Kaye 2000), where headedness refers to the arrangement of elements that make up phonological expressions (i.e., combinations of elements). An [i] can be represented as ($\underline{\mathbf{I}}$), with the (sole) element **I** as the head, and [ɪ] as (**I**, __), with again only one element and no head. Branching nuclei are required to link to headed expressions (for reasons of government), thus deriving long vowels to be headed (i.e., T-type). While this explains (2c) in that branching nuclei are universally forbidden to precede coda–onset clusters by the Binarity Theorem (Kaye et al. 1990),[5] it does not explain (2a–b) without extra stipulation.

Moraic accounts usually assume that L-type nuclei are mono- and T-type nuclei bimoraic. Hammond (1999) derives (2a–b) by stipulating that syllables contain exactly two moras. However, this requires ad-hoc adjustments such that the final consonant in *lip* counts as moraic while that it *limp* must not, leaving unclear to what extent (2c) is actually explained. The account also requires problematic assumptions about ambisyllabicity to allow words like *bitter*, *bigot*, *busy* etc., where the first syllable would otherwise not meet Hammond's requirement.[6]

Summing up then, previous accounts capture different behavioural aspects of the T-type/L-type, but never the full range. In this paper I want to submit that a unified explanation is in fact possible. In a way, Wells's dichotomy free/checked makes a lot of sense once defined clearly and interpreted in the right way, i.e., as a particular structural configuration. Other aspects, in particular the distributional properties in (2) and durational differences, can then be derived easily from that configuration, and so can quality (to a lesser extent, though). We will turn to that proposal now.

[5] T-type nuclei do occur before coda–onset clusters under very special conditions (*fiend, paint* etc.), cf. Harris (1994) for an account, Pöchtrager (2010; 2013) for discussion and an alternative proposal, and also section 4.

[6] For discussion of the problematic nature of ambisyllabicity, cf. Harris (1999).

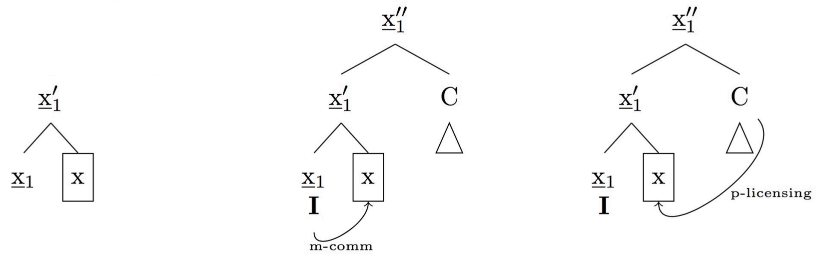
3. The proposal

The key idea that I want to put forward is given in (3).

(3) T-type and L-type nuclei have the same basic structure (a nuclear head and a complement) but differ in what constituent makes use of the complement.

This will become clearer in (4a), which shows the basic scaffold for a stressed nucleus: a nuclear head $\underline{x}$ and its complement x, which together form $\underline{x}'$, a projection of the nuclear head. This builds on the general structure of vowels developed in Pöchtrager (2018). Those two positions will be relevant for the expression of the T-type/L-type difference. To be more precise, the fate of the nuclear complement, highlighted by boxing in all the structures in (4), will determine the nature of the nucleus. (4b–c) illustrate the vowel and the final consonant in *beat* [biːt] and *bit* [bɪt], two words that only differ in the nucleus (T-type vs. L-type). The word-initial onset and the final empty nucleus (Kaye 1990) are left out, as they are not relevant here. The nuclear head $\underline{x}$ is annotated with the element **I** as we are talking about a front vowel, and the consonant *t* sits in the specifier position of that same nucleus.[7]

(4) a. basic b. T-type [iː] c. L-type [ɪ]



In [iː], the T-type nucleus in (4b), the nuclear head $\underline{x}$ claims the complement x by m-command, a relation that ensures that the m-commandee receives

[7] For arguments that the *t* sits in the specifier, cf. Pöchtrager (2006). The exact position of *t* is not relevant to the discussion; what is of interest is only whether there is a consonant following the nucleus and whether it is involved in p-licensing, discussed in just a moment. In a word like *city*, the entire sequence -*ty* would sit in the position where the consonant sits in (4b–c), i.e., the consonant *t* in *city* would be further removed from the stressed nucleus (Pöchtrager 2006), but again, what is crucial is that there is a consonant at all.

the same interpretation as the m-commander.[8] In contrast, in L-type [ɪ] (4c) the complement is not m-commanded by the nuclear head. Instead, it gets licensed by the following consonant (marked "C"; its internal structure is irrelevant here) and remains silent. That licensing relationship has to be such that it will not lead to the consonant being longer, i.e., it cannot be a relation of m-command from the consonant to the nuclear complement, since the final consonants in *bit* and *beat* are of equal length, though one follows an L-type nucleus, the other one a T-type.[9] I will assume that it is a p-licensing relationship (Pöchtrager 2006), i.e., it serves to silence the position in question. As such, the p-licensing relationship extends the Empty Category Principle that forms a central part of GP (Charette 1991; Kaye et al. 1990; Kaye 1990; 1995). Whether the nuclear complement is m-commanded by its head or p-licensed by the following consonant is a lexical property: the difference between *bit* and *beat* is not predictable but lexically stored.

Note that the nuclear complement could not be claimed by both neighbours at the same time, but in sections 4–5 we will look at whether it could be left unclaimed by either one. This leads to a more general question: to what extent can the properties of [iː] and [ɪ] in (4b–c) be generalised to all T-type and L-type nuclei? Will there always be an m-command relationship from head to complement in the T-type and a p-licensing relationship from the following consonant in the L-type? The answer is: not quite. While it will be true that in an L-type there will always be an influence from the outside (from outside $\underline{x}'$, the basic scaffold), we will see that in T-type nuclei more can happen than m-command from head to complement. In other words, while the L-type is relatively uniform, the T-type is more varied. This also lines up with our observations about the existence of diphthongs in (1b).

The proposal in (3–4) has several interesting consequences, all of which seem to be on the right track. Firstly, as we said, an L-type nucleus can be characterised by reference to an outside influence: in (4c) p-licensing from a following consonant. This is very similar to Wells's notion of 'checking'. But notice that in the present account we do not need to add an extra

[8] See Pöchtrager (2006) for details. M-command is similar to spreading in autosegmental accounts, except that m-command, unlike spreading, does not depend on there being melody in the m-commander.

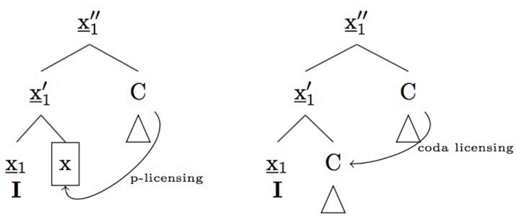
[9] But cf. Polgárdi (2012) for a similar proposal, where she argues that consonants following L-type nuclei are "virtually" long, which can actually be heard (i.e., length is no longer virtual but real) in Welsh English. It is unclear how this extends to other varieties.

stipulation like checking, rather we make use of what we already have and reinterpret checking in the language of our theory: nuclei have internal structure, and part of that structure can be claimed by a following consonant. The ungrammaticality of L-type nuclei finally (**b*[ɪ]) and in hiatus position (**l*[ɪ]*o*) follows, since for such forms we would need a consonant after the vowel in question in order to p-license the nuclear complement, and there is no consonant.

Secondly, the durational difference between T- and L-nuclei (a ratio of about 3:2) follows as well. Since L-type means that part of the nucleus is claimed by a following consonant (and, in (4c), made inaudible), while that is not the case in the T-type, it is unsurprising that the former are shorter than the latter. The longer duration of the T-type as a whole can be read off the structures employed.[10] As pointed out in section 2, this must be kept separate from other factors playing a role in the duration of English vowels, most notably the effects of a following consonant, as in pairs like *bid/bit* (both of the L-type) or *bead/beat* (both of the T-type), where the second member of each pair has a vowel duration that is nearly twice as long as that of the first member. This is discussed extensively in Pöchtrager (2006), which the present account elaborates on. Wells's criticism of the terms long/short stems from the neglect to keep those two factors apart.

Thirdly, the ban on T-nuclei preceding coda–onset clusters can also be derived. Consider (5b), a simplified representation of *limp*, in contrast to (5a), which repeats the basic structure of [ɪ] with a single final consonant (as in *lip*) from (4c).

(5) a. *lip*: L-type plus single C b. *limp*: L-type preceding coda–onset cluster



[10] A reviewer points out that the durational difference follows, but not the exact ratio. This is correct, but there is no claim that exact ratios can be predicted. (In any theory, as far as I am aware.) The possibility of predicting the exact durational realisation of any given x-slot would be a welcome result, but this seems non-trivial, as several factors are involved: the position of the x-slot within the foot, whether it is part of a consonant or a vowel, whether the slot is a head or not—these are just some potential parameters that come to mind.

While T- and L-nuclei share the same basic structure, L-type nuclei are characterised by a following consonant exerting influence on the nuclear complement, as we saw above. In (5c) we see a further example of this. The coda consonant *m* of *limp* sits in the position of the nuclear complement, while *p* sits where all final consonants sit: in the specifier. In earlier versions of GP the final onset *p* was assumed to govern the coda *m*.[11] In this cluster, the requirement on homorganicity is taken as visible proof of that relationship. Expressed in terms of GP 2.0, there must be a relationship of m-command to express the sharing of the element **U** for labiality. Of course, not every coda–onset cluster will display homorganicity, but still there will be some conditions on its members, and the final onset will exert some influence on the coda, cf. Kaye's (1990) Coda Licensing. Yet again we see that the L-type can be defined in such a way that there is an outside influence on the nuclear complement.

Another way to think about the problem, possibly even more straightforward, is in terms of space. In a sense, L-nuclei are the "shorter versions" of T-nuclei. Hence, if part of the structure is taken up (by a coda), only L-nuclei are possible: *l*[ɪ]*mp*, **l*[iː]*mp*.

4. Extending the vowel space

So far we have only looked at one T-type/L-type pair, the front high vowels. Replacing the element **I** by **U** in (4) gives us the back high vowels of *put/boot*. The discussion of **non**-high vowels will require a short detour into a basic assumption of GP 2.0, viz., the non-existence of the element **A**. In earlier versions of GP non-high vowels and coronal consonants were characterised by **A**,[12] which enjoyed the same combinatorial possibilities as any other element. In GP 2.0, **A** is replaced by structure, based on the observation that **A** behaves unlike the other elements and interacts in subtle ways with structure, thus suggesting that it is itself structure (Pöchtrager 2006; 2010; 2013; 2018). Since the T/L-distinction is interpreted here as a structural one as well (in terms of the nuclear comple-

[11] For references on the definition of coda within GP, cf. fn. 4. Within GP 2.0, most instances of government have undergone heavy revision; the reader is referred to Pöchtrager (2006, chapter 6) for proposals on substantive conditions on coda–onset clusters; the representation of the coda in (5) is in line with those proposals.

[12] Within GP, it is not universally accepted that **A** encodes coronality, cf. Pöchtrager (2013) for discussion and justification, and Backley (2011) for an opposing view. The claim that **A** represents coronality seems to originate with Broadbent (1991) and can also be found in Cyran (1997); Goh (1997); Ploch (1999), and many others.

ment), it is to be expected that there will be interactions with **A** yet again. This is borne out by the facts, and the analyses of T/L-nuclei as structurally different and **A** as structure support each other. We must keep in mind that in L-type nuclei the nuclear complement is claimed from the outside, but not in T-type nuclei. As we saw towards the end of the last section, one can also rephrase this in terms of space, with the L-type being "smaller" than the T-type. (The same amount of space is provided, but it is divided up differently.) English does not normally allow T-nuclei before coda–onset sequences (*apt*, **aypt*/**eept*/**oopt*),[13] for reasons given in the previous section: there is simply not enough room. But that limit can be exceeded if the final consonants are coronal: *fiend*, **fiemp*, **fienk*; *count*, **coump*, **counk*. Similar and also more complex observations can be made for German, Finnish, and Hungarian. Those findings were interpreted in Pöchtrager (2010; 2013) to mean that **A** should be replaced by a structural configuration comprising several positions, with only part of that structure used up. As a result, coronal consonants are fairly big, but much/most of that structure is unused and can therefore provide the necessary positions required for T-vowels. Thus, T-vowels can appear in cases where they otherwise could not, as in *fiend* [fiːnd]. Coronal consonants have unused space inside them, and a preceding vowel can "borrow" that unused space.

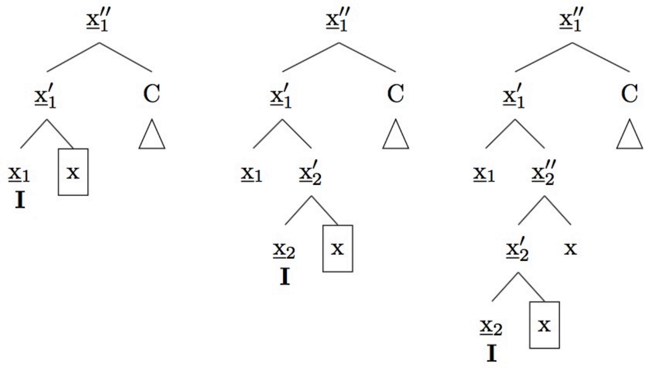
Replacing **A** by more (empty) structure also has repercussions for vowel quality, leading to low(er) vowels being reúinterpreted as structurally bigger. This lends itself to a natural account of vowel reduction, explored in detail in Pöchtrager (2018), where vowel reduction is uniformly interpreted as the loss of structure: for example, while Italian has both mid-open (bigger structure) and mid-close vowels (smaller structure) in stressed position, only the smaller structure (mid-close) survives in unstressed position.[14] Yet another rationale that could be adduced is inherent durational differences between different degrees of openness, with high vowels inherently shorter than mid ones and mid ones shorter than low ones (Lehiste 1970). This falls out from the representational format.

With this in mind, we are now ready to combine the insights on **A** and the proposal on the T/L-distinction in (6), which illustrates various degrees of opening for the English front vowels (all annotated by **I**).

[13] Cf. fn. 4 for references on coda-onset sequences.

[14] I avoid phonetic symbols on purpose. The vowels in question are usually transcribed as [ɛ] and [e], but the transcription is no guarantee that they are to be treated as L-type vs. T-type, cf. sections 1–2. Rather, they are to be treated as differences in openness, cf. Pöchtrager (2018) for a detailed analysis.

(6) a. high, e.g., [ɪ/iː] *bit/beat* b. mid, e.g., [e/eː] *bet/bate* (Northern) c. low, e.g., [æ/ɛː] *bat*/(*bare* or) *bad* (NYC)



(6a) equals the structures discussed before; the fate of the nuclear complement will decide between T- and L-type. (6b–c) are non-high and thus require more structure in the shape of another nuclear head ($\underline{x}_2$) and its projection embedded within the first. This is the replacement of old **A**.[15] As in (6a), the use of the nuclear complement (boxed) in (6b–c) will determine whether it is a T- or an L-nuclei, only this time it is the nuclear complement of the lower nuclear head that is relevant, as the complement to the upper nuclear head is itself complex (it is a projection of $\underline{x}_2$).

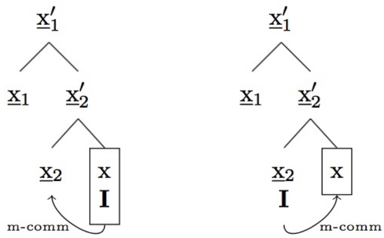
Given that many T-vowels in RP are diphthongs (while L-vowels are not) it is not always easy to find matching pairs that differ only in whether they are T- or L-type. Consider (6b): the difference between [e/eː] as in *bet* vs. (Northern, non-RP) *bate* lies in whether the nuclear complement (sister to $\underline{x}_2$) is p-licensed by a following consonant (*bet*) or m-commanded by the nuclear head (*bate*). This raises the question which of the two nuclear heads is doing the m-commanding, and I will assume that the head in question is $\underline{x}_2$, i.e., the head that the crucial position is a complement of. (6b) also raises the question how to deal with RP [eɪ], but before addressing this we complete the discussion of height: (6c) gives [æ/ɛː], the first of which, an L-vowel, we find in *bat.* Its T-counterpart [ɛː] here, can be found in RP in words like *hair* or *bare*, but not before a final consonant (except for morphologically complex forms like *bare-d*). It has a wider distribution

[15] Simply saying that **A** is replaced by more structure is somewhat incomplete. The difference between (6b) and (6c), both of which used to contain **A**, shows that the amount of empty structure corresponds (roughly) to whether **A** was assumed to be the head or not. (6c) represents the maximal structure; cf. Pöchtrager (2018) for details.

in some North-American varieties, e.g., New York City English, where we also find it (in overlong quantity) in words like *bad*.[16]

Let us finally address varieties with [eɪ] instead of [eː], such as RP. Building on the representation of diphthongs presented in Pöchtrager (2015) and Živanović & Pöchtrager (2010), I will assume that a diphthong has a non-head position annotated with melody.[17] In the case of [eɪ] the nuclear complement of the lower head is annotated with **I**. This gives us the offglide transcribed as [ɪ]. The first part of the diphthong is also a front vowel, so the **I** must be passed on from the complement to the head via m-command relationship (the equivalent of **I**-spreading). (7a) illustrates this, contrasting it to (7b), repeated from (6b), but without the final consonant. Both structures contain two empty positions as we expect from something that is essentially mid (recall that the amount of empty structure is assumed to correspond to openness); they differ in where the melody sits. What is peculiar about RP is that many T-nuclei, like [eɪ], have their melody outside the head position.

(7) a. [eɪ] b. [eː]



To what extent can differences in quality be derived from the structures? After all, [ɪ] is a bit more open than [iː]. It is tempting to blame this on the existence of an m-command relationship from the nuclear head in [iː] vs. its absence in [ɪ]. It is easy to show, however, that this will not work, at least if transcriptions like *bet* [e] and (Northern) *bate* [eː] are to be trusted. The first one does not involve m-command from the nuclear head, the second

[16] See Pöchtrager (2006) for a formal definition of overlength and Kaye (2012) for discussion of that particular vowel quality. For the vowel in *bare* Wells (1982) writes [ɛə] but notes that there is really only very little movement. Other authors have [ɛː]. We will come back to centring diphthongs and monophthongisations in section 5. Note furthermore that the structure in (6c) has another unused position, the specifier of the lower nuclear head, whose role is still unclear to me.

[17] At least those with a peripheral vowel in offglide position; centring diphthongs will be discussed in section 5.

one does, but the difference in quality seems small or non-existent. Given common disagreement on the exact quality of vowels, even within the same variety (cf. section 1 on *bet* with [e] and [ɛ]), it is unclear whether predicting exact qualities is a goal worth pursuing. I will remain agnostic about this issue for the time being, and move on to something more positive, i.e., further properties that can be teased out of the basic T/L-structure.

5. Hiatus and schwa

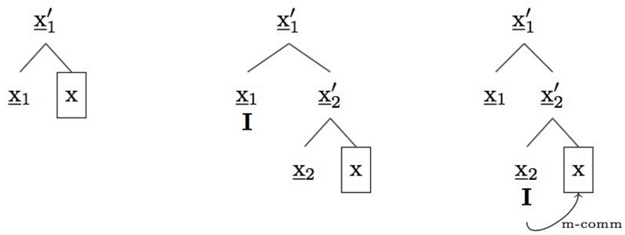
In section 3 the ungrammaticality of a form like **l*[ɪ]*o* was derived from the lack of a consonant following the L-nucleus. This, however, cannot be the full story. Preceding schwa, things look somewhat different: alongside *th*[iːə]*tre* we find *th*[ɪə]*tre*, and Wells (1982, 215) points out that bisyllabic *id*[ɪə] has more or less completely replaced older trisyllabic *id*[iːə].[18] The reference to syllable count is crucial here: while *th*[iːə]*tre/id*[iːə] is a true hiatus (two separate nuclei in a row), *th*[ɪə]*tre/id*[ɪə] have a vowel sequence that has changed from hiatus to diphthong (*one* complex nucleus). The former forms follow from the current proposal (*th*[iːə]*tre* parallels *l*[iː]*o*), the latter, which are really centring diphthongs, will be addressed imminently.

The same issue occurs before *r*, even in non-rhotic varieties where that *r* has been lost and only the vowel sequence remains, as in *fear*. Wells (1982, 213ff) gives a diachronic account, whereby vowels preceding the rhotic first underwent breaking (giving rise to vowel plus schwa) and then pre-schwa laxing (his term). The result is a form like *fear* [fɪə], which rhymes with *id*[ɪə]. Not only do they rhyme, they also undergo the same phonological phenomena: today, both (non-rhotic) *fear* and *idea* can show intrusive *r* before a following vowel.[19] This makes clear that the origin of the schwa is irrelevant synchronically and that the rhyming portion is identical in both words. How to represent [ɪə] then? Consider (8).

[18] Since there is (at least inter-speaker) variation for *theatre* but possibly no longer for *idea*, I do not assume that [ɪə] and [iːə] are related by synchronic rule, but that for some words and some speakers they are variant forms stored as such in the lexicon. Whether this is the result of sound change or due to other factors is irrelevant here.

[19] I follow Broadbent (1991), who derives the emergence of *r* from the non-high quality (expressed by old **A**) of the final vowel of the first word (as in *idea_ r_ of*), and do not assume floating material as Harris (1994) did. Under the present account, intrusive *r* must be treated as the spell-out of an empty onset position claiming (some of) the empty positions that characterise the final vowel of the first word.

(8) a. basic b. [ɪə] (*fear*, *idea*) c. [eː]



(8a) repeats (4a), the basic structure for the T/L-distinction. (8b) shows the proposal for [ɪə]: the higher head x_1 is annotated with **I** and the complement to x_1 is complex, it contains a projection of x_2. This projection of x_2 by itself (i.e., the combination of x_2 and its complement) is the representation of schwa (Pöchtrager 2018). (8b), in comparison to (8a), expresses the idea that [ɪə] is basically a high vowel with a schwa embedded into it, which captures the acoustic facts and also makes sense from a diachronic perspective. The **total** result is a structure similar to a close-mid vowel, and thus (8c) repeats the representation of [eː] as in (6b) in order to illustrate another crucial difference, viz., the position of the element **I**. While (8c) has it in the lower head, we find it in the higher head in (8b). That higher head in (8b) does not m-command anything, neither does the lower head. This explains two things at the same time: firstly, the complement of the lower head x_2 is not claimed by either head in any way. In addition, in *idea* or (non-rhotic) *fear*, there is no consonant following the nucleus in question, so that there can be no p-licensing from a following consonant, either.[20] The question then is, what happens if positions are not claimed? I submit that in accordance with the Empty Category Principle (ECP; Charette 1991; Kaye 1995) those positions must be phonetically realised. Following Pöchtrager (2018) I take [ɨ] to consist of one empty position that gets spelled out, while [ə] consists of two. This is the fate of the complement of x_2 and x_2 itself; they both get spelled out as schwa by the ECP. Secondly, we have a structure here where the complement to the higher head x_1 is somewhat self-contained and spells out as [ə]. This is similar to the scenario where there is a coda consonant, i.e., (5b). Given that x_1 is annotated with **I** and does not m-command anything, it will be realised as

[20] While *fear/idea* etc. have no consonant following, there is no reason to exclude one in principle. English has monomorphemic sequences of [ɪə]C as in *beard, weird, pierce*, but C is limited to coronals. This is unsurprising given their historical origin as superheavy rhymes (cf. the discussion in section 4) but also shows that even if there *is* a consonant, p-licensing is contingent, not necessary.

a short [ɪ], the same length (and quality) as if there was a coda consonant. Wells's "pre-schwa laxing" is a direct consequence of our representation: since we have put schwa in the position necessary to express the T-type, the first part of the centring diphthong will make do with the space left. Since x_1 is realised as [ɪ] and x_1 and its sister as schwa, we get [ɪə] as the realisation of the entire structure.[21]

6. A speculation on stress

That we find a wider range of vowels in stressed position than in unstressed position is common in phonological systems, and has been used as evidence for the internal structure of vowels for a long time in GP (Harris 1994; 1997; Harris & Lindsey 1995). The evidence has been interpreted in different ways in order to back up the presence or absence of certain elements, headedness relationships, length etc., cf. Backley (2011) for a helpful overview. Pöchtrager (2018) attempted a structural reinterpretation, arguing that the amount of structure available in stressed position is the uniting factor: stressed nuclei have more structure at their disposal, and thus allow for a wider range of contrasts to be expressed. (As an example, Italian was mentioned briefly in section 4.) This harks back to an idea in the CV-offspring of GP, according to which stress could be represented as additional structure, cf. Larsen (1995) for one of the earliest references.

English also limits its unstressed position, where we find (i) schwa (*sofa*) but also some (ii) other, non-central vowels (*happy*, *into*...). As Wells (1982, 165) points out, the identification of unstressed with stressed vowels is "usually [...] debatable", but takes, for example, RP *happy* as ending in

[21] A reviewer inquires about speakers monophthongising [ɪə]. The result is similar to a long L-type [ɪː]. I stress "similar" because I do not believe those monophthongs to be of the L-type. L-type means that (i) there is a following consonant to license an empty position, or, if not, (ii) that those empty positions get spelled out in some way. Neither option applies in a final **monophthong**. I assume that the monophthongisation result of [ɪə] comes about by x_1 m-commanding the complement of x_2 in (8b), leading to a *T-type* monophthong that is acoustically similar to a very close [eː], but still different from it (formally by the position of **I** in the tree). Incidentally, this is also equivalent to the representation assumed in Pöchtrager & Kaye (in preparation) for high +ATR vowels, which often get confused with close-mid vowels. This will have to be explored further in future work. Finally, the vowel in *bare* (clearly T-type) is a parallel case. (6c) presents it as a monophthong [ɛː], with justification given in footnote 16. If it is rather to be treated as a diphthong [ɛə], it would have to have the element **I** in the higher head with part of the lower structure spelled out as schwa.

[ɪ], equating it in its acoustic properties (though not necessarily in its structure) with an L-type nucleus.

How would that fit in with our account? In order to express [iː], two positions would be required: a nuclear head m-commanding its sister, as we saw in sections 3–4. If we assume that being in unstressed positions means that the head has no sister, the T/L-distinction becomes inexpressible there. That a head (with no sister) annotated with **I** is spelled out as [ɪ] is exactly what we have seen elsewhere. One could take this even further and propose that the position allowing for the expression of the T/L-contrast **is** in fact the formal representation of stress, similar in spirit to what many CV-accounts do, and thus simply absent in positions lacking stress.

Tempting as that might be, it is easy to see that this is too radical a move. For one thing, schwa requires two positions (section 5), and if the unstressed position banned complements to nuclear heads altogether, we would have a problem representing *sofa* etc. In addition to that, we do find certain T-nuclei (or at least, what seem to be T-nuclei) in unstressed position, as in *yellow*, *potato* etc. What is interesting, though, is that in unstressed position we find high and close-mid vowels (schwa counts as such as well, cf. Pöchtrager 2018), but no open-mid or open ones. Given that openness is expressed by structure in the present account, as the equivalent of old **A** (section 4), it does make sense to say that the unstressed position is limited in structure. The challenge for future work will then be finding out which position exactly is inaccessible/removed in the unstressed position in English.

7. Conclusion and outlook

This article proposed an underlying similarity between T-type/L-type (commonly: tense/lax) vowels. It has certain similarities to the account of Polgárdi (2012), which, however, was not couched within a general account of the internal structure of vowels. There is also a certain parallel here to Kaye's (1990) Coda Licensing, which requires a coda to be licensed by a following onset. (Hence there are no final codas, only final onsets followed by an empty nucleus.) Stressed L-type nuclei are similar in that their occurrence is linked to a following consonant, too. The comparison is not complete, though: firstly, the absence of a following consonant can be "remedied" by spelling out the empty material. Secondly, while Coda Licensing is assumed to be universal, the conditions on T/L-nuclei presented here might not be.

This leads to a pressing question to be addressed in further work: as the system is set up at the moment, it predicts that before coda–onset clusters we should only find the L-type, since the position that is needed for a T-type is taken up by a coda, cf. the discussion in section 3. What does this imply for languages like Spanish, whose vowels are generally treated as equivalent to English T-type nuclei? Even a preliminary look at Spanish vowels in the Oxford Acoustic Database (Pickering & Rosner 1993) confirms that stressed vowels are of the T-type, at least going by quality and in length – independently of whether they occur before a single consonant or before coda–onset clusters. But if they are T-nuclei like in English, how could they be followed by a coda–onset cluster? Either there is something still missing in our account of the T/L-distinction, or there is something about constituent structure that needs clarification (or both). That constituent structure is not fully understood is also clear for English, which **does** allow T-nuclei before coda–onset clusters, albeit under very special circumstances. (Recall the discussion of *fiend* in section 4.)

Despite those open questions for further research, a number of properties of the T/L-contrast could be made to follow from one simple assumption: that both are merely different instantiations of the same basic structure. One can hope that this is a step in the right direction towards a general theory of vowels.

References

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

Bauer, Laurie. 1980. The feature "tense/lax" with special reference to the vowel system of (American) English. Zeitschrift für Anglistik und Amerikanistik 28. 244–253.

Broadbent, Judith M. 1991. Linking and Intrusive r in English. UCL Working Papers in Linguistics, 3. 281–301.

Charette, Monik. 1991. Conditions on phonological government. Cambridge: Cambridge University Press.

Cyran, Eugeniusz. 1997. Resonence elements in phonology. A study in Munster Irish (PASE Studies and Monographs 3). Lublin: Wydawnictwo Folium.

Durand, Jacques. 2005. Tense/lax, the vowel system of English and phonological theory. In P. Carr, J. Durand and C. J. Ewen (eds.) Headhood, elements, specification and contrastivity: Phonological papers in honour of John Anderson. Amsterdam & Philadelphia: John Benjamins. 77–98.

Durand, Jacques and Francis Katamba (eds.). 1995. Frontiers of phonology: Atoms, structures, derivations. Harlow: Longman.

Goh, Yeng-Seng. 1997. The segmental phonology of Beijing Mandarin. Crane: Taipei.

Hammond, Michael. 1999. The phonology of English. A prosodic Optimality-Theoretic approach. Oxford: Oxford University Press.

Harris, John. 1994. English sound structure. Oxford & Cambridge, MA: Blackwell.

Harris, John. 1997. Licensing inheritance: An integrated theory of neutralisation. Phonology 14. 315–370.

Harris, John. 1999. Release the captive coda: The foot as a domain of phonetic interpretation. UCL Working Papers in Linguistics 11. 165–194.

Harris, John and Edmund Gussmann. 1998. Final codas: why the west was wrong. In E. Cyran (ed.) Structure and interpretation. Studies in phonology (PASE Studies & Monographs 4). Lublin: Wydawnictwo Folium. 139–162.

Harris, John and Geoff Lindsey. 1995. The elements of phonological representation. In Durand & Katamba (1995, 34–79).

Kaye, Jonathan D. 1990. 'Coda' licensing. Phonology 7. 301–330.

Kaye, Jonathan D. 1995. Derivations and interfaces. In Durand & Katamba (1995, 289–332).

Kaye, Jonathan D. 2000. A user's guide to Government Phonology (GP). Manuscript.

Kaye, Jonathan D. 2012. Canadian Raising, eh? In E. Cyran, H. Kardela and B. Szymanek (eds.) Sound, structure and sense. Studies in memory of Edmund Gussmann. Lublin: Wydawnictwo KUL. 321–352.

Kaye, Jonathan D., Jean Lowenstamm and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. Phonology 7. 193–231.

Ladefoged, Peter. 1968. A phonetic study of West African Languages (2nd ed.). Cambridge: Cambridge University Press.

Ladefoged, Peter and Keith Johnson. 2010. A course in phonetics (6th ed.). Boston, MA: Wadsworth.

Ladefoged, Peter and Ian Maddieson. 1996. The sounds of the world's languages. Cambridge, MA & Oxford: Blackwell.

Larsen, Uffe Bergeton. 1995. Vowel length, Raddoppiamento Sintattico and the selection of the definite article in Modern Italian. In L. Nash, G. Tsoulas and A. Zribi-Hertz (eds.) Actes du deuxième colloque Langues et Grammaire. Paris: Université Paris 8. 110–124.

Lass, Roger. 1976. English phonology and phonological theory: Synchronic and diachronic studies. Cambridge: Cambridge University Press.

Laver, John. 1994. Principles of phonetics. Cambridge: Cambridge University Press.

Lehiste, Ilse. 1970. Suprasegmentals. Cambridge, MA: MIT Press.

Lindau, Mona. 1979. The feature expanded. Journal of Phonetics 7. 163–176.

Pickering, John and Burton Rosner. 1993. The Oxford Acoustic Phonetic Database. Oxford: Oxford University Press.

Ploch, Stefan. 1999. Nasals on my mind. The phonetic and the cognitive approach to the phonology of nasality. Doctoral dissertation. SOAS, University of London.

Polgárdi, Krisztina. 2012. The distribution of vowels in English and trochaic proper government. In B. Botma and R. Noske (eds.) Phonological explorations: Empirical, theoretical and diachronic issues. Berlin & New York: Mouton de Gruyter. 111–134.

Pöchtrager, Markus Alexander. 2006. The structure of length. Doctoral dissertation. Universität Wien.

Pöchtrager, Markus Alexander. 2010. The structure of A. Paper presented at the 33rd GLOW Colloquium, 13–16 April 2010, Wrocław.

Pöchtrager, Markus Alexander. 2013. On A. Paper presented at A Workshop on Melodic Representation, 12 March 2013, London.

Pöchtrager, Markus Alexander. 2015. Binding in phonology. In H. van Riemsdijk and M. van Oostendrop (eds.) Representing structure in phonology and syntax. Berlin & New York: Mouton de Gruyter. 255–275.

Pöchtrager, Markus Alexander. 2018. Sawing off the branch you are sitting on. Acta Linguistica Academica 65. 47–61.

Pöchtrager, Markus Alexander and Jonathan Kaye. in preparation. The grand unified theory of vowels. Manuscript.

Pulleyblank, Douglas. 2011. Vowel height. In M. van Oostendorp, C. Ewen, E. Hume and K. Rice (eds.) The Blackwell companion to phonology. Malden, MA & Oxford: Wiley-Blackwell. 491–518.

Raphael, Lawrence J. and Fredericka Bell-Berti. 1975. Tongue musculature and the feature of tension in English vowels. Phonetica 32. 61–73.

Wells, John Christopher. 1982. Accents of English 1–3. Cambridge: Cambridge University Press.

Živanović, Sašo and Markus Alexander Pöchtrager. 2010. GP 2.0 and Putonghua too. Acta Linguistica Hungarica 57. 357–380.