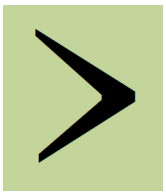


Papers in Historical Phonology
http://journals.ed.ac.uk/pihph
ISSN 2399-6714
Volume 4, 83–135 (2019)
DOI: 10.2218/pihph.4.2019.4192

Phonotactics, prophylaxis, acquisitionism and change: *RIME-xxŋ and *ash*-tensing in the history of English

PATRICK HONEYBONE
University of Edinburgh

Abstract

This article revisits, extends and interrogates the position advocated in Honeybone (2019) — that phonotactic constraints are psychologically real phonological entities (namely: constraints on output-like forms), which have a diachrony of their own, and which can also interfere with diachronic segmental change by inhibiting otherwise regular innovations. I focus in the latter part of the article on the role of one phonotactic constraint in the history of English: *RIME-xxŋ. I argue that we need to investigate the emergence of such constraints in the history of languages and I show how this particular constraint, once innovated (which occurs through *constraint scattering*), can be understood to have inhibited the patterning of *ash*-tensing in certain varieties of American English (and also that it may now have been lost in some varieties). To do this, I adopt a phonological model which combines aspects of rule-based phonology and aspects of constraint-based phonology, and which is firmly rooted in the variation that exists when changes are innovated. Finally, I evaluate the extent to which the type of phonotactically-driven process-inhibition that I propose here involves prophylaxis in phonological change (I show that it doesn't), and I consider the interaction of these ideas with the proposal that all change occurs in language acquisition ('acquisitionism').

1 Introduction

Phonotactics plays a rather minor role in historical phonology. Most diachronic work deals with segmental change, or with (higher-level) prosodic phenomena, such as stress or tone. This article, in contrast, aims emphatically to be a contribution to *diachronic phonotactics*; that is, to the part of historical phonology which deals (i) with the innovation of new phonotactics, (ii) with the loss or change of existing phonotactics, and (iii) with other ways in which phonotactics might interact with phonological change. There has been *some* serious work in diachronic phonotactics, such as Lutz (1988, 1991), Dziubalska-

Kołaczyk (2005) and the work of Ritt and others at Vienna (e.g. Baumann et al. 2016) which also led to a Special Issue of *Folia Linguistica Historica (*number 50:2) on diachronic phonotactics. Additionally, general volumes on historical phonology and/or phonological history do sometimes contain some consideration of relevant matters (as in Hogg 1992 and Minkova 2014). It is not that there is *no* work in diachronic phonotactics, but it is surprising how little phonotactic issues are considered when dealing with cases of phonological change.

For diachronic phonotactics to be worthwhile, we need to be sure that phonotactic entities are real phonological phenomena, and we need a way of integrating them into phonological derivations and into a model which allows for their interaction with other aspects of phonology in phonological change. I discuss that in section 2. Section 3 then focuses on the role of one particular phonotactic constraint in the history of English: *RIME-xxŋ. I consider how it entered the language, how it has interacted with the innovation of other phonological phenomena, and also how it seems to be being lost in certain varieties of the language. Section 4 considers some fundamental implications of the positions on diachronic phonotactics that are proposed earlier in the article, broadening the scope of the piece to address more general questions in historical phonology, such as whether change involves prophylaxis or repair. Section 5 concludes.

This article is a companion piece to Honeybone (2019),[1] where I consider some fundamental issues in phonotactics and in phonotactic change. I do not repeat here most of the points made there, although I summarise (in section 2) the model that I developed there, in order to show how it aims to understand the ways in which phonotactic constraints can interact with other aspects of phonological diachrony (specifically, to account for the way in which already-existing phonotactics can affect changes which are otherwise general in their phonological environment, but which have been inhibited in one very specific phonological environment). I extend here the ideas considered in Honeybone (2019), I consider some of the implications of them, and I show how they also apply to the history of the phonotactic *RIME-xxŋ.

2 Phonotactics in phonology and in phonological change

To start at the start, it is obvious that there are sequences of segments which do not occur in the phonological lexicon of specific languages. For example, [baɪŋ] is not the surface representation of any word in the

[1] This means that, in my mind, that article is Honeybone (2019a) and the current piece is Honeybone (2019b).

variety of English that I speak[2] — it is a 'gap' in the lexicon. It is *not* obvious, however, that speakers *realise* that there are such gaps in the lexicon and that they conceive of them as a genuine property of their language. In other words, a key question for phonology is: are such sequences absent from languages because they are forbidden by the phonological grammar? If so, then this must be due to some type of phonological entity which forms part of a speaker's phonological knowledge. And if so, a whole host of further questions arise: what *kind of thing* are the entities involved (rules, constraints, statistical generalisations...)? over what kind of structures do they *generalise* (words, morphemes, syllables, feet...)? are they all the *same* kind of thing? are they the same or different from the entities that are involved in other aspects of phonology (alternations, non-contrastive distributions, prosodic structure...)? are they categorical or gradient? are they all always equally important in the phonology of a language? what kind of conscious access do native speakers to have to them? and what is the precise inventory of the entities involved for any individual language?

In Honeybone (2019) I set out a model which answers many of these questions by adopting the position (proposed and defended in a considerable amount of phonological work, but not accepted by all) that phonotactics are psychologically real constraints on surface forms. In this paper I modify that position slightly (as not all constraints work on strictly *surface* forms), but the spirit of the position developed here is very much the same as that set out in the earlier article (in fact, the position that I end up adopting here was already suggested there). One crucial point of the model is that it recognises the traditional categorical distinction between different kinds of gaps in languages' lexicons. Some gaps are assumed to exist in languages because the grammar forbids them, while others could easily be filled. In Honeybone (2019), I call the former 'S-gaps', recognising that they have been called a number of things in the literature: 'systemic', 'systematic' or 'structural gaps'. I call the latter 'A-gaps', as they are often described as 'accidental gaps'.

Some work on the type of topics discussed here dissents from the idea that the A-gap/S-gap distinction is a valid one (Algeo 1978 and Bauer 2015 are cautious, for example). In section 2.1, I consider some observations that make most sense if we assume that it *is* indeed a real distinction, and that S-gaps really are enforced by the grammar, a position which I assume in the rest of this article. In section 2.2, I set out

[2] It, or something very similar, may occur in certain other varieties of English, which differ crucially in this one phonotactic. That will be part of the point of section 3.4. The variety that I speak is a generalised northern English dialect of British English.

how I assume phonotactics are implemented in phonology, and in section 2.3, I consider the interaction of phonotactics and diachrony.

2.1 Do S-gaps really exist?

If S-gaps are enforced by the grammar and A-gaps are not, we predict that they should behave differently. For example, A-gaps should easily be fillable (by the adoption of loanwords, for instance), but S-gaps should not (and they should force loanwords to be adapted). Loanword adoption and adaptation are in fact more complex than this simple dichotomy supposes — for example, it is possible for English speakers to produce forms like *carte blanche* with a [ʁ] and a nasal vowel which is not followed by a nasal coda, which is not expected to be possible in most varieties of English, but this involves a kind of code-switching, as an attempt to consciously preserve the original French phonology of the form, importing aspects of French into English utterances (perhaps as part of a phonological 'periphery' which is not allowed to affect the phonological 'core'); also it is surely the case that when speakers of a language are overwhelmed by large numbers of loanwords from a prestigious superstrate language (such that there is pressure to preserve the phonology of the borrowed forms) that this could break their current phonotactic grammar (penetrating into the phonological core of a language) and leading to phonotactic change.

Issues surrounding loanword adaptation deserve serious consideration in their own right, to a degree that cannot be given here (see, for example, Uffmann 2015), but as I discuss in some more detail in Honeybone 2019 (and also, a little, below), the behaviour of loanwords can indeed be important in determining the status of a type of gap in a language: is it forbidden by phonotactics or not? Some phonological work has argued carefully that certain quite noticeable gaps can nonetheless be shown to be A-gaps, rather than S-gaps, and to behave as A-gaps would be expected to behave in terms of loanword adoption (see, for example, Iverson & Salmons 2005 on the English 'tense vowel plus final [ʃ]' gap). This can only be done if the distinction is a real one, such that it allows us to expect that the different types of gap will exhibit different types of phonological behaviour.

A further prediction of an A-gap/S-gap distinction which is often assumed is that we might expect that speakers will have intuitions that forms which violate S-gaps are phonologically 'wrong'. A substantial strand of research has considered the extent to which speakers' judgements about 'wordlikeness' or 'wellformedness' of segmental sequences reflect the gaps that can be found in the phonological lexicon. The results of this research are, however, complex and conflicting, and

it is widely recognised that understanding such judgements requires more of us than simply distinguishing between S-gaps and A-gaps. This need not vitiate the distinction between the two, however. Wordlikeness judgements are (like all reflections on linguistic intuitions) not *grammaticality* judgements (because grammaticality is a property of a linguistic system which is not directly accessibly to speakers), but are rather acceptability judgements which are likely to be influenced (as well as by some reflection of our knowledge of phonological structure), by our knowledge of other things that we might be aware of in some sense, such as lexical frequency and lexical neighbourhood density. The study of acceptability judgements in syntax has developed to a sophisticated level in interpreting how informants react to judging linguistic data and has shown that the interpretation of such results can be complex (see, for example, Schütze to appear and Sprouse to appear), so we should not expect phonological wordlikeness judgements to simply reflect a categorical distinction between S-gaps and A-gaps.

Some of those who have worked on related issues argue that intuitions about which gaps are A-like and which are S-like are simply due to statistical generalisations by speakers over the items in their lexicon. If this is true, there is no role for categorical phonological statements.[3] However, much evidence indicates that "a simple model based purely on statistical properties of the linguistic data is inadequate, just as one based purely on phonetic biases would be", as Albright (in preparation) argues.[4] Gorman (2013, 88) adds that "there are many static phonotactic constraints which are statistically reliable

[3] There are thus claims in the literature that judgements of phonotactic well-formedness are entirely gradient, which would bring the categorical S-gap/A-gap distinction into question. However, as Gorman 2013 shows, there are reasons to be sceptical of such results, especially of the conclusion that they might be evidence against the existence of multiple specific synchronically-active phonological generalisations. For example, as Gorman (2013, 16) emphasises, Chomsky & Halle (1968, 416-417) already point towards a model which aims to "define the 'degree of admissibility' of each potential lexical matrix" beyond a simple two-way (grammatical vs ungrammatical) distinction, on the assumption that, while [bnɪk] is an impossible surface representation in English, [bznk] is taken to be "even less English-like than [bnɪk]." In comparing two sequences like this, which are both absent from English, the point is easy to accept because [bznk] violates multiple phonological requirements of English (onset sequencing, presence of a vowel in a stressed syllable), whereas [bnɪk] only violates one (onset sequencing). In addition to this accumulation of violations by whole strings of segments, it is not impossible that some phonologically-enforced gaps are more serious than others.

[4] Albright is arguing in favour of there being a role for universal phonotactic knowledge (which does not derive from the statistical tendencies present in the phonological lexicon of a language) in understanding speakers' judgements about both attested and unattested phonological forms in a language. A similar line of argumentation has been set out persuasively by Berent (2013, and elsewhere).

but synchronically inert" — that is, the statistical generalisations that speakers *should* make over their lexicon (because they can be shown to exist if we analyse the lexicon of a language computationally) do not in fact replicate the kinds of differences that speakers actually make in terms of the kinds of phonological behaviour predicted above for A-gaps and S-gaps, so we still require a distinction between the two.

It strikes me that understanding phonology (like other aspects of language) involves a delicate journey to work out how several kinds of thing influence both what comes out of our mouths and what judgements we report when questioned on intuitions — many of these things are grammar-external, and may involve usage-based generalisation, but some have all the hallmarks of grammar-internality. Shatzman & Kager (2007) and Lentz & Kager (2015) are examples of careful work probing wellformedness judgements which shows that both categorical phonotactics and probabilistic knowledge play a role and that the two are of different natures. If this is right (and I assume that it is), we need the notion of the S-gap, and of psychologically real phonotactics to enforce them. In this article, I adopt a robust position that phonotactic constraints are indeed phonologically real, and go on to investigate the possibilities that this gives us for understanding phonological change.

It is worth nothing that other work which assumes fundamentally compatible ideas to those adopted here has not always talked of 'phonotactics' — the ideas existed before the word was coined, and it is not always used to discuss them now. Bloomfield (1933), for example, includes some detailed consideration of relevant matters in a chapter on "phonetic structure" (so this is a case where 'phonetic' really means 'phonetic-and-phonological'), and this kind of terminology was also used elsewhere in earlier writings, as in Kruisinga's (1943) detailed volume on the subject, which is called *The Phonetic Structure of English Words*. Hockett (1955, 92) talks about "the distributional classification of consonants" and describes "the phonologic system: a stock of phonemes (or phonologic units) and the arrangements in which they occur relative to each other" (1955, 14). The extent to which such earlier work is describing *what occurs* (in terms of gaps) rather than *what can possibly occur* (that is, assuming that they are S-gaps) is sometimes moot but the idea that phonotactics forbids things is by no means exclusive to generative material (where it is, however, robust). For example, Abercrombie (1967), writing the 'British phonetic tradition', talks about how "[s]uch structural regularities in the phonology of a language produce in its speakers, deep-rooted habits of speech which are difficult to change. This is shown by the way new words introduced into a language — slang, trade names, borrowing

from foreign languages — conform to the existing structural patterns." This is not far from saying that, if the structural regularities of a language avoid a specific sequence, these 'habits' forbid such sequences in loanwords.

The *Oxford English Dictionary* (*OED*) has attestations of the actual word *phonotactics* from 1956 onwards, and refers to Hill (1958) who attributes the invention of the term to Robert Stockwell's use of it in a lecture in 1954.[5] The term *phonotactics* was not used in early core work on phonology from the generative paradigm (such as Halle 1959, 1962, and Chomsky & Halle 1965), but this does not mean that relevant ideas were not important in it. For example, Chomsky & Halle (1965) made famous the distinction between [blɪk] as a non-occurring but admissible form and [bnɪk] as a non-occurring but "inadmissible" form,[6] and Harms (1968, 85), in a textbook introducing standard generative phonology written right at the point of initial pre-eminence of the model, talks of "sequential constraint rules" which "describe many of the same phenomena traditionally treated by phonotactic statements" (thus acknowledging the notions involved, but distancing himself from the term).

The standard terminology of early generative work talks of 'Morpheme Structure Rules'. These include 'sequence structure rules', which enforce what I have been calling S-gaps. Stanley (1967), an influential article in this sphere, explains that "sequence structure rules, though included in the grammar to characterize redundancy, provide as a by-product a characterization of the notion 'possible morpheme'; this obviates the need for a separate set of statements to characterize this notion". Stanley (1967) was influential in the move within generative phonology away from the use of rules to describe phonotactic generalisations towards the use of constraints (which he called 'Morpheme Structure Conditions', which include 'sequence structure conditions'). This is an important development in terms of the model that I adopt in section 2.2. Stanley writes that we can "interpret the set of sequence structure rules as a statement of constraints on systematic phonemic sequences". He explicitly advocated the "need for negative conditions" — that is, constraints which rule out sequences. Stanley also argues for other types of constraints, but it is these 'negative'

[5] Stockwell has confirmed his hand in the creation of the term (Lutz 1988, 231). He was thinking in a context in which Hockett (e.g., 1947) had been using *tactics* to talk of "the study of the relation and arrangement of linguistic units, esp. the study of the arrangement of morphemes" (*OED*). Although Hockett himself abandoned the term in 1955 (16-17), this usage was likely influential on Stockwell's creation.

[6] At this point Halle and Chomsky assumed a simple two-way distinction between what I am calling A-gaps and S-gaps. Chapter 9 of Chomsky & Halle (1968) complicated things (as it often does) as discussed in footnote 3.

constraints which forbid the sequences which form S-gaps, that have taken root as a way of "separating the accidental sequential constraints from the systematic ones" (1967, 401) in phonological theory.[7]

The absence of the term *phonotactics* (but not the ideas that I associate with it here) is also notable in later generative work. For example, Hammond's (1999) volume (of which about half deals with what is here called phonotactics, from an OT perspective) does not use the term — rather, the volume is described as focusing on "surface distributional regularities" and on "the allowable configurations of consonants [and] vowels". There are several traditions of usage which have kept the term alive since the 1950s, however, as in Scholes (1966), Sommerstein (1974) and Singh (1987), some of which are broadly generative and some of which are not. More recently there are signs that it has become a usual term in generative work, too (e.g., Albright 2006, Hayes & Wilson 2008, Gorman 2013). Whatever we call it, the topic in focus here (what we could call 'phonotactic knowledge') is a clear focus for phonology.

2.2 The structure and status of phonotactics

Given the position on S-gaps just discussed, I assume that phonotactics is[8] fundamentally about understanding where the segments of a language can occur (and thus, in part, how they combine in sequences — especially in languages which allow consonant clusters, like English). Once the basic phonological structures of a language and its segmental inventory are understood, the phonotactic question is: do all segments occur in all environments where a language allows segments in principle? I assume here (as in Honeybone 2019, which is the source of all the discussion in this section) that the relevant 'basic phonological structures' of a language are provided by the slots in syllable structure that the language allows, and I assume a classic basic syllable structure of the type given in (1), where 'O', R', 'N' and 'C' stand for 'onset', 'rime', 'nucleus' and 'coda'.[9]

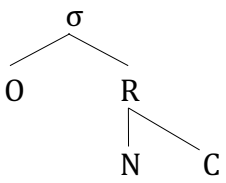
---

[7] Some work insists on using positive constraints (one of the other types that Stanley considers) — for example Taylor's (2002, 250-251) usage-based approach, in which the "grammar of a language comprises only 'positive statements' about what does occur, there is no need for 'negative statements' of what does not occur" — but this is a minority position, not even held by all usage-based linguists.

[8] A specific grammatically-enforced gap in a language is sometimes described as 'a phonotactic' (shorthand for 'a phonotactic generalisation'), so we can ask both what phonotactics *is* and what the phonotactics *are* in any language.

[9] I follow the convention of using the spelling 'rime' for the syllabic constituent, to differentiate it from the poetic notion of 'rhyme', which is not the same thing.

(1)



This assumption that syllable structure is relevant in describing phonotactics is not universally agreed — it is possible to focus on the simple occurrence of strings of segments (as for example Steriade 1999 has advocated), or on describing the segmental structure of words or of morphemes (as is implied in the term 'Morpheme Structure Rule', which cropped up above). However, the slots provided by basic syllable structure have shown themselves to be so useful in describing phonotactic generalisations insightfully that most phonotactic work uses them without comment, or ends up reinventing them by talking about 'word/morpheme initials' instead of 'onsets' or 'word/morpheme finals' instead of codas or rimes.

As mentioned above, I see phonotactics as being implemented through 'static' constraints on phonological forms. The position that I follow here (developed in more detail in Honeybone 2019) is that these constraints apply to surface-like forms, at the 'end' of the phonological derivation, with earlier aspects of phonology (those dealing with 'dynamic' phonological processes) modelled as rules. This mixed rule-and-constraint model follows work in Optimality Theory in placing importance on surface-oriented constraints, but it also follows rule-based diachronic work in allowing the straightforward depiction of (new) processes as the addition of a rule.[10] The model is not new — it is highly reminiscent of classic rule-based work which takes phonotactics seriously, such as Sommerstein (1974), a piece which proposes essentially the model of the grammar that I follow here (and which was described by Goldsmith (1993, 9) as "very prescient"). The interest

[10] Or as other classically-recognised types of change to the rule component of the grammar (following such work as Kiparsky 1968 and King 1969), such as rule loss or rule reordering (and recognising that 'rule reordering' might really be the 'rising' of a rule through the grammar, as in the life cycle of phonological processes). My overall approach echoes Salmons (to appear) in the rejection of OT as "ill-suited to application to sound change" in its most common form, which I take to be the addition of a rule (but it also recognises that the focus in OT on constraints is important, and indeed argues that we need to consider the emergence and loss of the effect of constraints in diachrony, as in historical work on OT). Salmons writes further that "synchronic phonology may well look monostratal but any reasonable historical record of complex sound changes shows obvious patterns of layering of change, the accumulation of stepwise changes" — this requires a derivational model of the type that I adopt. The same position is adopted in Ringe & Eka (2013).

from a diachronic perspective is that this kind of model means we need to consider both the innovation of new rules (and change in the rule component) *and* the innovation of new phonotactic constraints (and change in the phonotactic component). This may seem a complex model, but it strikes me as exactly what we need in order to account for everything that we might want to account for in phonology, and in phonological change. I give a number of derivations along these lines in Honeybone (2019), and I work though one example in section 2.3.

As background to that, we can recognise that one example of a psychologically real (language-specific) phonotactic constraint in Present-Day English is the fact that sequences like [pn, bn, fn, θn, tn, dn, kn, gn] are absent in onsets. A phonologically-informed observer (and, we might expect also, a learner) can easily see that the basic structural facts of English allow two-member complex onsets with [p, b, f, θ, t, d, k, g] as the first member and a coronal sonorant as the second (as in *brief, through, climb, gloom etc*.), and that nasals can occur in onsets in principle (as in all nasal-initial words); it is also the case that no universal constraint (such as the Sonority Sequencing Principle) rules out combinations involving obstruents and nasals, so this must be a fact about the phonology of English. This is the kind of S-gap that needs to be enforced by a phonotactic, and I represent such phonotactics as exemplified in (2). In this case, 'T' stands for 'all and only the non-sibilant obstruents',[11] and 'N' stands for 'all and only the nasals'.

(2) *ONSET-TN

a sequence of a non-sibilant obstruent followed by a nasal is forbidden in an onset

Evidence for the reality of this constraint comes from the fact that when words are borrowed into English from other languages which allow such clusters, they are adapted, either by deleting the obstruent or by epenthesising a schwa between the two consonants. For example, the *OED* includes *knackwurst,* with a first attestation of the word in English from 1939). This word is borrowed from German where the first syllable is [knak], but the *OED*'s transcription of the first syllable

[11] In and of itself, this constraint does not rule out onset sequences involving the English affricates, such as [t͡ʃn], but I assume that they are accounted for by a more general constraint on the combinability of affricates, given that they do not occur in any complex onsets. The behaviour of [s] in onset phonotactics needs some sort of special treatment, as is well known, given the initial strings in words like *string* and *skill*, which violate sonority-sequencing if they are viewed as onsets. I do not adopt any particular analysis of this here (many have been proposed), but I take for granted that some sibilant-specific phonology accounts for this, and that 'non-sibilant obstruent' is a natural class.

for the word in English is [nak]. The case of *gnu* is instructive: the precise etymology of this borrowing into English is complex — the *OED* gives 'Khoekhoe' and 'San' and its origin, but other authorities consider other languages to have also been involved (as part of a transmission chain), including Dutch. It is clear from the spelling, however, that the word in the donor languages is assumed to have had a complex onset involving [n], and it is also clear that the word is not pronounced with a complex onset involving [n] in English. The 'classical' English pronunciation according to the *OED* is [nju:] or [nu:], but the *OED* also recognises the 'jocular' (but in fact common UK) pronunciation [gənu:], which originates from a comic song (Flanders & Swann 1957), where the humour in part derives from the fact that the initial ‹g› is actually pronounced as [g]), and it is notable for our purposes that the [g] can only be pronounced with a following [n] if a vowel intervenes between the two (and, of course, the vowel that is used for this is schwa, which is the minimal English vowel).

Other phonotactics considered in Honeybone (2019), which will likely be familiar to those who know the phonology of English, include *CODA-h '[h] cannot occur in codas' and OCP-ONSET(CORONAL) 'a sequence of two anterior coronal segments (such as [tl]) is forbidden in an onset' (again, ignoring [s]). There are many more. I assume that gaps need to be *plausible* and *systematic* in order to be S-gaps and hence implemented in the phonological grammar as phonotactic constraints, and that they should not replicate the basic structural facts of a language (which may be language universal, as in what is possible in basic syllable structure, as given in (1), or may be accounted for by some form of parameterised principle, such as whether a language allows complex onsets or not). The precise definition of these notions are not necessarily simple, but I think they make sense. It is *plausible*, for example, that English would allow *ONSET-TN-violating clusters such as [pn] and [dn] because other languages allow them and English allows similar things, but nonetheless English *systematically* excludes them, on the basis of the natural classes of (non-sibilant) obstruents and nasals.[12] It might seem to be cheating to use capital letters to refer to a set of segments in the representation of phonotactics like in (2), as I could in principle define the set involved in any way that I choose, but this is not the intention — the capital letters are used to represent natural classes. I assume therefore that learners recognise phonological gaps (which involve plausible systematic gaps) in their lexicon as S-gaps, and that they assume that they are due to phonotactic constraints.

[12] See also Honeybone (2016) for a further consideration of 'plausibility' in phonological generalisations.

I discuss some more of what it takes for a gap to be an S-gap in section 3.1. I turn now to consider some aspects of the diachrony of phonotactics. If we take such S-gap-enforcing constraints to be as phonologically real as segments and syllables and rules, then we need a serious branch of diachronic phonotactics which investigates their innovation and loss, as well as change in phonotactics (as I go on to show in the next section), and also other ways in which phonotactics might intervene in diachrony (which is the focus of section 3 and 4).

2.3 Phonotactics in diachrony

Phonotactic entities of the type argued for in section 2.2 have a diachrony, just like other phonological entities (such as segments and foot structure). Part of the point of this article is that we should take this seriously — interesting results can emerge if we do so. It is clear that phonotactics can be innovated into (or 'become active' in) a language, and that they can be lost over time. It is also the case, however, that phonotactics already in a language can be observed to change. For example,[13] the constraint as precisely formulated in (2) has not always been part of the phonology of English. The full diachrony of the case involves more interesting developments than simply the addition of a phonotactic, however.

Firstly, we need to separate out stop-nasal and fricative-nasal clusters. While there is reason to believe that the current phonotactic affects the whole class of (non-sibilant) obstruents, as assumed in (2), the history of fricative-nasal onset clusters requires some separate thought. Initial [sn] and [sm] have been possible in English for as long as we can know, so this has not changed: for example, words like *snow* and *smear* descend through transmission from forms which are robustly reconstructed with initial [sn] and [sm] in Proto-Indo-European. In early stages of English, however, other fricative-nasal initials were possible, too: the onset sequence [fn] is firmly attested in Old and Middle English. While it has only ever occurred in a few words — Bosworth & Toller's (1898-1921) dictionary of Old English and the *Middle English Dictionary* (*MED*) both have 11 headwords with initial ‹fn› — there is no doubt that it was phonotactically possible: these two sources include, for example, words like *fnesan* 'to sneeze' and *fnæd* 'border, fringe'. The full history of the phonotactics that are relevant here deserves serious attention, and is not fully understood (perhaps because so few [fn]-initial words are involved). We know that [fn] was possible all the way back to Proto-Germanic (for which Kroonen 2009

[13] This example also comes from Honeybone (2019), but I extend the discussion of it here.

reconstructs *fneusan~fnūsan* 'to sneeze' for example), and we know that [fn] is no longer possible in English, and that the loss of this may have led to the emergence of the full form of the constraint in (2), but the precise manner in which [fn] was lost deserves serious investigation of a sort that would take us off track here. I think that this *is* something that we should consider in historical phonology, and I think it will likely be insightful if we see the diachrony of [fn] in this phonotactic context, but because it is still somewhat unclear, I set fricative-nasal onset clusters aside in this discussion and focus on stop-nasal onsets. I therefore redefine 'T' in this section to mean 'all and only the stops' (rather than 'all and only the non-sibilant obstruents', which is was it meant in section 2.2). If we do this, we have access to an intriguing understanding of the diachrony involved in the loss of [kn] and [gn] onsets in English: the constraint forbidding complex onsets with nasals was simpler after the change than before it.

We know that Proto-Germanic, Old English and Middle English (PGmc, OE and ME) did not have onset sequences of the type [pn, bn, tn, dn] (which are those that we would expect if *any* complex onsets are possible with nasals as the second member, given that [n] is coronal, which is the unmarked place of articulation — the languages, in fact, lacked all kinds of stop-nasal onsets involving [p, b, t, d]).[14] This can be established with a fair degree of certainty by searching full dictionaries of the languages for words beginning with reconstructed sequences of these types in PGmc, and letter sequences of the type ‹pn, bn, tn, dn› in OE and ME, on the assumption that the much more phonemic spelling of OE and ME would show these clusters in written forms if they existed, and that we would expect such onsets to show up word-initially if they are at all possible in a language. Kroonen (2009), Bosworth & Toller (1898-1921) and the *MED* are extensive dictionaries of these three stages of the language, and none of them feature any words that can be interpreted as featuring initial sequences of those types.[15] This is a fair

[14] I focus in the discussion here on clusters with [n] as they are phonologically the most likely, but what I say also holds for onset clusters with other nasals. No stop has ever combined with [m] or [ŋ] in an onset in English, as far as I am aware, and this may be due to a long-lasting fundamental phonotactic that requires the second element of an onset to be coronal when it is a non-glide sonorant (thanks to Ben Molineaux for discussion of this). Where relevant I also searched for combinations with [m]/‹m›, and also found nothing, so I assume that this is not a potential complicating factor in what follows.

[15] Apart from a few scribal errors or variants of forms which normally have a vowel, such as the ME forms *tnykyllere* which is likely an error for *tinkler* 'tinker', and *pniger*, which is likely an error for *winiger* 'vinegar' (with what looks like a ‹p› in fact possibly a late attestation of the letter wynn — thanks to Ben Molineaux for this suggestion).

proxy for establishing the systematic absence of [pn, bn, tn, dn]-type onsets.

On the other hand, the sequences [kn] and [gn] were perfectly possible as onsets in PGmc, OE and ME. All three dictionaries list multiple words beginning with these sequences (for example, the *MED* has 140 headwords beginning ‹kn› and 34 beginning ‹gn›), as in words such as *knee*, *know*, *gnat*, *gnaw*, which even now retain the spelling of these earlier sequences. We can be sure that they had initial stops up till Middle English at least because alliterative poetry from that period alliterates words with [kn] onsets with those with [k] onsets, as Minkova (2003, 313) shows (for example *cneouwen* ‘knees’ alliterates with *king* in *Lagamon’s Brut* from around 1200). On the other hand, early pronouncing dictionaries such as Spence (1775) and Walker (1791) list several words with initial ‹kn› and ‹gn›, and all are transcribed as being [n]-initial in their pronunciation, so the stops had clearly been lost in these onsets by the late 18th century.

If we consider the gaps involved in the three early historical stages considered here (PGmc, OE and ME), we can recognise that, while [kn] and [gn] are possible, and [pn, bn, tn, dn] are missing, other complex onsets involving all stops (like [pl, dr, kl, gr]) are entirely possible at all three stages, and all this means that the [pn, bn, tn, dn] gaps could *plausibly* be filled. In addition, the sets of segments involved in the gaps are *classes*: the second member of the S-gap is ‘all and only the nasals’ and the first member is ‘all and only the labial and coronal stops’. The latter class can be expressed, but any system of features will struggle to express it simply. I return to this point below.

If we assume that there is an S-gap in early English, banning sequences of labial-or-coronal stops and nasals, we predict that the effect of the constraint should be visible in the history of English. There is some evidence that it was. Thus, for example, the early loanword *pneumatic*, borrowed from Greek and/or Latin (for which the *OED*’s first attestation is from 1624), is nasal-initial in Present-Day English, and is already nasal-initial in Spence (1775) and Walker (1791) — while English spelling often preserves the orthography of a donor language, there is no evidence that there was ever an initial [p] in this word, thus it may well be that it was adapted at the point of its borrowing into English, which may just have been while velar-stop-nasal onset were still allowed. The pattern of adapting loanwords which violate this kind of phonotactic by deleting the stop is certainly robust in earlier stages of English — as another example, Baldwin (1846) transcribes the Russian river *Dnieper* as nasal-initial in English, too. While the source languages for these loans allowed labial and coronal

stop-nasal onset clusters, the initial stop was lost in English, likely as soon as the words were borrowed into the language.

In terms of the dating of the loss of the initial stops in [kn] and [gn] onsets, there is agreement in the literature that were lost by the end of the 17th century (there is some dispute over the precise dating — see Minkova 2014, Lass 1999 — Lass argues on the basis of comments on pronunciation from the period that the key change for our purposes began in the 17th century). The precise stages involved in the change may in fact be different in different parts of Britain, and the change considered here likely involved some briefly-existing intermediate stage between stop and deletion, but it can be set out in simple segmental form as in (3), with the full-stop/period indicating a syllable boundary. The *phonotactic* point is that this innovation was accompanied by (or can equally be seen as) something like the change given in (4). Any phonotactic from earlier stages of English, banning onsets combining [p, b, t, d] with a nasal, would have been relatively complex: no matter how it is formulated with phonological features, *ONSET-{T–k,g}N or *ONSET-{p,b,t,d}N is more complex than *ONSET-TN.

(3) k, g > ∅ / . __ n

(4) *ONSET-{T–k,g}N > *ONSET-TN

The change in (4) expresses the diachronic phonotactics involved in the change, and it misses part of the point to ignore this (and to only consider (3)). Honeybone (2019) speculates that we might even perceive a pressure to simply phonological generalisations, as in (4), as part of the explanation for the change. This is reminiscent of productive work which has aimed to establish if there is directionality in change in terms of how one set of rules can change into another, or in terms of how a rule can change if it stays in a language but changes its structural description or environment, as in Kiparsky (1968, 1971) and King (1969), for example (and see Ross 2011 for a cautionary investigation of possible patterns in rule change). We should be cautious with such speculation — such structural pressures cannot be seen to *cause* any particular change, but it may be that they constrain which changes are possible in any phonological state.

In any case, if we understand the change in hand here as just discussed, we can see it as a case of phonotactic *change*, rather than the addition of a new phonotactic. Indeed, we are called to consider which of (3) or (4) 'came first', or whether they are in fact separable at all. I foresee criticism that this approach risks a 'duplication' dilemma if we can model a change as having both an effect in the rule component and

the constraint component. I would counter that failing to consider both risks missing a point, and while the approach makes things more complex, it also makes them more interesting (and, I think, more likely to be closer to the truth). If we take phonotactics seriously in phonology at all we need to consider these issues related to phonotactic change (and to other types of change in the phonotactic component of the grammar).

The main point of Honeybone (2019), however, is that there is even more to the understanding of diachronic phonotactics than this: we also need to recognise that already-existing phonotactics in a language can inhibit newly innovated segmental changes. By this, I mean that there are cases of segmental change which fundamentally have a straightforward environment of application, but where the signature of a phonotactic is visible, explaining a complexity in its patterning.

One example of this that I deal with in Honeybone (2019) is the role of a constraint like that in (5) in inhibiting an aspect of Late Middle English syncope, which Luick (1914–40) assumes was completed by 1500 (what I call *Spätmittelenglischer Schwund* 'late Middle English loss', following Luick).

(5) OCP(SIBILANCE)
a tautosyllabic or tautomorphemic sequence of sibilant segments is forbidden

The phonotactic in (5) forbids sequences of the six sibilant segments of English: /s, ʃ, z, ʒ, t͡ʃ, d͡ʒ/. We can see that it is necessary in Present-Day English because, while affricates may not combine easily with other segments, sequences of other fricatives are perfectly possible (that is, the basic structural facts of the language allow them) — thus [sf] occurs as an onset sequence, as in *sphere, sphinx, sphincter*,[16]

[16] Chomsky & Halle (1968, 416) in fact assume that words with initial [sf], such as *sphere*, are ungrammatical ("inadmissible") despite surviving perfectly well in a phonological derivation. While an intriguing idea, this would be surprising given their long presence in the language. These words, while all loanwords, have been in English for a considerable time (doubtless loaned and reloaned from learned sources several times), and we might expect them to have been adapted by now if they are problematic phonologically: the *OED* gives a robust number of attestations of all three words (with ‹sph›) from the middle of the 16th century, with much earlier attestations, too. Interestingly, the very first attestations of *sphere* and *sphinx* are spelt *spere* (from before 1300) and *Spynx* (1420-1422), both with ‹sp›, which may show that English has changed in this regard and did not then allow initial [sf] sequences so that the words were adapted when first borrowed to have [sp]. The earliest spelling in the *OED* of *sphere* and *sphinx* with ‹sph› (indicating that [sf] then survives the reloaning process) are from around 1533 and 1579, respectively, just after the late ME period in question here.

as may also [sθ], as in *sthenic* and *sthenia*; medial tautomorphemic sequences of fricatives are possible, thus [sf] occurs in *asphalt* and *blaspheme*, [fθ] in *diphtheria* and *ophthalmology*, and [sθ] in *aesthetic* and *anaesthesia*; and also final sequences of fricatives such as [θs] and [fθ] are possible, as in *meths* and *twelfth*. None of these kinds of combinations occur when the fricatives are sibilants: no initial clusters occur (unlike in Polish, for example), word-medial sibilant clusters always indicate a morpheme boundary in English, with one sibilant of either side of it (as in *misshapen* and *newssheet*), and word final sibilant clusters are absent.

There is good evidence, therefore, that the phonotactic in (5) currently regulates the combinability of sibilant segments in English. Most of the six current sibilants have been in English since at least the Middle English period (all except /ʒ/ which emerged in Early Modern English), and there is evidence that their combinability has been constrained in at least some of the ways described (5) since then, too. The diachrony of this phonotactic is an interesting question in its own right, and while I think it is the type of a question that we can and should pursue, I lack the space to do that fully here. As I hint in footnote 16, it may well be that OCP(SIBILANCE) has changed over the centuries in its precise formulation, and in its interaction with other phonotactics. Depending on how gemination is analysed (if geminates involve clusters, they are relevant), the 'tautomorphemic' aspect of (5) may have changed, too. Thus, for example, the ME word *kissen* 'to kiss' had a medial geminate (which was underlying — in the base — unlike present-day cross-morpheme derived geminates), and this kind of structure was possible in English until the loss of gemination (the completion of which Lass 1992 dates to around 1400). All that is crucial for my purposes, in fact, is that there was an S-gap in the lexicon due to the inability of sibilants to cluster in final sequences — I need to assume that this was active in Late ME. Absent clear evidence that the constraint was different from the formulation in (5), however, I simply assume that the LME phonotactic has the same form as the current one.

Honeybone (2019) sets out the way in which a phonotactic like (5) interacted with the innovation of part of Late Middle English syncope (*Spätmittelenglischer Schwund* 'SpSchw'), which is one of the ways in which English has lost unstressed vowels. This is a case of (post-tonic) syncope, which targeted the last unstressed vowel in a word — the part considered here is that with the specific segmental conditioning given in (6), which seems to have been regular. As shown in the syllabic representations, this involved the loss of a syllable.

(6) CV́C<u>V</u>z > CV́Cz σ́σ > σ́
CV́CVC<u>V</u>z > CV́CVCz σ́σσ > σ́σ

Some example forms relevant to understanding the change are given in (7), gathered from Honeybone (2019), to which I have added acute accents to indicate stressed vowels. The first three show orthographic forms of morphologically simple originally disyllabic words which underwent the syncope (showing how spelling change records aspects of the change), the remaining forms are plurals, represented in part following the conventions that Lass (1999, 142) adopts to discuss this case, using [V] to represent the "weak vowel" which syncopates when the rule applies and otherwise semi-orthographic forms, which allow us to fudge the issues surrounding the nature of the tonic vowels. Given that the pre-syncope form of the regular plural morpheme was /-Vz/ the environment for the syncope was met in all regular plurals. The second three forms in (7) are from Lass (1999), giving the ancestors of *cats*, *dogs* and *kisses*, and the remaining forms area added to show that not just disyllabic forms were affected. The left-hand column gives forms as they occurred before the change was innovated into the language, and the right-hand column gives forms that existed once the syncope had stabilised (the syncope can feed laryngeal assimilation, as shown in *cats*).

(7) *Témys* > *Thámes*
álmis > *álms*
áddis > *ádze*

kátVz > kátz [→ káts]
dógVz > dógz
kísVz ... kísVz

*máiden*Vz > *máiden*z
*héaven*Vz > *héaven*z
*físher*Vz > *físher*z
*ábbess*Vz ... *ábbess*Vz

The absence of syncope in *kisses* and *abbesses* (indicated by the absence of '>' and the same forms before and after the change) is the important thing to note. It makes sense if we assume (i) that the key change was the introduction of this case of syncope as a phonological rule, which applied to every occurrence of [-Vz]; (ii) that, like all changes, this new rule of syncope was initially variable; and (iii) that it was inhibited due to interaction with OCP(SIBILANCE). The 'SpSchw' rule of syncope can be understood as in (8), although this is really simply a

description of the process. The rule is partially prosodically defined, targeting the final unstressed vowel in a word, so a linear rule format like (8) does not show this insightfully, but it will do for exemplification.

(8) $V \rightarrow \emptyset \; / \; C\_\_ z\#$
[unstressed]

The interaction between the rule and the phonotactic is shown in (9), which sets out the model proposed in Honeybone (2019) to understand cases like this. It shows synchronic derivations for three representative words once the syncope (*SpSchw*) had been innovated (as a variable rule), and assuming that morphology has concatenated the base and the plural morpheme before handing the forms to phonology. The left-hand derivation for each word shows what happens when the rule applies and the right-hand derivation shows what happens when it does not, so two possible derivations are present in (9) for each word. Both derivations are grammatical for words like *dogs* and *heavens*, but the derivation when the rule applies to *kiss* (and other words which end in a sibilant) is rendered ungrammatical once it enters the phonotactic component (which I represent at the end of the rule component, surrounded by a box to flag it up) because it violates OCP(SIBILANCE), and so it cannot surface. Even if laryngeal assimilation applies after syncope (as in *cats* in (7)) as is likely, the derivation is not rescued as the violation of OCP(SIBILANCE) remains. The alternative derivation of *kiss*, without the application of *SpSchw*, can surface without problem. This means that at this point in the language, there would have been two possible surface forms for words like *dogs* and *heavens*. As the syncope stabilised and ceased to be variable, it underwent rule inversion, to give the situation that remains in Present-Days English (as discussed further in Honeybone (2019)).

(9)

	kisses		*dogs*		*heavens*	
	/kis+Vz/	/kis+Vz/	/dog+Vz/	/dog+Vz/	/hevən+Vz/	/hevən+Vz/
SpSchw	kisz	—	dogz	—	hevənz	—
OCP(SIB)	*	—	—	—	—	—
	*	[kisVz]	[dogz]	[dogVz]	[hevənz]	[hevənVz]

A central point of Honeybone (2019), which I repeat and reinforce here is that we can understand the change involved in the introduction of (8) only if we recognise that the phonotactic in (5) intervened to affect its patterning. The 'exceptions' to SpSchw, such as *kisses* and *abbesses* are not due to the patterning of the syncope itself, but are due to the fact that the rule was innovated into a grammar which had a

phonotactic which ruled out forms with a sequence of sibilants. It is thus important for this model to work that OCP(SIBILANCE) was already in the phonology when the syncope was innovated. Surface forms in which the syncope applies are prevented from occurring because — if they did — they would violate a phonotactic. This has resonances with the notion of prophylaxis in diachrony: that is, a language avoiding change in order to prevent a 'problem' of some sort (such as a constraint violation). I return to this point in section 4.

The model adopted here was shown in Honeybone (2019) to also allow for an insightful analysis of the patterning of the innovation of Mid-Scots θ-debuccalisation, and I go on to show that it allows us to understand other cases of change, too, in section 3. The basic model combines aspects of rule-based and constraint-based phonology in the way that it models synchrony, and this is a crucial part of the way that it explains diachrony. While it is derivational, the model has certain characteristics that are reminiscent of Optimality Theory. The existence of constraints in the model, evaluating forms that are close to the surface is an obvious parallel with OT, but the assumption here is that any constraint violation is fatal for a form, unlike in OT but typical in pre-OT work on phonological constraints. Another parallel to OT is that the model assumes that (in cases of change) there is more than one candidate surface form, and that some of these forms can in principle be ruled out by the grammar. The model adopted here has only two candidates, however (rather than the infinite number in OT), which both can surface, and this number of candidates is determined by something else in the model (the fact that a rule can either apply or not when it is first introduced).

The conclusions of this whole section are: that there is a lot for diachronic phonotactics to consider, that we can investigate the history of phonotactics seriously, and that some diachronically and phonologically interesting ideas emerge if we do so. As well as the innovation and loss of phonotactics, and change in already existing phonotactics, we can recognise that that language-specific phonotactic constraints on phonological forms can inhibit otherwise regular changes. This latter point gives us a model for the interaction of phonotactics and segmental changes. I retain the focus on this in section 3, to show that it can account for the patterning of further cases of the innovation of phonological phenomena.

3 *RIME-xxŋ in English synchrony and diachrony

The kind of lexical gap discussed at the start of section 2 of this article is well known. I return to it here to consider it in detail. It is clearly true

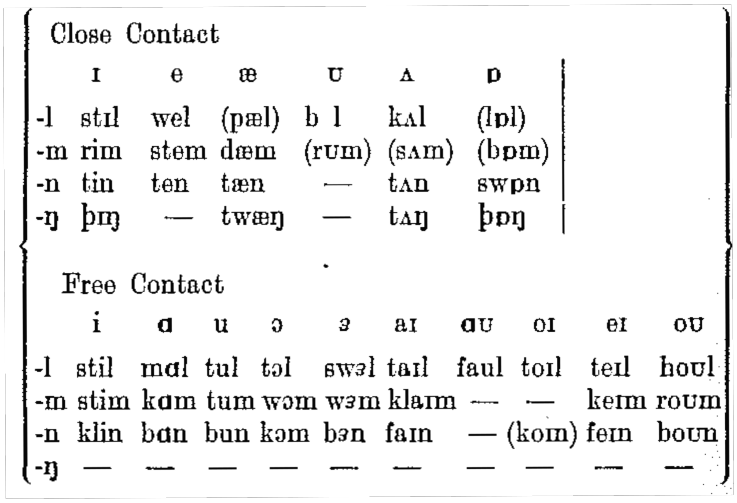
that [baɪŋ] is a gap in at least some forms of English, as I mention there — native speakers of such varieties agree that there is no such thing in their lexicon, and searches of electronic phonological databases find nothing. For it to be phonologically interesting as a gap, given the discussion in section 2, it needs to be plausible (that is, it must be plausible that we might expect to find words with the structure in question because the structure fits in with the 'basic structural facts' of the language) and systematic. [baɪŋ] is clearly plausible for most English forms of English, for example: [b], [aɪ] and [ŋ] occur as segments in these varieties, and there are surface representations of words which are similar: [baɪt] is *bite*, showing that [baɪ] is a possible start of a word; and [bɪŋ] *bing* is a possible sequence — it has, in fact, been used in a range of ways in English (as a new word), for example as the name of a children's character in a series of books and a television series (see, for example, Dewan 2003) — showing that monosyllabic morphemes that end in [ŋ] (and start in [b]) can be freely added to the language. I show in section 3.1 that [baɪŋ] is an S-gap — it is forbidden due to a systematic phonotactic (at least in most varieties of English). I consider a little of its history in section 3.2, and I go on to show in section 3.3 that it has acted in an inhibitory way in diachrony, as we saw for OCP(SIBILANCE) in section 2. In section 3.4, I consider how robust the relevant phonotactic is — I show that in most varieties it *is* robust, but is not surface-true, which has implications for where phonotactics apply in phonology; I also show, however, that it might not be robust in some other varieties.

3.1 *RIME-xxŋ in English

There has been some detailed work on the phonotactics of English which sets out a range of generalisations about the phonological gaps that exist in the language. In a remarkable early volume on the topic, Kruisinga (1943) shows (among much else) that [baɪŋ] is a gap in English for a robust and systematic reason. His table which shows this is reproduced in (10), taken from Kruisinga (1943, 54). This shows all the monophthongs of (General British) English (also known as 'RP') and whether they can be followed by single sonorants (in forms with initial consonants). The table uses '—' where no words with a particular phonological shape can be found and an example word where they can be found (and "when a word is between parentheses, this means that the type occurs in a few words only"). Vowels in 'Close Contact' are the

lax/short/checked vowels, and vowels in 'Free Contact' are the tense/long/free vowels.[17]

(10)



Close Contact

	ɪ	e	æ	ʊ	ʌ	ɒ
-l	stɪl	wel	(pæl)	b l	kʌl	(lɒl)
-m	rim	stem	dæm	(rʊm)	(sʌm)	(bɒm)
-n	tin	ten	tæn	—	tʌn	swɒn
-ŋ	þɪŋ	—	twæŋ	—	tʌŋ	þɒŋ

Free Contact

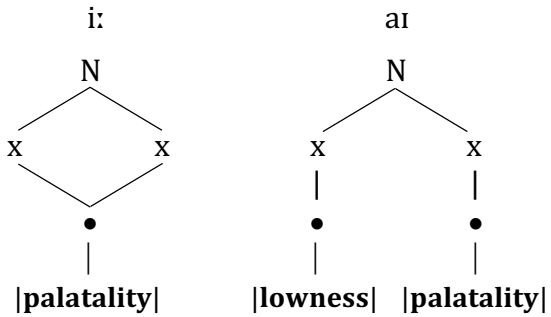
	i	ɑ	u	ɔ	ə	aɪ	ɑʊ	oɪ	eɪ	oʊ
-l	stil	mɑl	tul	tɔl	swəl	taɪl	faul	toɪl	teɪl	hoʊl
-m	stim	kɑm	tum	wɔm	wəm	klaɪm	—	—	keɪm	roʊm
-n	klin	bɑn	bun	kɔm	bən	faɪn	—	(koɪm)	feɪn	boʊn
-ŋ	—	—	—	—	—	—	—	—	—	—

The systematicity of the gap in question is clear from the absence of examples in the bottom row: there is no word in English which has an [ŋ] at the end if preceded by a diphthong or a tense vowel.[18] This represents a fundamental S-gap in the phonology of English, showing that the absence of [baɪŋ] is due to a constraint that involves a whole natural class of vowels: those which can be described as tense/long/free. It has often been proposed that the constraint involved can best understood in prosodic terms. It makes sense to see this as a constraint on possible rimes, expressing which vowels cannot combine with [ŋ] in a rime. It is not crucial how we express the natural class of vowels involved (only that there *is* one). I adopt the autosegmental convention of representing these vowels using two rimal x-slots, assuming such representations as those given in (11), where the units in vertical slashes at the bottom are intended as privative featural specifications (following a convention used in Honeybone 2001, 2005). Tense monophthongs have two x-slots linked to their one melody, and diphthongs have one melody linked to each of their two x-slots.

[17] Kruisinga uses some now outdated but interpretable conventions, for example ʊ = ʊ and þ = θ. He also uses [e], for the lax DRESS vowel — I use [ɛ] for this in Present-Day varieties of English elsewhere in this article, following a different convention. The first and last rows of words in (10) contain inconsequential typos — the fourth word in the first row should be [bʊl], and [kɔm] should be [kɔn].

[18] This holds quite fundamentally of the English lexicon, and any candidate words which violate it are marginal in some sense (e.g. *boing*, *oink* are clearly onomatopoeic).

(11)



These vocalic representations are essentially those of the 'Dependency/Government' approach (as it is described in Carr, Durand & Ewen 2005), and are shared with certain other approaches, too. In models like this, it is further assumed that each singleton consonant has one x-slot. There is no substantial difference in terms of the ideas covered here between rimal x-slots and moras, which are used for equivalent purposes in much other work — the two notions can be seen as identical for the purposes of this paper (so I sometimes use terms like 'monomoraic' and 'bimoraic' to refer to structures with one or two x-slots, as they are handy terms to use). Diphthongs and tense vowels in the varieties of English in focus here are thus seen as long/heavy at the surface. On this basis, the phonotactic involved in the case in question here can be understood along the lines of (12).[19]

(12) *RIME-xxŋ
a sequence of two x-slots and ŋ is forbidden in a rime

This phonotactic rules out [baɪŋ] and enforces all of the S-gaps identified in the bottom row of (10). The phonotactic is indeed clearly specific to [ŋ] — as the other rows in (10) show, there is no systematic gap involving vowels before other similar rimal consonants, such as [l, m, n]. As well as the systematicity involved in the pre-*ŋ* gaps, it is important for our purposes is that the gaps identified in the bottom row in (10) might very plausibly *not* be gaps in English — this is shown by the fact that most of the other cells in the table are filled.

A number of other ways have been proposed in the literature to model this gap, but I think they fall short. Hammond (1999), for example, assumes that [ŋ] has two moras, and that rimes may not have more than three moras (coronals can be moraless), but it seems arbitrary and *ad hoc* to multiply moras so freely. Jensen (1993), among others, assumes that all relevant occurrences of [ŋ] are underlyingly /ng/, and a late rule of *g*-deletion making the assimilation that derives

[19] It could equally well be represented as *RIME-μμŋ.

them opaque (other occurrences of [ŋ] are derived from /nk/, but the /k/ does not delete). There is much to recommend this approach, as it can allow the phonotactic in question here to be taken care of by a general constraint enforcing a rimal trimoraic maximum, which also rules out forms like [biːlk] and [baɪmp] (assuming a special status for coronals), but there is evidence that it cannot be right. I show in section 3.4 that there is firm reason to believe that at least some occurrences of /ŋ/ must exist in underlying representations. We are left, therefore with *Rime-xxŋ. I say something about how it must have emerged diachronically in the next section.

The table in (10) also shows other gaps, and this is highly relevant to the understanding of the notion of the S-gap developed in section 2. For example, the gap in the top part of the second column indicates that there is no word which ends in [ɛŋ]. This might be taken to show that the rime [ɛŋ] is ungrammatical in English, like [aɪŋ]. However, there is every reason to believe that [ɛŋ] is an A-gap. Other lax vowels occur in the [__ŋ] environment, so we might be cautious in assuming that it is an S-gap. The *CUBE* searchable dictionary (Lindsey & Szigetvári 2013, which includes a vast range of words of English, transcribed for General British) also does not find any word which ends in that way in the native English lexicon, but it gives *ginseng* and *nasi goreng* as ending with [ɛŋ] — these are both loanwords which have been easily adopted into English without adaptation (so, in fact, for speakers which use them commonly, [ɛŋ] is not an A-gap).[20] A similar story can be told about the [ʊŋ] gap that Kruisinga recognises: loanwords like *Samsung* and *Kung* (*Fu*) show that this is not an S-gap. We would predict that the non-xxŋ gaps in (10) are A-gaps, not S-gaps, because they do not feature natural classes of segments (unlike the xxŋ gap), and it seems that this prediction is met.

3.2 Where did *RIME-xxŋ come from?

If we assume that *RIME-xxŋ is indeed part of the phonology of varieties of English which follow the patterning described by Kruisinga (1943), as discussed in the previous section, diachronic questions arise: has English always been like this? and what would it look like if this situation changed? These are big questions, and especially the former

[20] There are also unambiguously English forms like *strength* and *length*, which end in [ɛŋθ]. This shows that [ɛŋ] is possible as *part* of a rime. It may even show that [ɛŋ] is possible as a full rime, if the final [θ] is shaved off from consideration because it realises a separate morpheme to the base (which itself ends in [ɛŋ]), but even considering evidence like this, [ɛŋ] does not occur finally in the native English lexicon in free morphemes.

would require serious study of the kind discussed in section 2.3. I do not explore this in detail here. I think that we *should* engage in such explorations and I hope to have shown that they are both necessary and possible, but that would be a job for a whole separate paper. Some parts of the diachrony of *RIME-xxŋ are relatively clear, however, and I set them out here. (I discuss some aspects of the latter question in section 3.4.)

A lot of serious work has been carried out on the length (or 'quantity') of Middle English vowels (see, for example, Luick 1898, 1914-40, Lass 1974, 1992, Minkova 1982, Ritt 1994, Minkova & Stockwell 1996, Bermúdez-Otero 1998 and Lahiri & Dresher 1999), and while it is clear that there were considerable lengthenings and shortenings in this period (some of which were prosodically driven, and some of which were driven in part by segmental environment in a way which is relevant to our purposes, as in Homorganic Lengthening and Shortening before Consonant Clusters), it is not possible to consider these issues fully here, in part because they are controversial. All this is relevant to our purposes, however, to the extent that English largely settled into a pattern of allowing long/bimoraic vowels in certain environments and not in others, and the constraints on this can be understood at least in part as constraints on rime structure. The situation that developed during these changes was the pre-*RIME-xxŋ state of the language. The *current* situation in varieties like that described by Kruisinga (1943), such as General British, is that rimes may have a maximum of three x-slots (as in *milk* [mɪlk], *lamp* [lamp], *meek* [miːk], *lime* [laɪm], but forbidding S-gaps like *[miːlk], *[laɪmp], [laɪŋk]), although extra segments may occur following such a rime if and only if they are coronals (as in *mind* [maɪnd], *monks* [mʌŋks]), in which case, they are often analysed as being in an 'appendix' — this is all well described in work such as Fudge (1969) and Giegerich (1992). The xxŋ gap is unusual in English as it is a case where the language allows only *two* x-slots in a rime.

If we set coronals aside (because they clearly complicate the picture and are not directly relevant to the development of the velar nasal), I assume, therefore, that the pre-change state (before the innovation of *RIME-xxŋ) was similar to the current situation just described, in that it allowed (i) rimes with a short vowel (= 1 x-slot, monomoraic), such as [i, e], as long as this was foot internal and another syllable supplied another x-slot (on the assumption that the foot was a moraic trochee), (ii) rimes with a long vowel or diphthong (= 2 x-slots, bimoraic), such as [iː, eː, aɪ], (iii) rimes with a short vowel and one consonant (= 2 x-slots, bimoraic), such as [ik, ek, ib, ep, im, em], (iv) rimes with a short vowel and two consonants (= 3 x-slots, trimoraic), such as [iŋk, eŋk, iŋg, eŋg,

imb, emp], and (v) rimes with a long vowel or diphthong and one consonant (= 3 x-slots, trimoraic), such as [iːk, eːp, iːb, eːg, iːm, eːm, aɪk, aɪb, aɪm].[21] The rime is thus fundamentally constrained by the constraint in (13).

(13) *RIME-xxxx
a sequence of four x-slots is forbidden in a rime

A change which is crucially involved in the innovation of *RIME-xxŋ was the 'cluster simplification' of [ŋg]. This change can be represented as in (14), but, as with (3) and (4), that innovation was accompanied by the a phonotactic change, as in (15). There is general agreement that this case of cluster simplification occurred (in stressed syllables — the story of *-ing* is different) during the Early Modern period, around the mid-to-late 16th century, or perhaps even later (see, for example, Dobson 1968, Lass 1999). As Lass (1999) discusses (and as has been picked up by Garrett & Blevins 2009 and Bermúdez-Otero 2011), this change worked its way gradually through the morphological environments in which [ŋg] occurred (following the pattern predicted by the 'life cycle of phonological processes', as Bermúdez-Otero shows). While the change has been lexicalised into underlying representations in most varieties of English (so that there is now /ŋ/, as I argue in section 3.4), Bailey (2018), among others, shows that (14) is still synchonically active (and variable) in certain western dialects in England.

(14) g > ∅ / ŋ __ . (= the introduction of a rule of g → ∅ / ŋ __ .)

(15) *RIME-xxxx > *RIME-xxŋ

Unlike the situation around (3) and (4), it is not the case that (14) can be seen as the 'same thing' as (15), because [ŋg] became [ŋ] after short-vowels as well as after long vowels and diphthongs. This means that we can perceive here an asymmetry between the change in the rule component and the change in the constraint component: the segmental rule can be perceived to have had precedence because it has a wider applicability, so it seems to make sense to see the phonotactic change in (15) as coming in the wake of (14).

[21] Other vowel-consonant sequences and other clusters also, of course, occurred. I only list a few that are relevant here, in order to demonstrate the range of the possible. I use the symbols [i] and [e] for short vowel here because laxing may not have occurred at this point.

The innovation of (14) meant that trimoraic rimes like [iŋg, eŋg] simplified to bimoraic [iŋ, eŋ]. No compensatory lengthening occurred to preserve the x-slot/mora that was lost — quite the opposite: the vocalic forms that the language allowed in this regard froze. No rimes of the type [iːŋg, aɪŋg] were possible before the change (because they would have violated *RIME-xxxx), and they did not become possible after the change, because the phonotactic *RIME-xxŋ was innovated. We could assume that, as [ŋg] was finally reinterpreted by learners as a single segment, the systematic gap that this left was noticed by them and was assumed to be due to a phonotactic, turning the xxŋ gap into an S-gap, as in (15).

A further point to note about the change in (15) is that, while a new phonotactic is added to the language in the change, *RIME-xxxx remains in the grammar, too. Forms like [iŋk, eŋk, imb, emp, iːk, eːp, iːb, eːg, iːm, eːm, aɪk, aɪb, aɪm] are still possible in the language after the change, and are still the *maximum* rimes possible (ignoring the complications of additional coronals), and *RIME-xxxx can do the job of enforcing this maximum in the same way as before — rimes involving [ŋ] could never reach four x-slots, as *RIME-xxŋ would stop them, but *RIME-xxxx does not need to know or care about this and can stay as general and systematically-justified as before the change.

The change involved here is reminiscent of 'rule scattering,' which, as Bermúdez-Otero (2015, based on Robinson 1976), explains, involves a rule rising through the grammar (as in the life cycle of phonological processes), but also leaving its original form behind. This involves a kind of splitting of a phonological phenomenon, both leaving the pre-change form in the grammar and also innovating a new (changed) form. It makes sense in the phonologisation of processes because, as Bermúdez-Otero (2015, 387) writes, "the original gradient process of phonetic implementation remains active in the grammar even after the new categorical rule enters the phonology". While it is not clear that the life cycle is involved in the case of phonotactic change considered here, (15) can be seen as a kind of 'constraint scattering', in which the original constraint remains active, while a new, more restrictive phonotactic has also entered the phonology.

If the origins of *RIME-xxŋ proposed here are correct, we have encountered further aspects of diachronic phonotactics in the process of working to understand the changes involved: constraint scattering can occur in phonological change. This shows again that there is much of interest to consider in this area. The precise details of the pre-change state (and their interaction with ME length changes) would be well worth investigating properly in connection with the issues that I have

discussed,[22] but it is clearly the case that *RIME-xxŋ was innovated relatively recently, during the Early Modern English period, around the 16th century.

3.3 Has *RIME-xxŋ inhibited any segmental change?

Given all the discussion up till now in this article, it will be obvious that I am aiming to answer the question in this section's title. I focus here on a case where there is evidence that *RIME-xxŋ has indeed inhibited the patterning of one well-known segmental change, in a similar way to how we saw OCP(SIBILANCE) inhibit an aspect of the *Spätmittelenglischer Schwund* syncope in section 2.3. The change in question has a number of names in the literature. I call it *ash*-tensing. It is the change which has led to what was originally a low, lax, front, short vowel becoming a raised, tense, often diphthongal, long vowel, as has been well described in American dialectology and sociolinguistics for the Middle Atlantic states of the US (see especially Labov 1994). The change is complex, and, as with the Middle English quantity adjustments above, I cannot

[22] The full story is complex. For example, the history of a word like *thing* from Old English to Present-Day English looks quite straightforward: [θiŋg] > [θiŋ] > [θɪŋ] (if we follow Lass 1999 in assuming that the laxing/lowering of high vowels was late), but there is in fact good evidence (from texts like the *Ormulum* — see Ritt 1994) that it first underwent Homorganic Lengthening ([θiŋg] > [θi:ŋg]), because [ŋg] was one of the clusters that triggered the change, and was then shortened again, perhaps even as part of Shortening before Consonant Clusters (a form of closed syllable shortening), before the introduction of *g*-deletion (and the phonotactic in focus here), so that the full history of *thing* (at least, in some varieties) is: [θiŋg] > [θi:ŋg] > [θiŋg] > [θiŋ]/*Rime-xxŋ > [θɪŋ]. We should also consider the history of the syllable-final [mb] cluster in English, which also simplified to [m]. At first glance, we might expect that this was part of a unified change with the ŋg > ŋ simplification given in (14), but in fact it seems that they were separate events. There is certainly no *RIME-xxm constraint in Present-Day English, as shown by the existence of words like *womb* [wu:m] and *climb* [klaɪm]. This makes sense if work like Dobson (1968) is right in arguing that mb > m occurred much earlier than ŋg > ŋ. Dobson argues that mb > m occurs by 1300, whereas ŋg > ŋ occurs about 1600, meaning that they are entirely unconnected events. It may even be that this took words like *womb* out of the realms of possible application for Shortening before Consonant Clusters because the cluster disappeared before the shortening occurred (if not, [mb] was ignored in the environment of the shortening), so that a possible scenario for the history of the word (once original /a/ had risen to /o/) is: [womb] > [wo:mb] > [wo:m] > [wu:m], that is: the chronological ordering of the possibility for application of shortening and cluster reduction are different for words with syllable-final [mb] and with syllable-final [ŋg], which meant that shortening did not apply in [mb] words, xxm occurs when [mb] is simplified, and no *RIME-xxm phonotactic could be innovated. The scenario sketched out here may not be right, but at least I hope to have shown that the question as to why there is no phonotactic *RIME-xxm in English (when there *is* a *RIME-xxŋ) is an interesting question.

deal with all aspects of its innovation, but certain points in the diachrony of the phenomenon are clear, and it is these that I focus on here. The input to the change has variously been described in the literature as 'short a', the BATH/TRAP vowel, and /æ/ (which is why it is also called 'ash', as that is the name of the letter ‹æ›, and is why the change can be referred to as '*ash*-tensing'). I use 'æ' as the symbol for the input. The output varies from variety to variety, and is transcribed in several ways in the literature (including non-transparent forms like [A], [E] or [æh]). It is often transcribed as a diphthongal form like [eə], and that is what I use as the symbol for the output of the change.

The ultimate origin of the change is linked by Labov (1994), based on earlier work such as Ferguson (1972), to 'BATH-broadening',[23] which in its full extent is a complex set of changes in its own right, with a changing environment and aspects of lexical conditioning. It began in Britain in the 17th century, and eventually turned short, lax, low vowels (originally [a]) into long, tense vowels which are back in many varieties, giving Present-Day General British forms like *bath* [bɑːθ] and *grass* [gɹɑːs]. At the time that this change was taken to Philadelphia and nearby areas in North America in early settlement from Britain, the vowel in these forms was still front (likely [æː]), and Labov's assumption is that it was still instantiated through a rule (æ → æː). The output of this set of interrelated changes has therefore clearly changed, and now varies from dialect to dialect. In the varieties that I consider here, it is typically described as diphthongal, and so (as mentioned above), the relevant part of the change as I will be considering it can be thought of as: æ > æː > eə. What will be important is that the output is long/tense — that is, it is associated to two x-slots/moras.[24]

The phonological environment in which the change occurs (and/or in which the associated rule applies) has also changed and generalised in different ways in different places, and it is this aspect of the phenomenon that I focus on. The patterning currently described for *ash*-tensing is famously lexically-specific in certain dialects (such as Philadelphia English) in, for example, the pre-*d* environment (see, for example, Labov 1981, Kiparsky 1988), such that — if we assume that it is still a rule — it has exceptions, although it is regular in various other

[23] This name is from Wells (1982), using his key-word for the lexical set of words which have undergone the change (which have a vowel which is distinct from the words in the TRAP set). Lass (1999) calls the change 'Lengthening I'.

[24] While we can't read phonology directly from phonetics, Sneller (2019) shows clearly that the duration of tensed vowels is longer than their untensed equivalents, so there is good reason to believe that the vowel in long. This change can this be seen as introducing a further long/tense vowel to the set listed (as in 'Free Contact') in (10), although varieties that have it typically do not have all of the other vowels in that set.

sub-environments. I follow Labov and Kiparsky in assuming that *ash*-tensing can be seen as a synchronic rule, which I represent as æ → eə, and the diachrony that I consider is not the innovation of this rule (or the change in its output), but the way that its environment has changed over time, since it was taken to North America. Sneller (2018, 2019) makes a good case that *ash*-tensing in classical Philadelphia English is a synchronic phonological rule, with a 'tolerable' number of lexical exceptions in terms of Yang's (2005, 2016) Tolerance Principle.

Labov (1994) argues that the pattern in the distribution of the tensed output of *ash*-tensing found in late 20th-century Philadelphia is "close to that of the original short **a** lengthening, and that the New York City distribution was originally quite similar". This gives us a picture in which the rule has generalised, first from the initial 'BATH-broadening' environment which was brought to the area (before syllable-final voiceless fricatives and some anterior nasal clusters),[25] then to a pattern like that found now in Philadelphia (which includes syllable-final anterior single nasals), and then to a pattern found in New York City (NYC), applying gradually in more environments, with the distribution in NYC being highly generalised. It has become conventional, following discussions like Labov (1981, 1994), to represent the environment which triggers *ash*-tensing as in (16), which sets out potentially relevant obstruents and sonorants of English and includes in a box those which trigger the rule when they directly follow the vowel, in syllable-final position.

(16)

p		t	t͡ʃ	k
b		d	d͡ʒ	g
m		n		ŋ
f	θ	s	ʃ	
v	ð	z	ʒ	
		l	r	

The inner box in (16) gives the environment for Philadelphia (where the dashed lines around /d/ indicate that this sub-environment has a considerable number of exceptions) and the outer box gives the

[25] Lass (1999) also includes the pre-r environment in the patterning of this change, but it may be that this is a separate development, linked to other British pre-*r* changes and perhaps even the loss of rimal-*r*, given that it does not fit as part of the pattern of *ash*-tensing in North America.

environment for NYC.[26] This means that *ash*-tensing occurs in all varieties mentioned here in words such as *ham, hand, man, pass, bath*. NYC has a very general pattern, taking in natural classes of stops, fricatives and nasals apart from the fact that only anterior nasals trigger the process: *ash* does not tense preceding a syllable-final velar nasal. Interestingly, Harris (1989) describes a similar situation in Northern Irish English, which has generalised the same kind of *ash*-tensing pattern to that found in the Mid-Atlantic states of the US. He writes that "[i]n Belfast, the tense reflex has a much wider distribution, occurring in all the New York City contexts as well as before voiced fricatives and *l*". This distribution includes the tensing/lengthening before syllable-final lenis stops at all places of articulation, and nasals at the labial and coronal place of articulation, but, as in NYC, explicitly excludes tensing before [ŋ]. In this type of *ash*-tensing, almost any consonant can trigger the process apart from [ŋ]. A featural characterisation of the surface pattern for NYC (and Belfast) would be complex because natural classes of segments are involved *apart from* in the nasals, where only [ŋ] is excluded. It is notable that in Philadelphia, too, syllable-final [m] and [n] trigger the rule, but [ŋ] is excluded. Why is the pre-*ŋ* environment always resistant to *ash*-tensing?

This is the phonotactic point: we can understand the patterning in the generalisation of ash-tensing, as the environment for the rule changed diachronically, if we recognise that the phonotactic *RIME-xxŋ was involved: *ash*-tensing itself generalised to take in full natural classes — it did not care about place in the nasal series — but the existence of *RIME-xxŋ in the constraint component of the grammar prevented pre-*ŋ* tensed *ashes* from surfacing. We can model this in exactly the same format as that set out in (9) as a way of modelling the interaction of OCP(SIBILANCE) and the *Spätmittelenglischer Schwund* syncope, as I show in (17).

(17)

	ban		*bang*	
	/bæn/	/bæn/	/bæŋ/	/bæŋ/
ash-tensing	beən	—	beəŋ	—
*RIME-xxŋ	—	—	*	—
	[beən]	[bæn]	*	[bæŋ]

[26] The full environmental patterning also requires a consideration of paradigm uniformity effects, morphological structure and a few other factors which are not relevant here.

The derivations in (17) are from the point when *ash*-tensing is generalising to apply to full natural classes of segments, including the nasals. At this point, as in the early stages of all such phonological changes, the application of the rule would have been variable, and as is the case in (9), (17) shows synchronic derivations for representative words at this point of variation, with the left-hand column for each word showing what happens when the rule applies and the right-hand column what happens when it does not, so two possible derivations are present in (17) for each word. Both derivations are grammatical for words like *ban*, which features an environment where tensed/lengthened vowels surface in all *ash*-tensing varieties discussed above, but the derivation when the rule applies to *bang* (and other words with syllable-final [ŋ]) is rendered ungrammatical once it enters the phonotactic component (again represented at the end of the rule component, surrounded by a box for ease of recognition) because it violates *RIME-xxŋ, and so it cannot surface. The other derivation of *bang*, without the application of *ash*-tensing, can surface without problem, with the underlying *ash* vowel. We can thus see that *RIME-xxŋ has indeed inhibited a segmental change, explaining what seems otherwise to be an exceptional environmental patterning.

We can understand the patterning of the generalisation of *ash*-tensing only if we recognise that a phonotactic intervened to affect it. *Ash*-tensing did not generalise to produced tense vowels in pre-*ŋ* words such as *bang*, *gang* and *fang*, but this is not due to the patterning of the rule itself, but to the fact that the rule was innovated into a grammar which had a phonotactic which ruled out tense/long vowels preceding [ŋ]. As was the case in the *SpSchw* syncope discussed above, the language seems to have prophylactically avoided enacting the change if it would have violated an already-existing phonotactic to do so. I return to these points in section 4.

3.4 Where does *RIME-xxŋ apply? Is it robust in all varieties?

We have seen that there is good evidence that *RIME-xxŋ constrains the phonology of English — both in the current patterning of the lexicon in terms of varieties like General British, as shown in (10), and from the fact that we can see its effects in diachrony, as in (17). In this section, I consider a number of pieces of data which seem to involve violations of *RIME-xxŋ. In the model that I have adopted here, this is problematic, given that any constraint violation is assumed to be fatal for a form.

Gimson (1962) is one of the classic discussions of English phonetics and basic phonology, giving a detailed description of aspects of General British, among other things. Gimson points out that [ŋ] *can* in fact occur

(syllable-finally) after long vowels, as in forms like those in (18), taken from Gimson (1962, 270).

(18) *I've been gardening.*
[biːŋ]

She'll soon come.
[suːŋ]

Does this mean that *RIME-xxŋ is not a real phonotactic after all? Or has the phonotactic been lost in this variety? We can answer both of these questions with a firm 'no', but this requires a recognition of the wider paradigm in which the phonological model set out above is embedded. Gimson considers this data as part of his discussion of assimilation, and it is clear that the two cases of [ŋ] in (18) are underlyingly /n/, and have assimilated in terms of their place of articulation to the following /g/ or /k/. This recognition alone does not solve the problem, however, given that the model adopted here assumes that phonotactic constraints are evaluated at the end of phonology — a rule like Nasal Place Assimilation must occur earlier in the derivation than the constraint component as it has been set out in examples of the model assumed up to here, so the constraint *RIME-xxŋ should rule out forms like [biːŋ] and [suːŋ] after assimilation has created them. Either the model is wrong, or *RIME-xxŋ does not apply.

The answer to this conundrum is straightforward: the model used up till now needs a refinement, or, rather, it needs to be understood in its proper context. It is a commonplace in many models of phonology to recognise a distinction between lexical and postlexical (or 'phrasal') phonology. This idea was developed in detail in work such as Kiparsky (1982, 1985), which proposed the Lexical Phonology model, but it also has longer roots, and it is currently pursued in Stratal Phonology, in such work as Kiparsky (2000), Bermúdez-Otero (2003, 2015). The key point in terms of the data just discussed is that the relevant assimilation-derived cases of [ŋ], such as those in (18), have been derived from /n/ by a post-lexical process — this is clear because it occurs across words. What this means is that, at the end of the lexical phonological stratum, *been* is [biːn] and *soon* is [suːn], and it is here that *RIME-xxŋ applies — in a constraint component at the end of the lexical phonology. Post-lexical nasal place assimilation then applies in the post-lexical phonological stratum, and all we need to assume is that *RIME-xxŋ is not part of the post-lexical constraint component. This all seems fits togther: if the idea is right that there is first a lexical stratum of phonology, after which words are concatenated into phrases, after which there is a post-lexical phonology, then the model proposed above

would expect there to be a rule component in each, followed by a constraint component in each, and we would expect that the contents of these components will be different in the two strata.

It is clear that *ash*-tensing is a lexical rule: it is word-bound, and the triggering consonant which follows the vowel needs to be in a coda (so potential resyllabification across word-boundaries is irrelevant). Kiparsky (1988) argues strongly that *ash*-tensing is indeed a lexical rule. We thus have evidence that both *ash*-tensing and *RIME-xxŋ are part of the lexical phonology, and everything slots into place if we assume that *RIME-xxŋ applies at the end of the *lexical* phonology, before words are concatenated to produce phrases in the postlexical stratum. *RIME-xxŋ is not surface-true, but it is still a part of the phonology of the variety that Gimson is describing, and it is true at the level at which is applies.

What this means is that the locus of phonotactic constraints is not necessarily 'the end of phonology', as assumed in section 2.2 and elsewhere above, but rather: phonotactics apply at the end of a phonological stratum. The crucial point remains the same: phonotactics apply at the end of a block of phonology, and can rule out certain forms which have been derived, so that they cannot 'escape' from phonology.

Like *RIME-xxŋ, there is good evidence that OCP(SIBILANCE) also occurs in the lexical phonology: by very definition, it does not apply across word-boundaries (cross-lexical sequences of sibilants are possible, and all are cross-morphemic by definition). The extent to which there are *post-lexical* phonotactics is an interesting research question, but the idea that phonotactics can apply at several places in phonology seems the right thing to assume — the patterning of syncope in present-day English implies exactly this. Thus, there are good arguments, from both lexical forms and speaker behaviour when confronted with loanwords, that OCP-ONSET(CORONAL) is a robust constraint of English, as defined in (19).

(19) OCP-ONSET(CORONAL)
a sequence of two anterior coronal segments is forbidden in an onset

This is discussed in some detail in Honeybone (2019), and is based on the facts that [tl], [dl] and [θl] are not found as onsets in English (despite the fact that similar onsets like [tr], [dr], and [θr] are well-attested, and again ignoring aspects of the behaviour of [s]), and that words which begin with those sequences in a donor language are adapted when borrowed into English. This seems to be a robust fact about the lexical phonology of English, but it is also well recognised that

pre-tonic syncope of unstressed vowels can create initial sequences with [tl], in words like *Toledo*, *telegraphy*, and with [dl], in words like *dilemma*, *delightful* (and can also create other clusters that we would want to rule out from the lexicon, such as [pt] in *potato* — see, for example, Algeo 1975 and Szigetvári 2007). It seems right to say that OCP-ONSET(CORONAL) applies in the lexical phonology, but not in the post-lexical phonology, like the other phonotactics discussed in this section, but it *also* seems right to say that this kind of syncope is still under phonological control: it is not the case that absolutely anything goes. This type of syncope cannot create a falling sonority slope in an onset and so is not possible in words like *retire, reduction, mature, laconic* (Roca & Johnson 1999, 35). We can make sense of this if we assume that OCP-ONSET(CORONAL) applies in the lexical phonology and that pre-tonic syncope of unstressed vowels applies in the post-lexical phonology, which must mean that the constraint that is responsible for ruling out forms like *[lkɒnɪk] for *laconic* is part of the post-lexical phonology, too.

The full phonological model that I assume is therefore that the locus for the application of phonotactics is at the end of the relevant phonological stratum. The discussion here ties in with the Lexical/Stratal Phonology-type principled explanation for the ordering of phonological phenomena and opacity (as discussed in Bermúdez-Otero 2003, for instance), which allows for ordering if phenomena can be shown to belong to different strata but does not require or allow extrinsic ordering. Although Gimson (1962) does not put it in this way, he is discussing the fact that *RIME-xxŋ is opaque, and that the ordering of the rules and constraints involved is predicted from their properties, as long as we allow for (at least) two fundamental strata in phonology.

In fact, putting everything together, the model developed here predicts that *RIME-xxŋ will also prove to be opaque in Mid-Atlantic English, if we consider its interaction with *ash*-tensing and nasal place assimilation (NPA) of the type just discussed. Mid-Atlantic *ash*-tensing is lexical, *RIME-xxŋ is lexical and NPA is postlexical, therefore in pre-*g* contexts like that in (20), *ban* should be able to surface as [beəŋ], as modelled in (21), which shows *ash*-tensing at a time that it is still variable.

(20) *The mad president might ban gardening*.

The mad president might bang garden gates.

For comparison, (20) and (21) also show the derivations for *bang* in pre-*g* context, where NPA does not apply, and which is predicted *not* to be able to surface as [beəŋ]. (21) shows the two phonological strata that are assumed in the above, separated by a broken line (and

indicated by the vertical labels 'Lx Phonology' = 'lexical phonology' and 'PL Phonology' = 'postlexical/phrasal phonology'). The output of the lexical stratum is indicated in vertical slashes (to indicate that it is neither an underlying nor surface form), and NPA is assumed to be non-variable (which may well be wrong — if so, [beən] and [bæn] will also be possible surface forms for *ban* in the sentence in (20)).

(21)

	ban		*bang*		
	/bæn/	/bæn/	/bæŋ/	/bæŋ/	Lx Phonology
ash-tensing	beən	—	beəŋ	—	
*RIME-xxŋ	—	—	*	—	
output of LPh	|beən|	|bæn|	*	|bæŋ|	
NPA	beəŋ	bæŋ	*	—	PL Phonology
[other constraints]	—	—	*	—	
output of PLPh	[beəŋ]	[bæŋ]	*	[bæŋ]	

There is data in the literature on English phonology which poses a more fundamental challenge to the assumption that *RIME-xxŋ is part of the phonology of English. For example, Donegan & Stampe (1979, 149) describe a variety of English which has a process in which /æ, ɛ, ɪ/ diphthongise to [æe̯, ɛi̯, ɪi̯] before tautosyllabic [ʃ, ŋ, g]. This produces forms like [bæe̯ŋ] *bang*. Hayes (2009, 156) describes a similar situation, in which "/æ/ is diphthongised to [æ͡ɪ̯] before /ŋ/". I assume that the varieties described in these two sources are the same in this regard and that the two transcriptions for the output of the diphthongisation are notational variants. I transcribe the output as [æɪ], in line with the conventions that I am most used to — what is important for current purposes is that it is a bimoraic vocalic unit: a diphthong, which takes up two x-slots. Given that Donegan & Stampe link the process to other lax vowels as well as *ash*, I call it 'pre-*ŋ* diphthongisation' (PŋD). Neither Donegan & Stampe nor Hayes describe exactly which dialect they are referring to (I assume the transcriptions reflect the speech of the authors, and Hayes[27] says as much). It is clear that these are North American varieties which are not the same in this regard as the Mid-Atlantic dialects described in section 3.3.

[27] Hayes' biography on his website say that he grew up in Ithaca in New York state: https://linguistics.ucla.edu/people/hayes/Personal/Index.htm.

Donegan & Stampe's transcription for *bang* makes the point, and Hayes' data makes it clearer — this is a situation where a lexical rule produces forms which violates *RIME-xxŋ. The forms are presented as categorical representations, due to an invariable rule (Hayes 2009, 156 describes it as "an allophonic rule"), and the velar nasals involved are not derived by assimilation. Furthermore, both Donegan & Stampe and Hayes point out that the PŋD rule which diphthongises /æ/ is opaque because occurrences of /n/ which are assimilated to [ŋ] by NPA do not trigger it. Hayes (2009, 157) gives clear and compelling data for this, some of which is reproduced as (22) — amended slightly to fit with the transcription practice that I adopt for vowels (including flagging up that tense vowels are bimoraic, using the [ː] diacritic), and to remove stress marks, but otherwise following Hayes' transcriptional decisions (including his analysis of rhoticity). The first three rows of words compare phonologically similar words with either underlying /n/ (in the first column) or /ŋ/ (in the second column). The fourth row shows what happens when a word with underlying /n/ occurs in a context which triggers NPA: NPA can apply, but this is postlexical, and so occurs after the lexical rule of PŋD has had its chance to apply. This makes PŋD opaque, as in forms like *Dan Gurney*, and *pancake*, where non-diphthongised [æ] occurs in an environment which is pre-*ŋ* at the surface. Hayes contrasts this with the form *pang cake* which has /æ/ before an underlying /ŋ/, which therefore has a pre-*ŋ* environment in the lexical phonology, so that PŋD can apply (Donegan & Stampe cite *mankind* [mæŋkaɪnd] as a form that shows the opacity of PŋD for similar reasons).

(22)

pan	*pang*
/pæn/	/pæŋ/
[pæn]	[pæɪŋ]
fan	*fang*
/fæn/	/fæŋ/
[fæn]	[fæɪŋ]
gander	*anger*
/gændɚ/	/æŋgɚ/
[gændɚ]	[æɪŋgɚ]

Dan Gurney (the name of an American racing driver)
/dæn gɚːni/
[dæŋgɚːni]

pancake	*pang cake* 'cake eaten to assuage pangs of hunger'
/pænkeɪk/	/pæŋkeɪk/
[pæŋkeɪk]	[pæɪŋkeɪk]

This situation is not like the case in (18) from the dialect that Gimson (1962) is describing (which is like other British dialects that I am aware of), or like the Mid-Atlantic dialects discussed above. Dialects with PŋD must have lost *RIME-xxŋ in order for forms like [pæɪŋ] and [æɪŋgəˑ] to be possible. It is not surprising if different dialects have different phonotactics, but this recognition opens an interesting avenue for investigation and makes predictions about how other aspects of PŋD dialects will behave (for example, regarding loanwords).

We would expect that most other aspects of the phonology of PŋD dialects and non-PŋD dialects will be the same, however, and PŋD thus provides interesting evidence in the argument about whether /ŋ/ really exists in English (that is, about whether or not all occurrences of [ŋ] derive from sequences of /ng/ or /nk/, as discussed in section 3.1, following the analysis of the xxŋ gap in classical generative work, including Jensen 1993). It must be the case that words like *pang* have underlying /ŋ/ in order to trigger PŋD in the lexical phonology[28] because the relevant case of NPA is post-lexical — this means that at least some occurrences of surface [ŋ] must derive from underlying /ŋ/ (while others, as in *pancake*, derive from underlying /n/, which has assimilated to a following velar through NPA). If *some* case of [ŋ] where no [k] or [g] follows clearly derive from /ŋ/, there is every reason to assume that *all* cases of [ŋ] (where no [k] or [g] follows in any morphological form of a word) are underlying. As promised in section 3.1, there is therefore firm reason to believe that at least some occurrences of /ŋ/ must exist in underlying representations, and this, in turn, can be seen as evidence for the necessity of a *RIME-xxŋ constraint in English and against relying on a general constraint enforcing a rimal trimoraic maximum to rule out forms like [baɪŋ] and [biːŋ] (in non-PŋD dialects).

One final issue related to the interaction of *RIME-xxŋ and *ash*-tensing is that the phonotactic may, in fact (despite the above), currently be in the process of being lost in Mid-Atlantic dialects such as those described in section 3.3. Mid-Atlantic *ash*-tensing classically patterns as described in section 3.3, with its patterning showing a robust signature of *RIME-xxŋ, but other dialects (such as those in New England) have a simpler pattern of *ash*-tensing, typically described as the 'nasal pattern' (for example, in Labov 1994) or the 'nasal system' (as in Labov et al 2016). This is often described as tensing /æ/ in front of all (and only) syllable-final nasals (and so is obviously relevant to our concerns: this means in front of /m, n *and* ŋ/); however, Labov et al

[28] Unless we allow for the unappealing prospect of extrinsic ordering to make the derivation work in a brute-force manner, with *ŋ*-creation ordered before PŋD.

(2016, 277) write that "[i]n earlier characterizations of the nasal system, vowels before velar nasals were often not included. Friedman (2014) summarized the lax tendency of velar nasals in a range of Eastern and Midland dialects." This implies that the classic nasal system of *ash*-tensing is in fact constrained by *RIME-xxŋ. The point of Labov et al (2016) is that the nasal system is spreading and is now also found in Mid-Atlantic areas, competing with, and likely replacing the classic Mid-Atlantic 'split-a' systems. Labov et al continue that "recent studies of the expansion of the nasal system show tensing velar nasals playing a prominent role. Becker & Wong (2009, 15) reported that younger White New Yorkers showed no significant difference in the F1 of /æ/ before velar and other nasals. Eckert (personal communication) found consistent tensing of short-a before velar nasals among young Californians." This flags up the spread of the nasal system, and implies (although Labov et al do not put it like this) that *RIME-xxŋ is being lost in these varieties: if pre-*ŋ* tensing is included as part of the new (propagating) pattern, it seems that it is accompanied by a loss of *RIME-xxŋ. It is not clear if the patterning of the diphthongisation of /æ/ in PŋD dialects was an influence in its loss in this generalisation of the nasal system, but this may be worth considering.

The overall conclusion of this section is that *RIME-xxŋ is robust in the lexical phonology in most varieties of English, including the British varieties discussed, traditional Mid-Atlantic varieties, and the varieties with the traditional nasal-system of *ash*-tensing. There is good evidence, however, that it has been lost in some other American varieties and perhaps also that it is being lost in others. A potential implication of this change is that speakers of varieties like this would more readily accept forms like [baɪŋ], [biːŋ] and [baʊŋ] as 'wordlike' than speakers of other varieties.[29] Indeed, the form [baɪŋ], or something very similar, *is* found in these dialects, as flagged up in footnote 2.

4 Prophylaxis or therapy? Acquisitionism or lifespan change?

The key point of this paper is to extend and interrogate the position advocated in Honeybone (2019) — that phonotactic constraints are psychologically real phonological entities (namely: constraints on output-like forms), which have a diachrony of their own, and which can also interfere with diachronic segmental change by inhibiting otherwise regular innovations. I hope to have shown all of this in section 3 for the phonotactic *RIME-xxŋ: there is evidence that we can use to understand its innovation into English, its effect in inhibiting the patterning of *ash*-

[29] I thank Joe Fruehwald for discussion of these issues.

tensing, and its loss in certain varieties. The story is different in different dialects, as we would expect (and I not told the history of *RIME-xxŋ in *all* dialects of English). In this final section of this article, I turn to some implications of a major point of Honeybone (2019) and of this paper: if it really is the case that phonotactics can inhibit segmental change in the way that I discuss, then there are resonances with two central issues in historical phonology.

One of these issues is the fact that we might query whether we *should* expect that already-existing phonotactics are able to interact with changes which are otherwise independent of phonotactics, in light of the impetus that is often recognised in historical linguistics to reject prophylaxis in change. As mentioned above, there are resonances between the position on phonotactically-inhibited change adopted here and the notion of prophylaxis in diachrony: the two cases considered in Honeybone (2019), and the further case considered here involve a language *avoiding* enacting a change if to do so would have violated an already-existing phonotactic. This could be seen as languages avoiding change in order to prevent a 'problem' of some sort (such as a constraint violation).

There is much resistance to the notion that languages engage in prophylaxis in change. Fertig (2013), for example, explains that "[t]he slogan 'languages practice therapy, not prophylaxis' is widely attributed to Lightfoot (1979, 123) but as Harris & Campbell (1995, 28) point out, the wording is actually several years older and the idea goes back much further than that." Lass (1997), for another example, is sceptical of both prophylaxis and therapy as explanatory concepts in historical linguistics. The concepts involved in discussions of diachronic prophylaxis are often linked to the interaction of analogy and regular change, but they are applicable way beyond this. For example, Kiparsky (1974, 190) argues that "language practises 'therapy' (Gilliéron) rather than prophylaxis", as part of a rejection of the idea that phonological change can be inhibited if it would create homophony (which he attributes to Martinet). Kiparsky's reference is to Gilliéron (1915-1921), who explicitly argues that phonological change regularly creates homophones, and that this is followed by languages engaging in 'therapy', in order to remove the homophony. This is one example of discussion of the notion much earlier than Lightfoot, and shows that the idea is applicable in principle to any way of preventing a regular phonological change.

Fertig (2013) goes on to consider how neogrammarians such as Osthoff (1879) and Paul (1886) rejected the idea that phonological change can be prevented from occurring in the environments where it is expected. This is all relevant to my concerns here because what we

expect it that changes will occur in environments that can be defined by simple phonological principles: we expect phonological conditioning to involve natural classes of segments, not complex segmental conditioning (as is shown to be the case in section 3); if a syncope affects final unstressed vowels, this is a prosodically-defined environment, and we do not expect complex segmental conditioning, as would be necessary to prevent it from occurring between sibilants (as in the case of change discussed briefly in section 2 and in more detail in Honeybone (2019); furthermore, if a lenition occurs in 'strong' environments, we would also expect that it should also apply in 'weak' environments, yet the case of Mid-Scots θ-debuccalisation described in in Honeybone (2019) has exactly this patterning. These three changes have these unexpected conditionings — in a sense the regularity of the changes is disturbed — and the model set out in Honeybone (2019) and developed in this paper aims to explain how the phonotactics of the innovating language in some sense *prevented* the changes from applying in particular environments, causing this odd patterning.

Is this prophylaxis? Prophylaxis is indeed highly controversial in historical linguistics. We should be cautious of invoking a model which adopts it — of claiming that phonotactics can inhibit a phonological change from occurring in a subset of the environments where it is expected. A change is not aware of its effects, so cannot be expected to be inhibited if its effects might present some problem for a language.

The framework that I propose to model cases of phonotactically-inhibited change, however, avoids the conceptual problems that are connected with this. Central to understanding this is to recognise that the changes under consideration involve the introduction or generalisation of phonological rules. At the relevant level of the phonological derivation where the rules apply, they are in fact *not* inhibited and *do* affect the relevant phonological forms. In this sense, the changes themselves (the rule additions or generalisations) are not inhibited — there is no prophylaxis. Forms which violate the phonotactics *are* derived in the derivations in (9) and (17). The point is that the phonotactic constraints apply later in a derivation after phonotactic-violating forms have been derived, and these forms are ruled ungrammatical at the end of the relevant phonological stratum. The phonotactics prevent the forms that violate them from surfacing, but the rule which is innovated, or the generalisation of the rule which already exists does not know anything about this. The innovations in the rules or their environments in these cases are straightforward, and are as would be expected in line with the expectation that natural classes should be involved in rules, or that lenitions should pattern in terms of the 'strong implies weak' predictions that form lenition theory.

The phonotactics that are important in the three cases considered either here or in Honeybone (2019) — *RIME-xxŋ, OCP(SIBILANCE) and *CODA-h — do not prevent the rules from applying (this would be true prophylaxis), although they do stop ungrammatical forms from escaping from the grammar.

The other issue that is raised by this model — the final point that I discuss — is the fact that the model assumes that the phonotactics were already existing in the pre-change state of the languages involved when they affected the segmental changes. The phonotactics considered here are clearly language-specific: while some other languages may also use OCP(SIBILANCE), *CODA-h and *RIME-xxŋ, others do not, so the fact that they were active in pre-change late Middle English and Mid-Scots, and in Mid-Atlantic varieties before *ash*-tensing generalised, were facts about the mature phonology of these varieties. This means that these are cases where pre-existing language-specific structure in an innovating language affected the (surface) patterning of changes when they were innovated.

This is surprising for certain models of phonological change. If we consider the question *what is the locus of phonological change?*,[30] one of the main responses is 'in acquisition'. This is standardly assumed in much current generative historical linguistics — for example, van Kemenade (2007, 158) writes that "[e]ver since Lightfoot (1979), the generative approach to syntactic change has considered that the key mechanism of change is reanalysis, essentially the language learner's attribution of a novel underling analysis to the same surface form", and Yang (2000, 9) writes that "[l]anguage change is observed when a generation of speakers produces linguistic expressions that differ from those of previous generations". This position that first language acquisition is a key locus for change has long been recognised as a possibility in historical phonology (it is found in Paul 1886 and Kiparsky 1965), and the certainty that change *only* occurs in acquisition has moved from generative historical syntax into historical phonology — for example, Hale (2003, 345) writes that "we believe that 'change' is to be conceived of as the set of differences between the grammar generating the primary linguistic data (PLD) used by an acquirer and the grammar ultimately constructed by that acquirer" and Reiss (2003, 143) writes that "reranking or rule loss/addition cannot be defined within a single grammar, but is, at best, a description of the relationship between grammars", neatly expressing the idea that change only occurs 'across speakers', during acquisition. This position — the all change occurs in acquisition — can be called 'acquisitionism' (as in Honeybone

[30] Honeybone & Salmons (2015) suggest that this is one of the key questions for historical phonology (formulated as 'where does change occur?').

2006 and Honeybone & Salmons 2015). It is also the most obvious interpretation of Ohala's influential position (e.g., 1981, 1993) that the key locus of change is in the hearer, due to acoustic confusability or misperception: it is implausible to assume that an adult would suddenly make a reanalysis (that is, a mistake) on misperceiving something when they have a steady state with which they can compare what they hear — learners engaged in First Language Acquisition, however, lack a grammar against which to check a confusable utterance, and are thus those who are open to the type of reanalysis that is crucial on this picture.

Strict acquisitionism is thus strong in some areas of historical phonology. It is challenged, however, by those who believe in 'lifespan change', such as Sankoff & Blondeau (2007) and Bowie & Yaeger-Dror (2015), where at least certain types of change are argued to be possible post-acquisition. Such work can be described as 'anti-acquisitionist'. Aspects of anti-acquisitionism are, in fact, also present in the earliest generative historical work, such as Halle (1962, 64), who assumes that "changes in later life are restricted to the addition or elimination of a few rules in the grammar", thus accepting that post-acquisition change is possible — while also assuming (1962, 65) that "it is not necessary that the child and his parents have identical grammars", and that change can thus occur across generations, too. Halle continues (1962, 64), that "a wholesale restructuring of [...] grammar is beyond the capabilities of the average adult", making clear that he assumes that only certain types of change are possible post-acquisition. We can describe the position seen here as 'soft anti-acquistionism', that is, the assumption that change can occur both in acquisition and in the post-acquisition period. We can also recognise 'strong anti-acquisitionism', that is, the assumption that *no* change happens in first language acquisition, which Foulkes & Vihman (2015), for example, tentatively advocate.

The model that I propose for understanding phonotactically-inhibited change is most straightforwardly compatible with an anti-acquisitionist position: if all change occurs in first language acquisition — across generations — then language-specific structure in the innovating language should not be able to inhibit a change, as the innovating children do not yet have it. The proposals in section 2.3 and 3.3 assume that the phonotactics involved are already active in the language at the point at which the changes are innovated or generalised. For this to be possible, the changes must be added to speakers' phonology after the initial stage of language acquisition

(which fixed the phonotactic constraints).[31] It strikes me that soft anti-acquisitionism is likely right: there is some lifespan change, perhaps only involving the addition of a rule or similar 'small' changes, as Halle (1962) assumes, but more fundamental reanalyses (involving underlying representations, for example) occur in acquisition.[32]

5 Conclusion

This paper has considered quite a lot. I hope to have shown how complex and interesting questions in diachronic phonotactics can be, and also that we can and should pursue them. Certain changes are best viewed as phonotactic change — as the addition or loss or change of an already existing phonotactic — and other changes involve both phonotactic change and segmental or prosodic change occurring together; still other changes may involve phonotactic change being caused by other types of change. And we can clearly see that phonotactics can play a role in inhibiting the full surface patterning that comes in the wake of a segmental change, although this does not involve prophylaxis, narrowly defined. This latter point becomes clear if separate out the effect on the patterning of a change of the rule component of the grammar and on the constraint component of the grammar, and if we are clear to consider whether lexical or post-lexical phonology is involved. The model that I propose to understand phonotactically-inhibited change is tied to variability in phonology, and needs to involve careful consideration of the acquisition of phonological entities by learners — it can be seen as an argument for (at least soft) anti-acquistionism in historical phonology. The ideas discussed here are intended to be provocative, but I hope they are also intriguing.

---

[31] This conclusion follows only if the acquisition of the precise and detailed phonotactics of the type considered here proceeds at fundamentally the same pace as other aspects of phonological acquisition. This is an empirical question (and there is evidence that some aspects of phonotactic knowledge are acquired early, as a comment on a version of this paper points out). An alternative conclusion (if *detailed* phonotactic learning occurs before the learning of rule-like aspects of phonology) is that — if acquisitionism is correct — then *only* phonotactic knowledge should be able to inhibit innovations in the fashion discussed in this article. I am cautious about this: does early phonotactic learning really involve such detailed and specific phonotactics as those considered here, which require a knowledge of the natural classes of a language?

[32] This is also Dresher's (2015, 515) position, for example: "without claiming that all language change originates in acquisition, it appears unavoidable that certain types of changes do." It is notable also that Fruehwald (2017), after a detailed investigation of data for several changes from several decades, finds lifespan change in only a minority of the changes involved: it seems that lifespan change is possible, but is not part of all changes.

Comments invited

PiHPh relies on post-publication review of the papers that it publishes. If you have any comments on this piece, please add them to its comments site. You are encouraged to consult this site after reading the paper, as there may be comments from other readers there, and replies from the author. This paper's site is here:

http://dx.doi.org/10.2218/pihph.4.2019.4192

Acknowledgments

Early versions of some of the ideas in this paper were presented at the Diachronic Phonotactics Workshop at the University of Vienna in September 2017, and then reviewed for publication in *Folia Linguistica Historica,* leading to Honeybone (2019). A version of some of the other ideas was presented at the P-Workshop at the University of Edinburgh in 2019. I am grateful for all the comments and ideas that I got from this, especially from Niki Ritt, and I am particularly grateful to Ben Molineaux, who read through the paper as an editor for *PiHPh*, and whose comments led to many improvements in it. I alone am responsible for its contents, of course, however.

Author contact details

Patrick Honeybone
Linguistics and English Language
University of Edinburgh
Edinburgh
EH8 9AD, UK.

patrick.honeybone@ed.ac.uk

References

Abercrombie, David. 1967. *Elements of general phonetics*. Edinburgh: Edinburgh University Press.

Albright, Adam. In preparation. Natural classes are not enough: Biased generalization in novel onset clusters. Available here: http://web.mit.edu/albright/www/

Albright, Adam. 2006. Gradient phonotactic effects: lexical? grammatical? both? neither? Talk presented at the meeting of the Linguistic Society of America, Albuquerque. Available here:

http://www.mit.edu/~albright/papers/Albright-GradientPhonotactics-LSA2006.pdf

Algeo, John. 1975. Syncope and the phonotactics of English. *General Linguistics* 15. 71–78.

Algeo, John. 1978. What Consonant Clusters Are Possible? *Word* 29. 206–224

Bailey, George. 2018. *Variation and change in Northern English velar nasals: Production and perception*. PhD thesis, University of Manchester.

Baldwin, Thomas. 1846. *A Universal Pronouncing Gazetteer*. Third edition. Philadelphia: Lindsay & Blakiston.

Bauer, Laurie. 2015. English phonotactics. *English Language and Linguistics* 19. 437–475.

Baumann, Andreas, Nikolaus Ritt & Christina Prömer. 2016. Diachronic dynamics of Middle English phonotactics provide evidence for analogy effects among lexical and morphonotactic consonant clusters. *Papers in Historical Phonology* 1. 50–75.

Becker, Kara & Amy Wong. 2009. The short-a system of New York City English: an update. *University of Pennsylvania Working Papers in Linguistics* 1(2). 9–20.

Berent, Iris. 2013. *The phonological mind*. Cambridge: Cambridge University Press.

Bermúdez-Otero, Ricardo. 1998. Prosodic optimization: the Middle English length adjustment. *English Language and Linguistics* 2. 169–197.

Bermúdez-Otero, Ricardo. 2003. The acquisition of phonological opacity. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.), *Variation within Optimality Theory: Proceedings of the Stockholm Workshop on 'Variation within Optimality Theory'*, 25-36. Stockholm: Department of Linguistics, Stockholm University.

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume & Keren Rice (eds), *The Blackwell Companion to Phonology*, volume 4, 2019–2048. Malden, Mass.: Wiley-Blackwell.

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.

Bloomfield, Leonard. 1933. *Language*. New York: Holt.

Bosworth, Joseph & Thomas Toller. 1898-1921. *An Anglo-Saxon Dictionary*. With supplement. Oxford: Clarendon Press. Electronic version: http://bosworthtoller.com/

Bowie, David & Malcah Yaeger-Dror. 2015. Phonological change in real time. In Patrick Honeybone & Joseph Salmons (eds.) *The Oxford handbook of historical phonology*, 603-618. Oxford: Oxford University Press.

Carr, Philip, Jacques Durand & Colin Ewen (eds). 2005. *Headhood, Elements, Specification and Contrastivity*. Amsterdam: John Benjamins.

Chomsky, Noam & Morris Halle. 1965. Some controversial questions in phonological theory. *Journal of Linguistics* 1(2). 97–138.

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Dewan, Ted. 2003. *Bing: get dressed*. Oxford: David Fickling Books.

Dobson, E.J. 1968. *English pronunciation 1500-1700.* Second edition. Two volumes. Oxford: Clarendon.

Donegan, Patricia & David Stampe. 1979. The study of natural phonology. In Daniel Dinnsen (ed.), *Current Approaches to Phonological Theory*, 126–173. Bloomington: Indiana University Press

Dresher, B. Elan. 2015. Rule-based generative historical phonology. In Patrick Honeybone & Joseph Salmons (eds.) *The Oxford handbook of the historical phonology*. Oxford: Oxford University Press.

Dziubalska-Kołaczyk, Katarzyna. 2005. Phonotactics of consonant clusters in the history of English. In Bertacca, Antonio (ed.), *Historical linguistic studies of spoken English*, 15-34. Pisa: PLUS – Pisana University Press.

Ferguson, Charles. 1972. 'Short a' in Philadelphia English. In M. Estellie Smith (ed.), *Studies in linguistics in honor of George L. Trager*, 259-74. The Hague: Mouton.

Fertig, David. 2013. *Analogy and morphological change*. Edinburgh: Edinburgh University Press.

Flanders, Michael & Donald Swann. 1957. *A gnu*. London: Parlophone.

Foulkes, Paul & Marilyn Vihman. 2015. First language acquisition and phonological change. In Patrick Honeybone & Joseph Salmons (eds.) *The Oxford handbook of historical phonology*, 289–312. Oxford: Oxford University Press.

Fruehwald, Josef. 2017. Generations, lifespans, and the zeitgeist. *Language Variation and Change* 29(1). 1–27

Fudge, Erik. 1969. Syllables. *Journal of Linguistics* 5. 253–286.

Garrett, Andrew & Juliette Blevins. 2009. Analogical morphophonology. In Kristin Hanson & Sharon Inkelas (eds), *The nature of the word: essays in honor of Paul Kiparsky*, 527–45. Cambridge, Mass.: MIT Press.

Giegerich, Heinz. 1992. *English phonology: an introduction*. Cambridge: Cambridge University Press.
Gilliéron, Jules. 1915–1921. *Pathologie et thérapeutique verbales*. Four volumes. Various publishers, including Paris: Champion.
Gimson, A.C. 1962. *An introduction to the pronunciation of English*. London: Edward Arnold.
Goldsmith, John (ed.). 1993. *Phonological theory: the essential readings*. Malden & Oxford: Blackwell.
Gorman, Lyle. 2013. *Generative phonotactics*. PhD dissertation, University of Pennsylvania.
Hale, Mark. 2003. Neogrammarian sound change. In Brian Joseph & Richard Janda (eds.), *The handbook of historical linguistics*, 343–368. Malden: Blackwell.
Halle, Morris. 1959. *The sound pattern of Russian*. The Hague: Moulton.
Halle, Morris. 1962. Phonology in generative grammar. *Word* 18. 54–72.
Hammond, Michael. 1999. *The phonology of English: a prosodic Optimality-Theoretic approach*. Oxford: Oxford University Press.
Harms, Robert. 1968. *Introduction to phonological theory*. Englewood Cliffs: Prentice-Hall.
Harris, John. 1989. Towards a lexical analysis of sound change in progress. *Journal of Linguistics* 25. 35–56.
Harris, Alice & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
Hayes, Bruce. 2009. *Introductory phonology*. Malden & Oxford: Wiley-Blackwell.
Hayes, Bruce & Colin Wilson. 2008. A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry* 39. 379–440.
Hill, Archibald. 1958. *Introduction to linguistic structures: from sound to sentence in English*. New York: Harcourt, Brace.
Hockett, Charles. 1947. Problems of morphemic analysis. *Language* 23. 321.
Hockett, Charles. 1958. *A course in modern linguistics*. New York: Macmillan.
Hogg, Richard. 1992. *A grammar of Old English*, vol. 1: *Phonology*. Oxford: Blackwell.
Honeybone, Patrick. 2001. Lenition inhibition in Liverpool English. *English Language and Linguistics* 5. 213–249.
Honeybone, Patrick. 2005. Sharing makes us stronger: process inhibition and segmental structure. In Philip Carr, Jacques Durand & Colin Ewen (eds.), *Headhood, elements, specification and contrastivity: phonological papers in honour of John Anderson*, 167–192. Amsterdam: John Benjamins.

Honeybone, Patrick. 2006. Review of: D. Erc Holt (ed.) 2003. Optimality Theory and Language Change. *Journal of Linguistics* 42, 726-731.
Honeybone, Patrick. 2016. Are there impossible changes? θ > f but f ⊁ θ. *Papers in Historical Phonology* 1. 316–358.
Honeybone, Patrick. 2019. Can phonotactic constraints inhibit segmental change? Arguments from lenition and syncope. *Folia Linguistica Historica* 40. 9–36.
Honeybone, Patrick & Joseph Salmons. 2015. Introduction: key questions for historical phonology. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford handbook of historical phonology*, 3–10. Oxford: Oxford University Press.
Iverson, Gregory & Joseph Salmons. 2005. Filling the gap: English tense vowel plus final /š/. *Journal of English Linguistics* 33. 207–221.
Jensen, John. 1993. *English phonology*. Amsterdam: John Benjamins.
King, Robert D. 1969. *Historical linguistics and generative grammar*. Englewood Cliffs, NJ: Prentice-Hall.
Kiparsky, Paul. 1965. Phonological change. PhD thesis, MIT.
Kiparsky, Paul. 1968. Linguistic universals and linguistic change. In Bach, Emmon & Robert Harms (eds.), *Universals in linguistic theory*, 171–202. New York: Holt, Rinehart and Winston.
Kiparsky, Paul. 1971. Historical linguistics. In William O. Dingwall (ed.), *A survey of linguistic science*, 576-649. College Park: University of Maryland Linguistics Program.
Kiparsky, Paul. 1974. On the evaluation measure. In Anthony Bruck, Robert Fox & Michael LaGaly (eds.), *Proceedings of the parasession on natural phonology, tenth regional meeting, Chicago Linguistic Society*, 328-37. Chicago: Chicago Linguistic Society.
Kiparsky, Paul. 1982. Lexical morphology and phonology. In In-Seok Yang (ed.), *Linguistics in the morning calm: selected papers from SICOL–1981*, vol. 1, 3–91. Seoul: Hanshin.
Kiparsky, Paul. 1985. *Some consequences of Lexical Phonology*. Phonology Yearbook 2. 85–138.
Kiparsky, Paul. 1988. Phonological change. In Frederick Newmeyer (ed.), *Linguistics: the Cambridge survey*, volume 1, 363–415. Cambridge: Cambridge University Press.
Kiparsky, Paul. 2000. Opacity and cyclicity. In Nancy A. Ritter (ed.), *A review of Optimality Theory*. Special issue of *Linguistic Review* 17. 351–67.
Kroonen, Guus. 2013. *Etymological dictionary of Proto-Germanic*. Leiden: Brill. Available here: http://dictionaries.brillonline.com/proto-germanic
Kruisinga, Etsko. 1943. *The phonetic structure of English words*. Bern: A. Francke.

Labov, William. 1981. Resolving the Neogrammarian controversy. *Language* 57. 267–308.
Labov, William 1994. *Principles of linguistic change*, vol. 1: *internal factors*. Oxford: Blackwell.
Labov, William, Sabriya Fisher, Duna Gylfadottír, Anita Henderson & Betsy Sneller. 2016. Competing systems in Philadelphia phonology. *Language Variation and Change* 28. 273–305.
Lahiri, Aditi & B. Elan Dresher. 1999. Open syllable lengthening in West Germanic. *Language* 75. 678–719.
Lass, Roger. 1974. Linguistic orthogenesis? Scots vowel quantity and the English length conspiracy. In John Anderson & Charles Jones (eds.), *Historical linguistics: proceedings of the First International Conference on Historical Linguistics, Edinburgh, 2-7 September 1973*, volume 2, 311-52. Amsterdam: North-Holland.
Lass, Roger. 1992. Phonology and morphology. In Norman Blake (ed.), *The Cambridge history of the English language*, volume 2: *1066-1476*, 23–155. Cambridge: Cambridge University Press.
Lass, Roger. 1997. *Historical linguistics and language change*. Cambridge: Cambridge University Press.
Lass, Roger. 1999. Phonology and morphology. In Roger Lass (ed.), *Cambridge History of the English* Language, volume 3: *1476–1776*. Cambridge: Cambridge University Press.
Lentz, Tomas & René Kager. 2015. Categorical phonotactic knowledge filters second language input, but probabilistic phonotactic knowledge can still be acquired. *Language and Speech* 58. 387–413.
Lightfoot, David. 1979. *Principles of diachronic syntax*. Cambridge: Cambridge University Press.
Lindsey, Geoff & Péter Szigetvári. 2013. *Current British English Searchable Transcriptions*. http://cube.elte.hu/.
Luick, Karl. 1898. Beitrage zur englischen Grammatik, III: Die Quantitatsveranderungen im Laufe der englischen Sprachentwicklung. *Anglia* 20. 335–62.
Luick, Karl. 1914-1940. *Historische Grammatik der englischen Sprache*. Stuttgart: Bernhard Tauchnitz. Reprinted 1964. Oxford: Blackwell.
Lutz, Angelika. 1988. On the historical phonotactics of English. In Dieter Kastovsky & Gero Bauer (eds), *Luick revisited: papers read at the Luick-Symposium at Schloss Liechtenstein, 15-18.9.1985*, 221-239. Tübingen: Gunter Narr.
Lutz, Angelika. 1991. *Phonotaktisch gesteuerte Konsonanten-veränderungen in der Geschichte des Englischen*. Tübingen: Max Niemeyer Verlag.
MED = *Middle English Dictionary*. Original version 1952–2001. Updated online version, revised 2018. Michigan: University of Michigan

Library. Available here: https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary

Minkova, Donka. 1982. The environment for open syllable lengthening in Middle English. *Folia Linguistica Historica* 3(2). 29–58.

Minkova, Donka. 2003. *Alliteration and sound change in early English.* Cambridge: Cambridge University Press.

Minkova, Donka. 2014. *A historical phonology of English*. Edinburgh: University of Edinburgh Press.

Minkova, Donka & Robert Stockwell. 1996. The origins of long-short allomorphy in English. In Jacek Fisiak & Marcin Krygier (eds), *Advances in English historical linguistics 1996*, 211–39. Berlin: Mouton de Gruyter.

Ohala, John. 1981. The listener as a source of sound change. In Carrie Masek, Roberta Hendrick & Mary Frances Miller (eds.), *Papers from the parasession on language and behavior*, 178–203. Chicago: Chicago Linguistic Society.

Ohala, John. 1993. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: problems and perspectives*, 237–278. London: Longman.

Osthoff, Hermann. 1879. *Das physiologische und psychologische Moment in der sprachlichen Formenbildung*. Berlin: Habel.

OED = *Oxford English Dictionary*. Original version 1884–1928. Supplements 1933-1986. Online version consulted in December 2019. Oxford: Oxford University Press. Available here: http://www.oed.com/

Paul, Hermann. 1886. *Prinzipien der Sprachgeschichte*. Second edition. Halle: Max Niemeyer.

Reiss, Charles. 2003. Language change without constraint reranking. In D. Eric Holt (ed.), *Optimality Theory and Language Change*, 143–168. Dordrecht: Kluwer.

Ringe, Donald & Joseph Eska. 2013. *Historical linguistics: toward a twenty-first century reintegration*. Cambridge: Cambridge University Press.

Ritt, Nikolaus. 1994. *Quantity adjustment: vowel lengthening and shortening in Early Middle English.* Cambridge: Cambridge University Press.

Robinson, Orrin. 1976. A 'scattered' rule in Swiss German. *Language* 52. 148–62.

Roca, Iggy & Wyn Johnson. 1999. *A Workbook in Phonology*. Oxford: Blackwell.

Ross, Erica. 2011. *Diachronic developments in the scope of phonological generalisations*. PhD thesis, University of Cambridge.

Salmons, Joseph. In preparation. *Sound Change*. Edinburgh: Edinburgh University Press.
Sankoff, Gillian & Hélèn Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83. 560–88.
Scholes, Robert. 1966. *Phonotactic grammaticality*. The Hague: Mouton.
Schütze, Carson. To appear. Acceptability ratings cannot be taken at face value. In Samuel Schindler, Anna Drozdzowicz & Karen Brøcker (eds.), *Linguistic intuitions*. Oxford: Oxford University Press. Available at: https://ling.auf.net/lingbuzz/004862
Shatzman, Keren & René Kager. 2007. A role for phonotactic constraints in speech perception. *ICPhC XVI*, *Saarbrücken*, 1409–1412.
Singh, Rajendra. 1987. Well-formedness conditions and phonological theory. In Wolfgang Dressler et al. (ed.), *Phonologica 1984*, 273-285. Cambridge, Cambridge University Press.
Sneller, Betsy. 2018. Mechanisms of Phonological Change. PhD dissertation, University of Pennsylvania.
Sneller, Betsy. 2019. Allophonic systems as a variable within individual speakers. In David Lightoot (ed.), *Variable properties in language: their nature and acquisition*. Washington, DC: Georgetown University Press.
Sommerstein, Alan. 1974. On phonotactically motivated rules. *Journal of Linguistics* 10. 71–94.
Spence, Thomas. 1775. *The grand repository of the English Language*. Newcastle Upon Tyne: printed by T. Saint, for the author, and sold by him at his school on the Keyside, and by all the booksellers in town and country.
Sprouse, Jon. in press. Acceptability judgment methods. In: Jon Sprouse (ed.), *The Oxford Handbook of Experimental Syntax*. Oxford University Press.
Stanley, Richard. 1967. Redundancy rules in phonology. *Language* 43. 393–436.
Steriade, Donca. 1999. Alternatives to the syllabic interpretation of consonantal phonotactics. In Osamu Fujimura, Brian Joseph & Bohumil Palek (eds.), *Proceedings of the 1998 Linguistics and Phonetics Conference*, 205-242. Prague: Karolinum Press.
Szigetvári, Péter. 2007. Branching onsets and syncope in English. *Language Sciences* 29. 408–25.
Taylor, John. 2002. *Cognitive grammar*. Oxford: Oxford University Press
Uffmann, Christian. 2015. Loanword adaptation. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford Handbook of Historical Phonology*, 644-665. Oxford: Oxford University Press.
van Kemenade, Ans. 2007. Formal syntax and language change: developments and outlook. *Diachronica* 24. 155–169

Walker, John. 1791. *A critical pronouncing dictionary and expositor of the English Language*. London: sold by G. G. J. and J. Robinson, Paternoster Row; and T. Cadell, in the Strand.

Wells, John C. 1982. *Accents of English*. 3 vols. Cambridge: Cambridge University Press.

William Labov, Sabriya Fisher, Duna Gylfadottir, Anita Henderson & Betsy Sneller. 2016. Competing systems in Philadelphia phonology. *Language Variation and Change* 28(3). 273–305.

Yang, Charles. 2005. On productivity. *Linguistic variation yearbook* 5(1). 265–302.

Yang, Charles. 2016. *The price of productivity*. Cambridge, Mass.: MIT press.