CrossMark

English Language and Linguistics, 25.1: 155–180. © The Author(s) 2020. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike licence (http://creativecommons.org/licenses/by-nc-sa/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the same Creative Commons licence is included and the original work is properly cited. The written permission of Cambridge University Press must be obtained for commercial re-use.
doi:10.1017/S1360674319000522

Middle English Open Syllable Lengthening (MEOSL) or Middle English Compensatory Lengthening (MECL)?[1]

DONKA MINKOVA and MICHAEL LEFKOWITZ
University of California, Los Angeles
(Received 8 July 2019; revised 28 October 2019)

This study addresses a controversial aspect of the change traditionally known as Middle English Open Syllable Lengthening (MEOSL): the variable results of lengthening in disyllabic (C)V.CVC stems, the *heaven–haven* conundrum. It presents a full philological survey of the recoverable monomorphemic input items and their reflexes in Present-day English (PDE). A re-examination of the empirical data reveals a previously unnoticed correlation between lengthening and the sonority of the medial consonant in forms such as *pa**p**er*, *ro**ck**et*, *ga**nn**et* and *ba**r**on*, as well as interplay between that consonant and the σ_2 coda. The alignment of disyllabic stems with a medial alveolar stop and a sonorant weak syllable coda (*Latin*, *better*, *otter*) with (C)V.RVR stems (*baron*, *felon*, *moral*) opens up a new perspective on the reconstruction of tapping in English. The results of lengthening in disyllabic forms, including those previously thought of as 'exceptions' to the change, are modeled in Classical OT and Maxent OT, prompting an account which reframes MEOSL as a stem-level compensatory process (MECL) for all inputs. We show that OT grammars with conventional constraints can correctly predict variation in the (C)V.TəR stems and categorical lengthening or non-lengthening in other disyllabic stems. Broadening the phonological factors beyond the open-syllable condition for potential stressed σ_1 inputs in (C)V.CV(C) stems allows us to apply the same constraints to stems whose input structure does not involve an open syllable and to propose a uniform account of stressed vowel quantity in all late Middle English mono- and di-syllabic stems.

Keywords: Middle English, phonology, vowel lengthening, apocope, syncope

[1] This article has benefitted from many earlier presentations and discussions. The data in sections 2 and 3 were first presented by Donka Minkova at the International Conference on Middle English in Murcia, Spain, in May 2013. A revised version was included in a joint presentation by Minkova and Nikolaus Ritt at the Second Edinburgh Symposium on Historical Phonology in December 2015 as part of a larger inquiry addressing both stem-level phonological processes and higher-level prosodic factors in quantitative changes. The authors of the current article thank Niki Ritt for many insightful exchanges and truly valuable contributions to our project. We are also grateful to the *ELL* editors and reviewers for helpful and challenging comments.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

1 Introduction

The research presented here addresses a familiar quantity change in English, commonly referred to as Middle English Open Syllable Lengthening (MEOSL). Earlier publications on the subject, including Minkova (1982, 1985, 2014), Ritt (1988, 1992, 1994, 1997, 2000), Bermúdez-Otero (1998), Lahiri & Dresher (1999), identified some problematic aspects of the existing accounts. In particular, the outstanding question is whether the set of diachronic correspondences covered by the change can be adequately explained – as its established name suggests – in terms of a single conditioning factor. This contribution extends the empirical base for the discussion, reveals new details in the array of previously recognized phonological factors, and provides a phonological account which shrinks the randomness in the relation between the inputs to the change and its Present-day English (PDE) descendants. The limited remaining cases where application of the change is unpredictable are modeled using Maxent OT, a stochastic grammar formalism, deriving a probabilistic grammar which successfully predicts that variation should occur in exactly these cases.

We start with an introduction to MEOSL, a brief survey of the basic regularities it describes, and highlight the problems it raises both empirically and theoretically. Focusing on (C)V.CV(C) stem-forms, we present the full input–output database in terms of the weight and the sonority of both the onset and the coda of the post-tonic syllable (σ_2). The syllable weight condition is categorical at the endpoints of a continuum, gradient in-between. The variable outcomes in (C)V.CV(C) stem-forms have been described as depending 'on chance' (Liberman 2015: 172), yet the examination of σ_2 onsets, which have been ignored in earlier accounts, reveals a new angle on this familiar change: the inhibiting effect of intervocalic sonorants on lengthening. On the other hand, the outcome in items with medial obstruents varies, leading to the hypothesis that the discrepancy is motivated by the phonotactic relations between the σ_2 onset and coda. We formalize the empirical findings using the OT formalism, and show that the ME quantitative changes in the entire set of disyllabic stem-form inputs can be accounted for in terms of compensatory lengthening, conditioned by σ_2 apocope and syncope or lack thereof. Our findings reaffirm the link between σ_2 apocope and syncope in the context of multiple factors which generate variability in the phonological representations of vowels, and thereby lead to an account of the variable yet systematic way in which vowel lengthening has come to be reflected in the English lexicon.

1.1 Some preliminaries: revisiting MEOSL

MEOSL is usually understood as a change affecting stressed, short, non-high vowels in open penultimate syllables of disyllabic words: #(C)V́.Cə# > (C)V́V.C(ə) > #(C)V́VC#. (1) is a reminder of what OSL stands for.

(1) Open-Syllable Lengthening (*c.* 1200–1400):
Early ME Late ME

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

(a)	(C)V.Cə~(C)V.CØ	(C)VV.CØ	
	[n**ɑ.**m(ə)]	[n**a**:m]	'name'
	[m**ɛ.**t(ə)]	[m**ɛ**:t]	'meat'
	[n**ɔ.**z(ə)]	[n**ɔ**:z]	'nose'
	… but:		
(b)	(C)V.CVC	(C)V.CəC~(C)VV.CəC	
	[s**ɛ.**vən]	[s**ɛ.**vən]	'seven'
	[b**ɛ.**vər]	[b**ɛ**:**.**vər]	'beaver'

Evidence of vowel quantity in ME is often ambiguous, and it allows for variability. This means that the OSL initiation point is as difficult to reconstruct as is the way in which the change evolved from what must have been an optional and phonetically conditioned process at first into the phonologically reinterpreted representations identifiable in the PDE lexicon.[2] Thus, crucially, the most informative long-term evidence of the change comes from its post-medieval reflexes in the form recorded in PDE 'standard' pronouncing dictionaries. We can be quite sure that *name*, *meat*, *nose* and *beaver* must have had phonologically short vowels in Old English (OE), judging from the way such items are treated in the verse. We are also sure that the majority of their realizations had long vowels when the Long Vowel Shift started, because otherwise they would not have been affected by it.

The time-line for the beginning of the change is fuzzy, especially in view of the significant absence of original compositions and autograph texts between *c.* 1066 and the second half of the twelfth century.[3] Doubling of vowels to indicate length in early ME is attested only sporadically for some items, but not for any of the OSL candidates pre-1325 (see *A Corpus of Narrative Etymologies from Proto-Old English to Early Middle English* (CoNE) on Orthographic Doubling of Long Vowels (ODLV)). The first rhyme evidence for OSL appears around the mid-thirteenth century in the northern dialects (Luick 1964: §391; Jordan & Crook 1974: 47–9; CoNE s.v. *MEOSL*). The gradual southward spread of the change is suggested by rhymes found in identifiably southern texts before the middle of the fourteenth century, e.g. *before : sore* (OE *beforan* 'before', OE *sār* 'sore').[4] The rhyme evidence for a full merger of the newly lengthened vowels with the pre-existing long vowels is unambiguous only after the middle of the fifteenth century (Luick 1964: §391; Ikegami 1984: 315–20; Ogura 1987: 126–8). Although the evidence points to an earlier start of OSL in the north, the details of the regional spread of the change in late ME/early Modern English are not

[2] PDE in this article refers to General British English (GB) and General American English (AmE). We do not address alternative forms in other regional varieties, e.g. Scots [hi(:)vən, ˈhɪvən, he:vən], n. or [hɛzl, he:zl].

[3] Hogg (1996) suggests the possibility of allophonic lengthening of the low vowel [ɑ] already in OE. Orthographic/diacritic indications of length marking for the relevant vowels cannot be precluded for OE (Ritt 2000). Mokrowiecki's (2015) findings of graphic marking of vowel quantity in OE are far from definitive, and he refers to the interpretation of the evidence with respect to OSL as the subject of further research.

[4] From *The King of Tars,* the *Auchinleck* manuscript (*c.*1330–40) version of the poem; see Minkova (2014: 222). The rhyme reflects the raising of OE [ɒ:] to [ɔ:] in the Southern dialects.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

clear-cut and continue to show variability. Dobson (1957: 467, 474) remarks specifically on the variable quantity of the stressed syllable in sonorant-final disyllables: *father*, *rather*, *water*, *haven*, *brazen*; *leaven*, *lever*, *heron*, *leper*.[5] Variable vowel quantities for stems preserving their disyllabic structure diachronically, the (C)V.CəC∼(C)VV.CəC subset, are projected in PDE variants such as GB [ˈlɛʒə]∼AmE [ˈliʒər] 'leisure', AmE /ˈsædər/∼/ˈseɪdər/ 'satyr'. The preferred association between stress and weight in open syllables is demonstrated in the realization of new lexical items: *preta* (1811), *NATO* (1949), *modem* (1958), comp. *model* (1570); *laser* (1960), *quasar* (1964), comp. *tassel* (1330), *dossal* (1658), precluding positing a *terminus ad quem* for the change.

Saying that evidence of MEOSL is scarce and often indirect does not mean there is none of it. On the contrary, one of the reasons why we think that the change is worth revisiting is precisely that more empirical/philological information is accessible now than twenty years ago. A new look at the data and the application of new theoretical models can refine our picture and deepen our understanding of the change considerably. Thus, some of the questions about OSL that had to be left open in earlier research have become addressable today, and that's the goal of the present study.

1.2 A note on the immunity of high vowels to OSL

Discussion of stressed vowels marked V in this article excludes high vowels. The OE short high vowels were subject to sporadic lowering and lengthening to [eː] and [oː] in the Northern dialects, but only three /ɪ/-lengthened forms survive in PDE: OE *wicu* > PDE *week*, OE *wifel* 'weevil', possibly OE *yfel* 'evil' ∼ Old Kentish **efel* (*MED*); the commonly cited item *beetle* is attested only as a trisyllabic form in OE, *bitela* (*DOE*) and in ME (*MED*). The only two examples of sustained OSL of /ʊ/ are OE *wudu* 'wood' and (possibly) OE *duru* 'door', but the vowel in both words can be explained without reference to OSL. The default development is preservation of the high vowels as monomoraic: OE *hype* 'hip', OE *dile* 'dill', OE *hulu* 'hull', OE *lufu* 'love', OE *munuc* 'monk'. We are ignoring possible variable length in ME; lengthened ME forms of e.g. OE *gyfan* 'to give', *sunu* 'son' are not quantifiable.

There is a reasonable explanation for why high vowels may have failed to reliably lengthen in lengthening environments. High vowels are intrinsically shorter than mid vowels in PDE (Peterson & Lehiste 1960)[6] and in many if not all of the world's languages (Lehiste 1970), so they were likely shorter in ME as well. If a historical process were to lengthen the short high vowels – the shortest stressed vowels in the inventory – this would have the effect of making them less perceptually distinct from mid vowels, and therefore be antithetical to the maintenance of a dispersed vowel space.

5 On the ambiguity of stressed open (CV.) syllables in Sidney's experiments in quantitative versification see Hanson (2001: 64–7). For Sidney (1554–86) the stressed syllables of e.g. *happy*, *blemish*, *echo* scan as heavy, while *many*, *merit*, *never* have light stressed syllables.

6 Westbury & Keating (1980) provide articulatory evidence that these differences in intrinsic duration are intentional, and not merely a biomechanical by-product of differences in jaw height.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

2 Focus on the post-tonic syllable

While a disyllabic input (C)V.CV(C) is the starting point, the macro-perspective shows that the probability of lengthening depends on multiple competing factors. In the course of ME, schwa apocope in (C)VCə stems was ubiquitous; a */-ə/# constraint was practically inviolable by *c.* 1400–50 (Minkova 2015); see section 4.1 for an Optimality Theoretic (OT) treatment of this constraint. Unstressed vowel syncope in (C)VCəC stems, on the other hand, was optional and generally preserved the disyllabic structure of the inputs.[7] A previously established predictor of lengthened and unlengthened stressed short vowels in the stem-forms is σ_2 stability (Minkova 1982, 1985; Ritt 1994: 30ff.): diachronic loss of σ_2 as in (2a) corresponds to a near-categorical lengthening.

(2) Lengthening input–output and σ_2 loss (low and mid vowels):

(a) OE: (C)VCə	(b) OE: (C)VCəC[8]
nama 'name' vs	*wæter* 'water', *baron*, *ganot* 'gannet'
mete 'meat' vs	*seofon* 'seven', *feþer* 'feather', *fetor* 'fetter'
nosu 'nose' vs	*coper* 'copper', *ofen* 'oven', *rocet* 'rocket, cloak'

The relevance of the stability of the second syllable is statistically testable: the ratio of lengthened to unlengthened forms in the bottom (C)V.Cə bar in figure 1 is reversed in the first and second bars.

For the (C)V.Cə inputs the 4.9 percent of unlengthened forms is based, very conservatively, on Minkova (1982: n. 5), but if we exclude CV.stə items: *cast*, *fast*, *hasp* (OE *hæpse*), *mast*, *rest*, which have northern and Scots lengthened forms, the unlengthened forms ratio drops to 3.3 percent. Further, there are ME lengthened forms for the remaining exceptions: *crack* (ME *crak(e)*, rhyming with *blake*, *spake*, *lake*), *beck* (Sc. *baik*), *drop* (ME *droupe*), *fret* (ME *freaten*), *get* (ME *geyt*, *gait*), *lap* (ME *lape*, *Sc laip*), *rot* (ME *roote*, *royte*), *tread* (ME *treide*). Excluding *knock* (OE *cnucian*) and *wag* (ME *waw*) makes the (C)V.Cə > (C)VVC change practically exceptionless, which is our justification for referring to it as 'categorical' in the rest of this article.

The inputs surviving as PDE disyllabic forms, the top two bars in figure 1, are of particular interest in all OSL accounts. The different results for monosyllabic and disyllabic outputs have been a central concern in the long OSL research history. The link between variability and the presence of coda sonorants in the unstressed syllable has always been part of the narrative (Luick 1964: §191, 192; Dobson 1962: 127–8; Lass 1992: 73; Bermúdez-Otero 1998). Bermúdez-Otero is very specific about 'the positive correlation between lengthening and the presence of a sonorant consonant in the post-tonic rhyme' (1998: 175). Separating .CəT# from .CəR# second syllables is

[7] A handful of syncopated CV.C(ə)C items survived as monosyllables: *heorot* ~ *heort* 'hart', *merisc* ~ *mersc* 'marsh', *waroþ* ~ *warþ* 'warth' are OE doublets, and *henep* ~ *hemp* 'hemp', *havek* ~ *hauk* 'hawk' coexist in ME. All such input forms fit the syncope patterns discussed in relation to the probability of lengthening in this article.

[8] We identify the second syllable peak as a generic /ə/, to the exclusion of [ɪ] + the palatal semivowel [j] (OE *-iġ,* e.g. *manig* 'many'), [ɪ] + [ŋ(g)], and [ɔ/ʊ] + [w] in inflected *-wa* stems.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Figure 1. OSL and post-tonic syllable preservation as percentage[9]

justified by the imbalance in the outcomes: a breakdown of the (C)V.CəC data shows that in 39 items the σ_2 coda is an obstruent: these are words of the type *gannet*, *planet*, *relic, rocket*, both native words and pre-1400 loanwords. The details on (C)V.CəT stems are presented in (3):

(3) (C)V.CəT items eligible for OSL (Minkova 1982; Bermúdez-Otero 1998: 193)[10]
 (a) Lengthened: 1 item: *naked*
 (b) Unlengthened: 39 items: *chalice*, *collop*, *eddish*, *gannet*, *haddock*, *planet*, *provost*, *radish*, *relic*, *trivet*[11] \\ *anet*, *anise*, †*barrat*,[12] *claret*, *damask*, *faggot*,[13] *habit*, *latchet*,[14] *marish*, *palace*, *palate*, *statute*, *brevet*, *ethic, jealous*, *legate*, *prelate*, *senate*, *trellice*, *bonnet*, *closet*, *crocket*, *forest*, *profit*, *rocket*, *rochet*,[15] *socket*, *solace*; *cherish*.

With the exception of *naked*, possibly influenced by the ME verb *nāke* (*c.*1350) and adjective *nāk(e)* 'naked' (*c.*1300),[16] all other items preserve the original short vowel. The inhibition of lengthening in stems with -.əT# σ_2 rhymes seems to require an

[9] T stands for any obstruent and R stands for a sonorant. The chart is based on the raw data in Minkova (1982), Bermúdez-Otero (1998: 175).

[10] The back slashes here and in (4) separate items attested in OE from items recorded in pre-1400 Middle English, mostly Old French/Anglo-Norman. Not included in the list in (3): *basket*, *jaspis*, whose intersyllabic clusters /sk, sp/ are cohesive, frequent word-initial clusters, and could therefore potentially be syllabified as second-syllable onsets. The only testably lengthened form of that type is OE *wesle* 'weasel'; note the sonorant coda in σ_2. Out of 39 **un**lengthened items only *damask*, *forest*, *provost* (Brit. /ˈprɒvəst/), have a -VCC# in σ_2. The distribution of the medial consonants in this subset is: 21 sonorants (53.8%); 10 voiced obstruents (25.6%); 6 voiceless stops and fricatives 4x/k, t/, 2x/f, θ/ (15.4%); 2 voiceless affricates (5%). Just over half of the items have an intervocalic sonorant. Medial OE (pre-) affricates block OSL systematically because they were still bi-segmental (Minkova 2016), though items from Old French (*latchet*, *rochet*) allowed medial singletons; for details see Minkova (2019: 174–5).

[11] OE *trefet*, Old Northern French *trevet.*

[12] Old French *barat,* ME *baret* 'fraud, grief, pain'.

[13] French *fagot* 'a bundle of sticks'.

[14] Old French *lachet*, dialect variant of *lacet*, 'lace, a thong'.

[15] Old French *rochet* 'ecclesiastical vestment'.

[16] Relevant CV.CəR and CV.Cə variants favoring lengthening are: ME *nāken* v. also *nake* (*MED*), e.g. 'O nyce men! why *nake* ye your bakkes?' ?a.1425(*c.*1380) Chaucer Bo. 4.m.7.72; ME *nāke* adj. also *naken* (*MED*), e.g. 'Kyng Thoas herte be-gan to *qwake*, He wende to be hanged al *nake*' *c.*1425(*c.*1400) *The Laud Troy Book* (LdMisc 595) 74.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

explanation independent of the framework in which the analysis is conducted. Ritt (1994, 1997) relates this to the weight of the second syllable. Lahiri & Dresher (1999: 694–5) acknowledge the imbalance between σ_2 -.əT# and -.əR# rhymes, but do not consider the presence of a stem-final obstruent a determining factor for the discrepancy; they refer to possible tri-syllabic forms and medial geminates as blocking lengthening, but these considerations also apply to forms with σ_2 sonorant codas, where the results are significantly different.

Since the results in the (C)VCəT# database in (3) are practically uniform, we leave that set aside for the moment, noting that the vowel in unstressed –.CəT#, the *rocket*-type, is stable compared to –əR#, where schwa alternates with the syllabic sonorant as in the set surveyed in the next section.

2.1 Target group: (C)V.CəR

Figure 2 compares lengthened and unlengthened reflexes of (C)V.CəR inputs.

The details on the items under discussion are presented in (4). Forms whose stressed vowel varies in length are boldfaced; they are recorded twice: both as 'unlengthened' and as 'lengthened'. The slashes separate inherited OE words from early ME loans.

(4) (C)V.CəR items (Minkova 1982; Bermúdez-Otero 1998):

(a) Lengthened: 38 items

acorn, *acre*, *beaver*, ***besom***, *chafer*, ***cradle/creddle***, *even*, *gable*, *haven*, *hazel*, *ladle*, *maple*, *navel*, *open*, ***over*** (cf. ***uvver***), *raven*, *staple*, *taper*, *weasel* \\ *bacon*, *basin*, *blazon*, *capon*, *favour*, *flavour*, *label*, *labour*, *mason*, *paper*, *patient*, *savour*, *razor*, *tabor*, *vacant*, *vapour*, ***azure***, *moment*, *odour.*

(b) Unlengthened: 128 items

saddle, *aspen*, *bastard*, *batten*, ***besom***, *better*, *blather*, *bottom*, *bracken*, *brothel*, *cackle*, *camel*, *canon*, *castle*, *chaffer*, *clatter*, *cocker*, *cockle*, *copper*, ***creddle/cradle***, *edder*, *father*, *feather*, *fennel*, *fester*, *fetter*, *fettle*, *gammon*, *gather*, *gavel*, *gravel*, *hammer*, *hatchel*, *heaven*, *hovel*, *hover*, *kettle*, *Latin*, *latter*, *leather*, *maslin (obs.)*, *nether*, *nettle*, *otter*, *oven*, ***uvver*** (dial.; cf. *over*), *pebble*, *pepper*, *pottle*, *rather*, *reckon*, *repple (obs.)*, *saddle*, *seven*, *shackle*, *shovel*, *smother*, *sollar*, *swaddle*, *talent*, *tetter*, *throttle*, *throstle*, *treadle*, *water*, *wattle*, *weather*, *wether*, *whether* \\ *alum*, ***azure*** (also long), *baron*, *barren*, *barrel*, *cattle*, *channel*, *chattel*, *clamour*, *dragon*, *flatter*, *gallon*, *hazard*, *latten*,

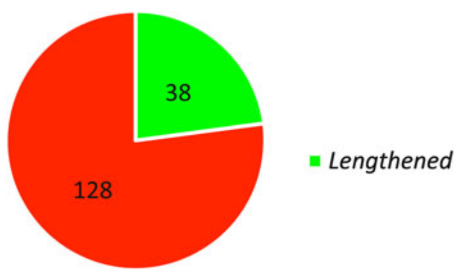

Figure 2. Reflexes of (C)V.CəR inputs

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

mallard, *manor*, *panel*, *satchel*, *satin*, *tabard*, *talon*, *tassel*, *travel*, *valour*, *vassal*, *warrant*, *beryl*, *bezant*, *cellar*, *deavour*, *desert*, *felon*, *herald*, *kennel*, *lecher*, *lesson*, *metal*, *pennon*, *peril*, *present*, *record*, *revel*, *second*, *tenant*, *tenor*, *treasure*, *tremor*, *venom*, *coffin*, *collar*, *common*, *coral*, *florin*, *foreign*, *honour*, *moral*, *volume; colour*, *covin.*

The variability of the results for (C)V.CəR inputs has been noted and commented on before (see Bermúdez-Otero 1998; Lahiri & Dresher 1999; Page 2006). We cite some representative statements in (5):

(5) Why are (C)V.CəR inputs interesting?

Lengthening applies *unpredictably* among unapocopated disyllables containing a sonorant consonant in the post-tonic rhyme. (Bermúdez-Otero 1998: 176)

Words ending in final sonorants can go in both directions. *It is not clear* why this should be the case, but it should be noted, that there are relatively few disyllabic words with final obstruents of Germanic origin in Old English to begin with.
(Lahiri & Dresher 1999: 695)[17]

The lengthenings [in items whose post-tonic rhyme contains a sonorant] *appear to be more the exception than the rule*. (Page 2006: 69)

If the application of OSL depends on the σ_2 coda, as one might be tempted to infer on the basis of the contrast of the two lower bars in figure 1, the substantial proportion of lengthened (C)V.CəR words, 22.9 percent, warrants further inquiry.[18]

The instability of [-ə-] in σ_2 in -.CəR can be traced back to the earliest stages of English.[19] In OE verse -.CəR forms can scan as monosyllabic; similarly, in ME and Renaissance verse the syllabic value of .-TəR continues to be variable.[20] Walker (1797: §102, §159, §170) also focuses on variants triggered by a σ2 coda sonorant

[17] The majority of the input items in (3) and (4) are nouns. Since the borrowed nouns would not have syllabic plurals, or other syllabic inflected forms, the paucity of Germanic inputs in the unlengthened forms is a challenge to Lahiri & Dresher's (1999: 689ff.) framework which accounts for these unlengthened forms in terms of trisyllabic shortening.

[18] The reference to σ_2 rhymes in the statements in (5) obscures the difference between σ_2 with just a sonorant coda and, potentially, a sonorant onset to a syllabic sonorant, .RR̩ σ2. Since that distinction is not made in the previous accounts, we take 'rhyme' in (5) to mean 'coda'.

[19] Syncope occurs typically in inflected disyllabic forms: *fæder* (sg.) – *fæd(e)ras* (pl.), 'father', *mōdor* (sg) - *mōdra* (gen. pl) 'mother', though variation is attested also in stem forms as in *ofn* ~ *ofen* 'oven', *weder* ~ *wedr* 'weather' (Fulk 1992: ch.1; Hogg & Fulk 2011: § 3.68–72). More variants of originally monosyllabic base forms undergoing epenthesis/parasiting (nom.sg.) in the *DOE* such as OE *æppel* ~ *æpl* 'apple', *fædm* ~ *fæðem* 'bosom, embrace', *hræfn* ~ *hræfen* 'raven' confirm the variability of the forms; see also *Sonorant Cluster Vowel Epenthesis* (SCVE) in CoNE.

[20] Syncope and resyllabification in unstressed .CəR# syllables followed by 'weak' <h-> or a vowel occur commonly in Chaucer:

(So fressh, so ȝong, so weldy semed he,)
It was an **heuen** vp-on hym forto see. (*Troilus and Criseyde* II, 636–7)

For examples of the syllabic variability of CVTəR strings in Renaissance verse, see Tarlinskaja (2014: 15–17, 93–4, 112–16) for monosyllabic scansion of *heaven*, *dazzle*, *lemon*.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

and recommends syncope in *weas'l*, *haz'l*, *mas'n*, *seas'n*, *blaz'n*, *dev'l*, but objects to it in *woolen*, *flannel*, *cavil*, *Latin*, allowing both forms in *novel*, *raven*. Bermúdez-Otero's compensatory analysis (1998: 175–8) attributed the potential lengthening in (C)V.CVR words, e.g. OE *beofor* > PDE *beaver*, to optional/variable schwa deletion in the second syllable, whereby unrealized input schwas are mora-donors. This hypocorrective compensation option has the advantage of aligning the *beaver*, *paper*-type forms with the *nama*-type forms in terms of moraic preservation. As we will show in section 4, this is on the right track, especially if one factors in the significant effect of the medial sonorants in (C)V.RəC inputs.[21] The compensatory factor contributing to the lengthening is therefore central to the process; a more detailed treatment of (C) V.CVR forms allows interesting predictions about the probability of lengthened vs unlengthened forms within that notoriously variable subset.

3 Medial consonants in (C)V.CVR inputs to OSL

No study of MEOSL other than Ritt (1994) considers the possibility that variable lengthening in (C)VCVC inputs may be sensitive to the nature of the medial consonant.[22] Ritt's attention to this matter is unique, but he bundles together uniformly unlengthened forms: OE *-ig* forms and the *gannet*-type forms shown in (2), with the (C)VCVR items which behave variably (1994: 63ff.). His basic proposal zeroes in on the structural association of the medial consonant to the stressed vowel.

For a fresh look at the postvocalic consonants traditionally regarded as onsets of the second syllable we rechecked and recounted just the 'variable' (C)VCVR subset; table 1 presents only the raw counts for the long and short PDE reflexes.

The distribution of lengthened and unlengthened forms in the 166 items with (C)V.CəR structure either in OE or in early ME is charted in figure 3. The subset of (C)V.RəR items shows practically no variability in the results for stems with a medial sonorant: 42 items are of the type *valor*, *venom*, *moral*, with only a single lengthened item: *moment*, whose etymology (Lat. *mōmentum*) may have interfered, making the intersyllabic sonorant in (C) V.RəR a factor which blocks OSL.[23] Further, there are no lengthened forms with –V.sCəR, where the medial clusters are also legal word-initial clusters: *-sp-, -st-, -sk-, -sm-*.

[21] Bermúdez-Otero (1998) treats all (C)V.CVR forms as one type of input, irrespective of the medial consonant. Lahiri & Dresher (1999) make the point that in both cases the moraic content of the second syllable is the same since the sonorant is the syllabic peak. As we will show below, their objection to Bermúdez-Otero's compensatory proposal, which extends the compensatory proposal in Minkova (1982) and is fully endorsed in the current study, disappears in the light of the new data on the interplay of syncope and sonority sequencing in σ_2.

[22] Bliss (1969 [1952/1953]: 196) noted in passing the absence of lengthening in Anglo-Norman loans with medial sonorants: *gallon*, *baron*, *gannet*, i.e. CV**R**VC, and he also observed that more recent borrowings show the same pattern (*helot*, *melon*, *zealot*), but he did not pursue the observation further.

[23] The *OED* dates the first attestation of *moment* to 1382. In ME verse the word is stressed initially, e.g. 'With baner reed is entred right anon / And in that selue ***móment*** Pálamón' (Chaucer, *KnT* 2583–4), or 'At the **m**ost in a **m**oment of his **m**old passe' (*Destr.Troy* (a.1400) (Htrn 388) 1820).

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 1. *Medial consonants in (C)V.CəR inputs to OSL*

	Long in PDE		Short in PDE	
Middle C	*OE*	*OF/AN*	*OE*	*OF/AN*
p	5	3	4	0
t	0	0	15	9
ʧ	0	0	1	2
k	2	1	6	2
b	1	4	1	1
d	3	1	7	0
g	0	0	0	1
f, s	1	2	1	3
v, z	9	6	8	10
ð	0	0	8	0
m, n	0	1	5	12
r, l	0	0	2	22
sp	0	0	1	0
st	0	0	4	0
sk	0	0	0	1
sm	0	0	1	0
Total		18	64	63

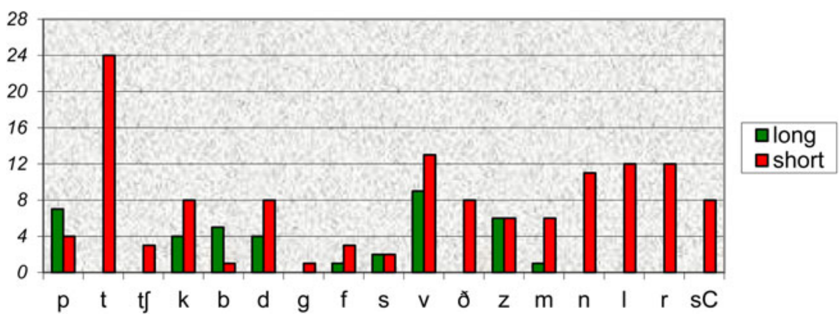

Figure 3. Medial consonants and lengthened/unlengthened (C)V.CəR inputs

Only 38 items (23 percent) in the database show a long vowel, if and only if the medial consonant is an *obstruent*. /t/ and /-ð-/ categorically inhibit lengthening.[24] Alveolar stop lenition, possibly resulting in tapping, can be traced back to early ME (6).

(6) Orthographic evidence for alveolar stop lenition in early English
OE *botm* ~ *bo**d**an* 'bottom', ME variant spellings *bo**þ**om*, *bo**dd**o*, *bo**d**me*, *bo**d**en*[25]

[24] All of the PDE /-ð-/ stems end in /-r/; historically they alternate with /-d-/ (Minkova 2014: 124–5), which also alternated with /-t-/ as in *otter*, *fetter*, *water* in (6). This may be the key to their parallel behavior in ME. We note, but do not explore further, the interesting phonotactics of /t, d, ð/ + (ə)r# in ME here.

[25] PDE northern English and Scots *bo**d**om(e)*, *bo**d**um*, *bo**dd**em* (*OED*).

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

OE Angl. *cetel*, WS *cytel* 'kettle', ME *ketel~che**d**el~chi**d**el~chi**th**el*,
(*Che**d**elhampton*, Devon 1242)
ME *oter* ~ *o**d**er-, other-* (1208, 1278, in place names, *MED*) 'otter'
ME *water(e)* ~ *wa**d**e-* as in *wa**d**etons*, *wa**d**ertons* 'water-casks'[26]
ME *catel* ~ *ca**d**el* 'property' (*c.* 1400)
ME *fete* 'fetter', *fe**d**erys* ~ *fe**þ**orys* 'fetters' (1440, 1450)
ME *potage* (1230) *po**d**ech* (1528) ~ *po**r**age* (1533) 'porridge'
ME *pottinger* (1415), *po**d**enger* (1463) ~ *po**rr**ynger* (1467) 'porringer'[27]

The data in (6) are fully compatible with ambisyllabicity of the intervocalic consonant, offering a previously unnoticed evidential angle on the early history of flapping in English. For now, if we take a bold step and project a lenited /-ɾ-/ realization of the alveolar stop, then we are looking at a segment whose features include [+sonorant] and [+continuant], which places it in the same set as (C)V.**R**VC words.[28]

Crucially, the new empirical finding on the behavior of (C)V.**R**VC stems adds a new perspective on both Ritt's (1994) focus on medial sonorants, and Bermúdez-Otero's (1998) focus on the presence of a second syllable sonorant coda and its link to the compensatory OSL model. Since variability occurs ***only*** in forms with both an intervocalic obstruent and a final sonorant, it is *their* exceptional behavior that calls for further scrutiny.

3.1 OSL and intersyllabic consonants

Figure 4 uses the same data as in table 1 and figure 3; (L) stands for 'lengthened', (S) stands for 'short'.

In the group shown in figure 4 the ratio of unlengthened to lengthened reflexes of items with a medial obstruent is 80:38; they are the locus of variability. Medial sonorants are a robust barrier to lengthening, *moment* being the only exception, and it is really not a good candidate for OSL, coming from Lat. *mōmentum.*

Before voiced consonants the vowel duration contributes to the perception of voicing. In our case medial /-b-, -d-, -v-, -z-/, in combination with a word-final sonorant, as in *gable*, *cradle*, *beaver*, *hazel*, would induce increased duration of the stressed vowel; perceived phonetic lengthening could then be reanalyzed as phonemic. This factor adds to the complexity of the overall picture, but it is clearly not the 'solution' of the

[26] (1399) Fabric R.Yk.Min.in Sur.Soc.35; note also <warter> (*MED*).

[27] Variability between [d] ~ [ɾ] is the most likely account for the current form of *paddock* < OE *pearroc* 'park, enclosed place'; the [d] and /r/ forms have coexisted since the middle of the sixteenth century. Positing early alveolar stop lenition for (C)V.**t**əR items (*bottom*, *cattle*) raises the question of why there are some (x4) lengthened (C)V.**d**əR forms: *cradle~craddle*, *ladle*, *treadle*, *odour*. We count them as 'lengthened', yet all of them have short variants: PDE *cradle~cra**dd**le*; ME *lādel* ~ *la**dd**el* n. (*MED*, *OED* frequency band 4), compare the verb *lade* (*OED* frequency band 5); ME *trēdel* ~ *tre**dd**le* (after the fifteenth century, *MED*, *OED*); *MED* **ŏ̄***dŏur* n. Since the lenited voiced alveolar dental would be spelled <-d> as in (6), it is only by consonant doubling that this suggestion can be tracked orthographically.

[28] Lenition may be the phonetic justification for the OE word for *Latin*: *le**d**en* ~ *læ**d**en* which the *OED* attributes to 'Celtic or early Romanic pronunciation of Latin *Latīnum* … confused with the native *léden*, *lýden*, *léoden* language, < *léode* people'.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

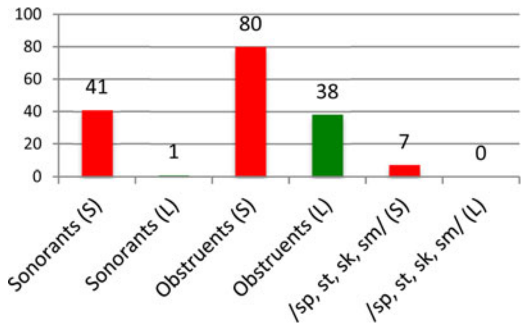

Figure 4. Sonorant vs obstruent medial consonants in (C)V.CəR inputs to OSL

problem of *travel* vs *haven*, *saddle* vs *ladle*. Nor does it tell us why (C)V.**T**əR items such as *bacon*, *capon*, *maple* were lengthened, while *baron*, *galon*, *florin* were not. So, based solely on the voicing/duration criterion of stressed (C)V[C]$_\sigma$, the uniform behavior of intervocalic sonorants is hard to account for – a problem noticed as 'particularly disturbing' and 'odd' by Ritt (1994: 40–1).

A summary of the stem-types discussed in sections 1–3 is shown in (7):

(7) A summary of stem-based OSL types

	Pre- OSL	Post-OSL		PDE example
(a)	(C)VCə	(C)VVCØ		*name, nose, meat*[29]
(b)	(C)VC**VV**	(C)VC**VV**		*many, body*[30]
(c)	(C)V**R**VC	(C)V**R**VC		*gannet, baron*
		(d₁) (C)V**T**V**T**		*habit, rocket*
(d)	(C)V**T**VC		(C)V**TVR**	*copper, heaven*
		(d₂) (C)V**TVR**		
			(C)VV**TVR**	*paper, haven*

This completes the survey of the disyllabic stem-based data relevant to OSL.

[29] Recall that prevocalic elision is early and ubiquitous in Middle English, so that the CØ variants in (7) were present in speech prior to the full-blown development of OSL. For the record: out of the 12 OE CVC input monosyllables that lengthened (lengthening attributed to inflected forms), 9 have a sonorant coda and 8 end in a liquid: PDE *bare*, adj. *coal*, *crane*, *dale, hole*, *stare* 'starling', *(a)ware*, *weir*, *whale*. The systematic resistance to lengthening in (C)VRVC words of the *gannet*, *baron* type suggests that for the *bare*, *coal* subset the V + liquid lengthening is unrelated to OSL/paradigmatic variants, but is simply an effect of the vocalization of the liquid in the uninflected form.

[30] This subset includes 24 items. In addition to (C)V.C**VV** outcomes from OE *-iġ* (*many, body)* and the loan *jolly* (1281), Old French (OF) *jolif, joli*, it includes inflected OE *-wa* stems: OE *nearo*, inflected *nearow-* 'narrow', OE *f(e)alo,* inflected *fealew-* 'fallow', similarly *barrow, callow, harrow, mellow, sallow, shadow, yellow*; see also footnote 8.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 2. *Overall patterns of lengthening in input disyllables*

	Example	% long PDE reflexes
(C)VCə	*name, nose, meat*	100
(C)VCVV	*many, body*	0
(C)VRəC	*gannet, baron*	2
(C)VTəT	*habit, rocket*	3
(C)VTəR	*copper, paper*	32

4 A phonological account of MEOSL

The following section offers a phonological account of the general pattern of lengthening and non-lengthening in the disyllabic ME stems in (7), making use of Classical OT, and, to account for the word forms which exhibit variation, Maxent OT.

A quantified version of the pattern to be accounted for is summarized in table 2. Other than the (C)VTəR subset, these word forms are effectively categorical in terms of the length of their PDE reflexes, the only two exceptions being *moment* for the (C)VRəC subset, and *naked* for the (C)VTəT subset, which, as already discussed, may be long for etymological or morphological reasons, respectively.

4.1 An OT grammar for MEOSL

Lengthening in our account is treated as primarily compensatory: a response to syncope or apocope, hypothesized to have occurred contemporaneously with lengthening. Compensatory lengthening can be modeled in OT by way of a constraint motivating the deletion of schwa, and any sort of constraint which forbids the output from being moraically shorter than the input, motivating compensation. These constraints must dominate the relevant faithfulness constraints which forbid deletion and lengthening (see table 3).

Table 3. *Constraints and definitions*

Constraint	Definition
*ə	No ə in σ_2[31]
Max-μ	The output must not have fewer morae than the input
Max-V	Don't delete a vowel
ID[long]	Don't change the length of a vowel

[31] For the moment we keep this constraint context-free; it motivates both apocope and syncope. Later on, we will justify the need for a separate constraint, *ə#, which motivates apocope only.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 4. *OT tableau for (C)VCə, (C)VCVV*

		*ə	Max-μ	ID[long]	Max-V
Input: (C)VCə	(C)VCə	*!			
nose, *name*	(C)VC		*!		*
	☞ (C)VVC			*	*
Input: (C)VCVV	☞ (C)VCVV				
many, *body*	(C)VVCVV			*!	

Since we now have a constraint that involves counting morae, some discussion here of our assumptions regarding weight is warranted. We assume that a short vowel (or a syllabic sonorant) is associated with one mora, a long vowel with two, and that coda consonants are associated with a mora as well, regardless of their sonority. Crucially, we treat intervocalic consonants as ambisyllabic, and therefore as mora-bearing. For example, a CVCə ('name') input has three morae, and not two. Deleting the final schwa to get CVC would bring the count down to two, and only by lengthening the vowel to CVVC would the original weight be preserved.[32]

These constraints are enough to explain the distinction between (C)VCə (*nose*, *name*) words on the one hand and (C)VCVV (*body*, *many*) words on the other: in the former, lengthening is a by-product of apocope, while in the latter, there is no motivation to lengthen at all, since the faithful forms are completely unmarked (see table 4).

Among (C)VCəC words, lengthening occurs only in the (C)VTəR (*paper*) variety, and in these only variably. We follow Bermúdez-Otero (1998) in proposing that lengthening in the *paper* words is, just as for *name* and *nose*, primarily compensation for the deletion of schwa in the second syllable.[33] While the word-final sonorant in these words syllabifies in these cases (violating a low-ranked faithfulness constraint on consonant syllabicity), this syllabification is not enough to compensate for the mora lost to schwa deletion, since CVR syllables have both a nucleus and a coda, and CR̩ syllables do not. Tables 5 and 6 show how the identity of the syllabic specification can affect the realization of CR̩ syllables. The careful reader may here interject that the deletion of a vowel in the second syllable in ME was not always accompanied by compensatory lengthening: (C)VVTTə words like OE *sōfte*, ME *soft(e)* 'soft', OE *cēpte*, ME *kept(e)* 'kept', underwent apocope, but in fact had their first vowel shortened, violating Max-μ twice. We can straightforwardly

[32] The question of why the post-apocope competition between -VC (**nam*) and -VVC (*naam*) was resolved in favor of the latter is addressed by Bermúdez-Otero (1998) with reference to Final Consonant Extrasyllabicity, which in his account does not apply to original -VC monosyllables because of the Unstressable Word Syndrome: since the minimal word in ME is bimoraic, **na*<*m*> is ill-formed. Apocope, -VCə > -VCØ results in a floating mora which 'must dock onto the root-vowel because, through extrasyllabification, the final consonant becomes unavailable as a landing site' (1998: 187). It seems to us that allowing the abstract notion of extrasyllabicity *only* in the case of apocope is insufficiently motivated, while our trimoraic preservation target is grounded in the robust presence of -VVC monosyllables in the language (ME *bōn* 'boon', *gōt* 'goat', *chēp* 'bargain', *shēf* 'sheaf').

[33] For now we describe a grammar which derives lengthening in CVTəR forms, when in fact this occurs only variably. Grammars which capture this variation will be given in the following sections.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 5. *Constraint definition for ID(syllabic)*

Constraint	Definition
ID[syll][34]	Maintain the identity of the [syllabic] specification

Table 6. *OT tableau for (C)VTəR*

		*ə	MAX-μ	ID[long]	MAX-V	ID[syll]
Input: (C)VTəR	(C)VTəR	*!				
paper	(C)VTR̩		*!		*	*
	☞ (C)VVTR̩			*	*	*

account for these cases with reference to an independent process, namely ME shortening before consonant clusters (Jordan & Crook 1974; Lass 1992; Ritt 1992; Minkova 2014). This process can be explained in OT by an undominated constraint like *VVCC. Except for the ambivalently weighted *-sp*, *-st*, *-sk*, a problem throughout the history of English because they are variably cohesive (Minkova 1982: 42; 2014: 334), and the special case of the so-called OE Homorganic Cluster Lengthening, e.g. *cild* > *cīld* 'child', *grund* > *grūnd* 'ground' (Luick 1964; Lass 1992; Ritt 1992; Minkova 2014), there are no -VVCC# stems in ME, or in PDE except a handful of real outliers (*traipse*, *oops*, *yikes*).

Since deletion occurs in the (C)VVTTə forms, some constraint motivating deletion must dominate MAX-μ. We must explain, then, why no deletion occurs in the (C)VTəT (*rocket*) forms. The easiest way to do so is to distinguish syncope generally from apocope, and have a constraint which enforces the latter be undominated, such that, unlike syncope, apocope applies without exception (which is in fact the empirical pattern observed across ME). Apocope (*ə#) and *VVCC are defined in table 7.

The ranking *ə#, *VVCC > MAX-μ allows us to explain *kept*, as well as other word forms where apocope occurred without lengthening, such as (C)VVTə words like *pope* where the vowel in the first syllable was already long in ME and could not lengthen any further since English prohibits overlong vowels.

The ranking *VVCC, MAX-μ > *ə explains why (C)VTəT (*rocket*) words do not undergo syncope: if syncope were to apply with compensatory lengthening in these words, this would create a long vowel in exactly the sort of environment where ME shortens long vowels. Conversely, if syncope were to apply without compensatory lengthening, this would violate MAX-μ (table 8).

[34] Alternatively, a markedness constraint *[+cons, +syll], prohibiting syllabic consonants, could be used in place of ID[syll], as it would be violated by the same candidates in the tableaux that follow.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 7. *Constraint definitions for apocope and *VVCC*

Constraint	Definition
*ə#	No word-final ə (apocope)
*VVCC	No long vowels before [-syllabic] clusters

Table 8. *OT tableau for (C)VVTTə, (C)VVTə, (C)VTəT*

		*VVCC	*ə#	Max-μ	*ə	ID[long]	Max-V	ID[syll]
Input: (C)VVTTə	(C)VVTTə	*!	*		*			
kēpt(e)	(C)VVTT	*!		*			*	
	(C)VTTə		*!	*	*	*		
	☞ (C)VTT			**		*	*	
Input: (C)VVTə	(C)VVTə		*!		*			
pōpe	☞ (C)VVT			*			*	
Input: (C)VTəT	☞ (C)VTəT				*			
rocket	(C)VTT			*!			*	
	(C)VVTT	*!				*	*	

Recall that (C)VTəR (*paper*) words can undergo syncope with compensatory lengthening without violating *VVCC because their final sonorant can be syllabified, avoiding a coda cluster. Because English (like most languages outside North Africa and the Pacific Northwest) forbids syllabic obstruents,[35] the same strategy is not available for the CVTəT (*rocket*) words, where syncope inevitably creates a branching coda.

The last thing to be explained is why there is no sonorant syllabification in the (C)VRəC (*gannet*, *baron*) words. This can be explained with reference to the sonority sequencing principle, or SSP.[36] A number of constraints are associated with this principle, but we here need only make reference to any constraint or constraints which require a rise in sonority from onset to nucleus (table 9). For conciseness, we here define SSP-Onset-Nuc such that it is also violated by onsetless syllables (which have no sonority rise). We could just as well

[35] In optimality theoretic terms, this is ruled out by an undominated constraint *[+syll -son]. This constraint and the candidates that violate it have been omitted for brevity.

[36] A much simplified SSP hierarchy sufficient to account for our data is vowels (V) > sonorants (R) > obstruents (T). We are using SSP instead of the pair of constraints SSP, Onset as a shortcut. On the functional co-occurrence of SSP and Onset see Vennemann's pre-OT Head Law (1988: 13–14): 'A syllable head is the more preferred: (a) the closer the number of speech sounds in the head is to one, (b) the greater the Consonantal Strength value of its onset, and (c) the more sharply the Consonantal Strength drops from the onset toward the Consonantal Strength of the following syllable nucleus.' While part (a) overlaps with onset, parts (b) and (c) translate directly into the SSP constraint as used here.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 9. *Constraint definition for the Sonority Sequencing Principle*

Constraint	Definition
SSP	Syllables must rise in sonority from the onset to the nucleus (and also must have an onset).

have used an additional constraint, Onset, to be added to the same stratum, and this pair of constraints would achieve exactly the same effect.

The medial sonorants of (C)VRəC (*gannet*, *baron*) words cannot syllabify because σ_2 would be [R̩C] with no onset, violating SSP (or Onset, see above). Similarly, if the final sonorants in the (C)VRəR (*baron*) words were to syllabify, it would create a syllable of the form [RR̩], which does not rise in sonority.[37]

Essentially, we are proposing that syncope is sensitive to the relative sonority of the consonants that it would bring into contact. Reassuringly, this exact sensitivity has been observed for trisyllabic syncope in PDE: Polgárdi (2015: 383) finds that 'English post-tonic syncope is sensitive to the quality of the flanking consonants: the consonant following the alternating vowel must be a sonorant which is more sonorous than the consonant preceding it (*dél*(i)cate*, *cól*(o)ny*).'

With the addition of SSP, all of the data can be accounted for, as shown in table 10.

4.2 Modeling variation with classical OT

The analysis in the previous section derives what kind of word forms *can* lengthen. In particular, it derives lengthening in the (C)VTəR (*paper*, *copper*) cases, when in fact lengthening in these forms occurs only variably.

In order to better understand how this variation could result from variable constraint ranking, we employed the constraint-ranking software OTSoft (Hayes *et al.* 2013). This software, when given a tableau indicating winning candidates and constraint violations, finds a constraint ranking which predicts the winners. We ran the software twice: first on the tableau in table 10, where the winning candidate for the (C)VTəR input was set to be categorically (C)VVTR̩ (deriving the *paper* cases), and then on a similar tableau in which the winning candidate for (C)VTəR was the faithful one (deriving the *copper* cases). OTSoft gave the constraint rankings shown in table 11, which we'll refer to as the 'paper' grammar and the 'copper' grammar, respectively.

[37] A reviewer points out that not all sonorant consonants necessarily have the same sonority, such that some sonorant clusters like <ml> could be argued to involve a sonority rise. While this is true in a phonetic sense, we can find no evidence that syncope is sensitive to the difference between nasals and liquids with regards to their sonority, though we do not rule out the possibility that other linguistic patterns based on sonority are sensitive to this distinction. For example, while it can be argued that *camel* (the only <-ml> sequence in the data in (4b)) has a rise in sonority (and should therefore conform to the SSP), more than half of the words with -.RR̩ σ2 do not, assuming equal sonority for [n, m]. These items (of the type *barrel*, *baron*, *moral*, *gallon*, *felon*, *venom*, *canon*, *pennon* …) suggest that the differences in sonority in -.RR̩ σ2 do not reach a sufficiently robust threshold to affect the outcome.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 10. *OT tableau for all (C)VCV(C) forms*

		SSP	*VVCC	*ə#	Max-μ	*ə	ID[long]	Max-V	ID[syll]
Input: (C)VCə	(C)VCə			*!		*			
nose, *name*	(C)VC				*!			*	
	☞ (C)VVC						*	*	
Input: (C)VCVV	☞ (C)V.CVV								
many, *body*	(C)VV.CVV						*!		
Input: (C)VTəT	☞ (C)V.TəT					*			
rocket	(C)VTT				*!				
	(C)VVTT		*!				*	*	
Input: (C)VTəR	(C)V.TəR					*!			
paper	(C)V.TR̩				*!			*	*
	☞ (C)VV.TR̩						*	*	*
Input: (C)VRəT	☞ (C)V.RəT					*			
gannet	(C)VRT				*!			*	
	(C)VVRT		*!				*	*	
	(C)VV.R̩T	*!					*	*	*
Input: (C)VRəR	☞ (C)V.RəR					*			
baron	(C)VRR				*!			*	
	(C)VVRR		*!				*	*	
	(C)VV.R̩R	*!					*	*	*
	(C)VV.RR̩	*!					*	*	*

Table 11. *Constraint ranking and strata learned by OTSoft for tableaux which differ only in whether lengthening applies to (C)VTəR forms*

'paper' grammar		'copper' grammar	
	SSP		SSP
Stratum #1	*VVCC	Stratum #1	*VVCC
	*ə#		*ə#
Stratum #2	Max-μ		ID[syll]
Stratum #3	*ə	Stratum #2	Max-μ
	ID[syll]		Max-V
Stratum #4	Max-V	Stratum #3	*ə
	ID[long]		ID[long]

Note that a 'stratum' is a set of constraints whose relative ranking does not affect the outcome. However, even this representation is over-specified, in that not all of the pairwise rankings implied by these strata are necessary. For example, while *VVCC is in a higher stratum than Max-V in both grammars, neither of the tableaux rely on the relative ranking of these two constraints. A better understanding of how the grammars work can therefore be gleaned from their Hasse diagrams (also generated by OTSoft), which represent the constraint rankings that are necessary to derive the winners. The

Hasse diagrams show more information than the constraint ranking. For example, SSP is in the top stratum because it is undominated, but it also doesn't strictly dominate anything else, which is why it's isolated in the diagram: it could be in any stratum and the grammar would make the same predictions. Hasse diagrams for the two grammars are given in figure 5.

From figure 5 we can see that the only necessary change in constraint ranking to block lengthening in the (C)VTəR words is the demotion of *ə (syncope) with respect to at least one of the constraints that is violated when syncope with compensatory lengthening applies: Max-V, ID[long] or ID[syll] (in conjunction with Max-μ). If *any one* of these three dominates *ə, then the change will be blocked. Thus, two grammars with minimal differences in constraint ranking could produce different outcomes in the (C) VTəR cases, holding the other cases constant.

4.3 Modeling variation with Maxent OT

The analysis presented in section 4.2 is sufficient if the lengthening variation in these cases is to be explained as variation between grammars: this could be synchronic inter-speaker variation, dialect variation, intra-speaker stylistic variation, word-specific diachronic variation due to lexical factors like frequency, or any number of other things.

However, stochastic varieties of OT are also capable of modeling variation using a single grammar, corresponding to the possibility of true free variation in these forms, even for individual speakers, at the time of the change. Maxent OT (Goldwater & Johnson 2003) is one such stochastic variety, in which constraints have weights rather than a strict ranking. Maxent grammars do not predict a winning candidate, but rather a probability (Pr) for each candidate based on its harmony (h), the weighted sum of its constraint violations, such that candidates with more violations of higher weighted constraints are less probable.

OTSoft is capable of learning the weights for a maxent grammar. The input is much the same as before – a tableau with candidates and constraint violations, but rather than indications of winning candidates, observed counts for candidates must be provided. We once again used the tableau in table 10, with an additional column indicating the number of attested stems matching the form of each candidate, summarized in table 12.[38]

Using the tableau in table 10 augmented with the candidate counts in table 12, and with maximum constraint weight capped at 500,[39] OTSoft learned the constraint weights in table 13.

A maxent grammar with these learned weights predicts candidate probabilities which match the patterns seen in the data almost exactly: it predicts categoricity for the categorical cases, and predicts variation for the *paper*/*copper* cases, with lengthening

[38] The exceptional forms *moment* and *naked* were omitted, as length in these words may be explained with recourse to non-phonological factors.

[39] Since some constraints are never violated, their optimal weights in a maxent grammar would be infinite, so a maximum allowable weight must be specified in advance to ensure convergence.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

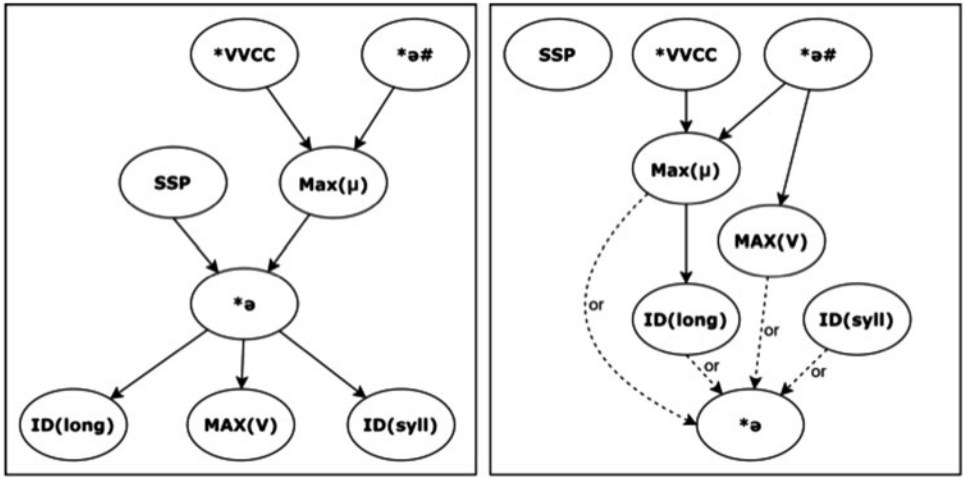

Figure 5. Hasse diagrams for the ‘paper’ grammar in which (C)VTəR words show lengthening (left) and for the ‘copper’ grammar in which (C)VTəR words do not show lengthening (right). In the latter, the dotted lines indicate rankings which need not all be present. In particular, any of the following would suffice to predict the data: MAX-V > *ə, ID[long] > *ə, or ID[syll], MAX-μ > *ə

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 12. *Candidate counts given to OTSoft for fitting the weights of a maxent grammar (candidates given counts of 0 are not shown in this table)*

Input	Output	Types
(C)VCə	—> (C)VVC (*nose, name*)	192
(C)VCVV	—> (C)VCVV (*many, body*)	24
(C)VTəT	—> (C)VTəT (*rocket*)	17
(C)VTəR	—> (C)VTəR (*copper*)	76
	—> (C)VVTR̩ (*paper*)	36
(C)VRəT	—> (C)VRəT (*gannet*)	
(C)VRəR	—> (C)VRəR (*baron*)	41

Table 13. *Constraint weights learned by OTSoft*

Constraint	Weight	Constraint	Weight
SSP	500	ID[long]	7.36
*ə#	500	*ə	6.61
*VVCC	500	ID[syll]	0
Max-μ	500	Max-V	0

occurring 32 percent of the time. A Maxent OT tableau with the learned weights is given in table 14.

This maxent grammar accomplishes its close fit to the data by relying on the interplay between ID[long] and *ə (syncope), which are assigned similar, comparatively small weights.[40] Because ID[long] has a small weight, changes to length will practically always occur when necessary to satisfy stronger constraints like *ə# (apocope) and *Max-μ. However, because syncope has a similarly small weight, ID[long] is enough to prevent syncope with compensatory lengthening from occurring 68 percent of the time. Syncope without compensatory lengthening is ruled out by the much higher weight of Max-μ.

Note that while the weights learned by OTSoft are mathematically 'optimal' in terms of their ability to match the observed, there could be alternative values for the weights that perform just as well, and derive the same qualitative pattern: for example, if ID[syll] or Max-V, rather than ID[long], had been given a weight similar to *ə, the interplay between that constraint and *ə might have resulted in variation in the (C)VTəR words in much the same way, since lengthening in these cases also involves the syllabification of sonorants and the deletion of vowels.

[40] Recall that since *ə# is defined in table 7 independently of *ə, our references to the relative ranking of *ə# and *ə imply that *ə, when it is the only violated constraint in items which remain disyllabic, refers to syncope, as in the winning candidates for *rocket* in table 8, *gannet*, *baron* in table 9.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Table 14. *Maxent tableau with weights learned by OTSoft. h = harmony (the weighted sum of a candidate's violations), Pr = predicted probability*

		SSP	*VVCC	*ə#	Max-μ	ID [long]	*ə	Max-V	ID [syll]		
		500	500	500	500	7.47	6.72	0	0	h	Pr
Input: (C) VCə *nose*, *name*	(C)VCə			1			1			506.72	0.00
	(C)VC				1			1		500	0.00
	(C)VVC					1		1		7.47	1.00
Input: (C) VCVV *many*, *body*	(C) V.CVV									0	1.00
	(C) VV.CVV					1				7.47	0.00
Input: (C) VTəT *rocket*	(C)V.TəT						1			6.72	1.00
	(C)VTT				1					500	0.00
	(C) VVTT		1			1		1		507.47	0.00
Input: (C) VTəR *paper*	(C) V.TəR						1			6.72	0.68
	(C)V.TR̩				1			1	1	500	0.00
	(C) VV.TR̩					1		1	1	7.47	0.32
Input: (C) VRəT *gannet*	(C) V.RəT						1			6.72	1.00
	(C)VRT				1					500	0.00
	(C) VVRT		1			1		1		507.47	0.00
	(C) VV.R̩T	1				1		1	1	507.47	0.00
Input: (C) VRəR *baron*	(C) V.RəR						1			6.72	1.00
	(C)VRR				1					500	0.00
	(C) VVRR		1							500	0.00
	(C) VV.R̩R	1							1	500	0.00
	(C) VV.RR̩	1							1	500	0.00

4.3.4 Future work

In our account, the OE (C)VCə inputs (*nose*, *name*) map to (C)VVC rather than (C)VC outputs due to a constraint which prefers preservation of word-level metrical weight in the face of apocope (table 4). A related, lexicon-specific question is whether the (C) VVC candidates could have been preferred because the lengthened vowel guaranteed better preservation of lexical contrast. We leave this line of inquiry for further research,

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

but we note that post-apocope moraic preservation maintains contrast with pre-existing CVC monosyllables in the lexicon: *bane–ban, bake–back, bale* 'evil' – *ball, bathe–bath, bead–bed, break–brack, cloak–clock, dean* 'valley' – *den, vane–fan, hate–hat, hope–hop, mane–man, lame–lamb, lease–less, mate–mat, meat–met, sake–sack, snake–*†*snack, soak–sock, stake–stack, tale–tell, weal–well, wane–wan*… All of these -VV-:-V- minimal pairs are pre-1400. The (C)VC–(C)VVC contrast in ME carried very high 'functional load' in ME. Wedel *et al.* (2013) have provided statistical evidence that 'the probability of phoneme merger is inversely related to functional load'. Another inquiry in that context would address the functional load of the contrast between pre-existing and newly lengthened (C)VVC items. We have not searched for minimal pairs of that type, but we note that potential mergers might be disfavored. The hyperarticulation which would initially characterize post-apocope lengthening would not necessarily be perceived as identical to a pre-existing target. In the set of the pre-lengthening inputs (Minkova 1982: 40–1), items with a low vowel, [a∼æ] in the stressed syllable represent over 50 percent of the entire database. The post-lengthening vowel [aː] was distinct from pre-existing OE [ɒː] > ME [ɔː], northern and Scottish [ɑː]. Lengthened OE [ɛ] could only merge with the monophthongized OE [ɛə], but not with OE [eː], and lengthened OE [ɔ] could only merge with the reflex of ME [ɔː] < OE OE [ɒː], but not with OE and ME [oː]. As noted in section 1, rhyme evidence for a merger of newly lengthened vowels with the pre-existing long vowels is unambiguous only after the middle of the fifteenth century. Not least, the potential functional load imbalance would be useful in establishing the distinctiveness of vowel quantity vs quality in providing information.

5 Summary and conclusion

As our title suggests, we propose a new descriptive name for the process discussed in this study: Middle English Compensatory Lengthening (MECL). In Middle English, the overall pattern of lengthening in (C)VCə(C) stems can be explained phonologically by postulating that lengthening in these cases was moraic compensation for the deletion of a stressless vowel in the following syllable. The open-syllable condition for lengthening is, under this account, completely epiphenomenal: in closed monosyllabic stems, no lengthening will occur because there is no deletion to be compensated for, while in closed first syllables of disyllabic inputs like *ende* 'end' the vowel is followed by a word-medial cluster, and vowels before clusters shorten in ME. The constraints used for the (C)VCəC inputs apply also to other disyllabic inputs (*kēpte*, *pōpe)* and can be generalized over much of the ME and PDE lexicon.

The fact of pre-cluster shortening in ME simultaneously explains why, of the (C)VCəC words, lengthening only occurs when the final consonant is a sonorant: after syncope, sonorant syllabification is necessary to avoid creating a cluster following a long vowel. Lengthening is less frequent in (C)VTəR (*paper*, *copper*) words than in (C)VCə (*nose*, *name*) words precisely because deletion is less frequent, which can be attributed to the fact that syncope, unlike apocope, is an optional process in ME.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

The various factors that have been described as creating large classes of 'exceptions' to lengthening in open syllables, such as the segmental identity of the following consonant and the nucleus and coda of the second syllable, are demystified: these factors straightforwardly relate to deletion in the second syllable, the trigger for lengthening.

Revisiting the entire set of (C)VCəC inputs and their reflexes in PDE in the context of the discovery that sonorant σ_2 onsets inhibit lengthening has led to a promising side-line to the main points in the article: the alignment of /-t-/ with the sonorants in σ_2 onset position (figure 3 and the attestations in (6)). The parallel results for the *gannet*, *baron* types and the *bottom*, *fetter* types is a previously unidentified argument for projecting alveolar stop lenition, a central process in PDE, back to at least Middle English, though not readily recognizable in the written records.

In sum, a re-examination of the database and a constraint-based analysis supports Bermúdez-Otero's (1998) insight regarding the compensatory potential in (C)VCəR inputs and accounts for the previously puzzling variability of the outcome. It also allows a broader perspective on the diachronic conditions shaping the phonology of Present-day English.

Authors' addresses:

Department of English UCLA
University of California, Los Angeles
149 Kaplan Hall
Los Angeles, CA 90095
USA
minkova@HUMNET.UCLA.EDU

357 Castro St
Suite 200
Mountain View, CA 94041
USA
lmlefkowitz@gmail.com

References

Bermúdez-Otero, Ricardo. 1998. Prosodic optimization: The Middle English length adjustment. *English Language and Linguistics* 2(2), 169–97.

Bliss, Alan J. 1969 [1952/1953]. Vowel quantity in Middle English borrowings from Anglo-Norman. In Roger Lass (ed.), *Approaches to English historical linguistics: An anthology*, 164–208. New York: Holt, Rinehart and Winston.

CoNE: *A Corpus of Narrative Etymologies from Proto-Old English to Early Middle English and accompanying Corpus of Changes*, version 1.1, 2013–. Comp. Roger Lass, Margaret Laing, Rhona Alcorn & Keith Williamson. www.lel.ed.ac.uk/ihd/CoNE/CoNE.html

Dobson, Eric J. 1957. *English pronunciation: 1500–1700*, vol. II: *Phonology*. Oxford: Clarendon Press.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Dobson, Eric J. 1962. Middle English lengthening in open syllables. *Transactions of the Philological Society,* 61(1), 124–48.
Fulk, Robert D. 1992. *A history of Old English meter.* Philadelphia: University of Pennsylvania Press.
Goldwater, S. & M. Johnson. 2003. Learning OT constraint rankings using a maximum entropy model. In Jennifer Spenader, Anders Eriksson & Osten Dahl (eds.), *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*, 111–20.
Hanson, Kristin. 2001. Quantitative meter in English: The lesson of Sir Philip Sidney. *English Language and Linguistics* 5(1), 41–91.
Hayes, Bruce, B. Tesar & K. Zuraw. 2013. *OTSoft 2.3.2* (software package). www.linguistics. ucla. Edu/people/hayes/otsoft
Hogg, Richard M. 1996. Old English open syllable lengthening. *Transactions of the Philological Society* 94(1), 57–72.
Hogg, Richard M. & R. D. Fulk. 2011. *A grammar of Old English*, vol. 2: *Morphology.* Chichester: John Wiley & Sons.
Ikegami, M. T. 1984. *Rhyme and pronunciation: Some studies of English Rhymes from Kyng Alisaunder to Skelton*. Hogaku-Kenkyu-Kai: Keio University, Extra Series 5.
Jordan, Richard & Eugene Joseph Crook. 1974. *A handbook of Middle English grammar: Phonology*, trans. and rev. by Eugene Joseph Crook (Janua Linguarum Series Practica). The Hague: Mouton.
Lahiri, Aditi & B. Elan Dresher. 1999. Open syllable lengthening in West Germanic. *Language* (75), 678–719.
Lass, Roger. 1992. Phonology and morphology. In Norman Blake (ed.), *The Cambridge history of the English language*, vol. 2, 23–155. Cambridge: Cambridge University Press.
Lehiste, Ilse. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
Liberman, Anatoly. 2015. The shortest history of vowel lengthening in English. In Michael Adamas, Laurel Brinton & R. D. Fulk (eds.), *Studies in the history of the English language VI* (Topics in English Linguistics 85), 161–83. Berlin: De Gruyter Mouton.
Luick, Karl. 1964. *Historische Grammatik der englischen Sprache*, vol. 1, 2 (1914–40), Leipzig: Tauchnitz (reprint).
MED: *Middle English Dictionary.* Online edition. Ann Arbor: University of Michigan Library, 2000–18. http://quod.lib.umich.edu/m/middle-english-dictionary/
Minkova, Donka. 1982. The environment for open syllable lengthening in Middle English. *Folia Linguistica Historica* 3(1), 29–58.
Minkova, Donka. 1985. Of rhyme and reason: Some foot-governed quantity changes in English. In Roger Eaton, Olga Fischer, Willem Koopman & Frederike van der Leek (eds.), *Papers from the Fourth International Conference on Historical Linguistics*, 162–79. Amsterdam and Philadelphia: John Benjamins.
Minkova, Donka. 2014. *A historical phonology of English*. Edinburgh: Edinburgh University Press (UK); New York: Oxford University Press (USA).
Minkova, Donka. 2015. A U-turn and its consequences for the history of final schwa in English. In Christina Sanchez-Stockhammer (ed.), *Studies in Variation, Contacts and Change in English* (VARIENG) 16. www.helsinki.fi/varieng/journal/volumes/16/minkova/
Minkova, Donka. 2016. From stop-fricative clusters to contour segments in Old English. In Don Chapman, Colette Moore & Miranda Wilcox (eds.), *Studies in the history of the English language VII. Generalizing vs. particularizing methodologies in historical linguistic analysis* (Topics in English Linguistics 94), 29–59. Berlin: De Gruyter Mouton.
Minkova, Donka. 2019. Examining the evidence for phonemic affricates: ME /t͡ʃ/, /d͡ʒ/ or [t-ʃ], [d-ʒ]? In Rhona Alcorn *et al*. (eds.), *Historical dialectology in the digital age*, 156–84. Edinburgh: Edinburgh University Press.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press

Mokrowiecki, Tomasz. 2015. Acute accents as graphic markers of vowel quantity in two Late Old English manuscripts. *English Language and Linguistics* 19(3), 407–36.

Ogura, Mieko. 1987. *Historical English phonology: A lexical perspec*tive. Tokyo: Kenkyusha.

Page, B. Richard. 2006. The diachrony and synchrony of vowel quantity in English and Dutch. *Diachronica* 23(1), 61–104.

Peterson, G. E. & I. Lehiste. 1960. Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32(6), 693–703.

Polgárdi, Krisztina. 2015. Syncope, syllabic consonant formation, and the distribution of stressed vowels in English. *Journal of Linguistics* 51, 383–423.

Ritt, Nikolaus. 1988. The processes amounting to MEOSL and its exceptions. In Manfred Markus *et al*. (eds.), *Historical English: On the occasion of Karl Brunner's 100th birthday*, 153–66. Innsbruck: Innsbrucker Beiträge zur Kulturwissenschaft.

Ritt, Nikolaus. 1992. Middle English vowel quantity reconsidered. In Matti Rissanen, Ossi Ihalainen & Terttu Nevalainen (eds.), *History of Englishes: New methods and interpretations in historical linguistics* (Topics in English Linguistics), 207–22. Berlin: De Gruyter Mouton.

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

Ritt, Nikolaus. 1997. Now you see it, now you don't: Middle English lengthening in closed syllables. *Studia Anglica Posnaniensia* 31, 249–70.

Ritt, Nikolaus. 2000. Theory, empiry, and textual witnesses: Acutes in the Lindisfarne Gospel and Early English vowel quantity. *Sprachwissenschaft* 25(4), 497–512.

Tarlinskaja, Marina. 2014. *Shakespeare and the versification of English drama, 1561–1642*. Farnham: Ashgate .

Vennemann, Theo. 1988. *Preference laws for syllable structure: And the explanation of sound change with special reference to German, Germanic, Italian, and Latin*. Berlin: Walter de Gruyter.

Walker, John. 1797. *A critical pronouncing dictionary, and expositor of the English language…: Abridged and adapted to the use of citizens of the United States*. New York: Alsop, Brannan and Alsop.

Wedel, Andrew, Abby Kaplan & Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2), 179–86.

Westbury, J. & P. Keating. 1980. Central representation of vowel duration. *Journal of the Acoustical Society of America* 67, S37.

https://doi.org/10.1017/S1360674319000522 Published online by Cambridge University Press