



PROJECT MUSE®

Old English vowels: Diachrony, privativity, and phonological representations

Thomas C. Purnell, Eric Raimy, Joseph Salmons

Language, Volume 95, Number 4, December 2019, pp. e447-e473 (Article)

LANGUAGE
A JOURNAL OF THE LINGUISTIC SOCIETY OF AMERICA

ISSUE	PAGES
Editorial Introduction	1-2
Old English vowels: Diachrony, privativity, and phonological representations	3-16
Book Reviews	17-20
Index	21-22

Published by Linguistic Society of America

DOI: <https://doi.org/10.1353/lan.2019.0083>

➔ *For additional information about this article*

<https://muse.jhu.edu/article/743116>

RESEARCH REPORT

Old English vowels: Diachrony, privativity, and phonological representations

THOMAS C. PURNELL

ERIC RAIMY

JOSEPH SALMONS

*University of Wisconsin –
Madison*

*University of Wisconsin –
Madison*

*University of Wisconsin –
Madison*

Old English underwent diachronic change in its vowel inventory between its predecessor West Germanic and Middle English. We provide an analysis of the addition and loss of vowels in Old English from the perspective of MODIFIED CONTRASTIVE SPECIFICATION (Dresher et al. 1994). Three main themes emerge from our analysis: (i) the phonological representation of contrast in the vowels in English has remained remarkably stable for over a thousand years, (ii) the proposed analysis improves upon and supersedes similar analyses proposed in Dresher 2015 and Purnell & Raimy 2015, and (iii) the adoption of privative features provides an improved representationally based understanding of phonological activity, feature geometry, and how phonology reflects general cognitive features of memory.*

Keywords: contrast, diachrony, distinctive features, privative features, Old English, successive division algorithm, phonological activity

1. INTRODUCTION. The goal of this research report is to analyze early English vowel changes in terms of phonological representation, highlighting the role of contrast and feature geometry, with two major results. First, following Oxford's (2015) proposals about diachronic change in phonological contrast, we argue that the core contrastive feature system underlying English vowels has remained remarkably stable for a millennium. Second, we provide new support for the CONTRASTIVIST HYPOTHESIS (Hall 2007) and privative features, developing responses to potential challenges to both. Our analysis is situated within MODIFIED CONTRASTIVE SPECIFICATION representation (Dresher et al. 1994). We propose a novel way of interpreting the unmarked side (null or 'Ø') of a privative contrast, thereby refining the concept of 'phonological activity', crucial for modified contrastive specification and the contrastivist hypothesis (Nevins 2015).

The *Handbook of historical phonology* (Honeybone & Salmons 2015) contains two analyses of early English vowel changes. Dresher 2015 (with further development in Dresher 2018) and Purnell & Raimy 2015 propose drastic changes to the contrastive hierarchies expressing the feature inventory leading to the Old English (OE) vowel system. We provide a more conservative analysis, while accounting for the generalizations of both analyses. Many questions about how contrast changed in OE are rendered moot because the specific contrastive hierarchy of English remains stable and changes to OE vowel phonology are expressed through rule addition, rule deletion, rule reordering, or phonetics, as opposed to change in contrast.

Modified contrastive specification (Dresher et al. 1994) and the proposals of Dresher 2015 and Purnell & Raimy 2015 are reviewed in §2. Subsequently, §3 discusses metrics for comparing different analyses of diachronic change, based on proposals made by Oxford (2015). The relevant changes from Proto-Germanic to the OE vowel system, important phonological rules of OE vowels, and our analysis appear in §4, while §5 develops proposals about privative feature representation in an MCS system. Finally,

* A very early version of this paper was presented at the Germanic Linguistics Annual Conference, University of Iceland, May 2016. In addition to that audience, we thank the referees and editors of this journal for helping sharpen the paper, as well as many students for conversations about the ideas here and Elan Dresher for comments on a draft. We are grateful to Bill Idsardi for bringing Jaccard distance to our attention. All errors are our own.

we discuss the ramifications for the contrastivist hypothesis and feature geometry (§6) and then conclude (§7).

2. MODIFIED CONTRASTIVE SPECIFICATION AND THE SUCCESSIVE DIVISION ALGORITHM. Modified contrastive specification (MCS; Dresher et al. 1994) and the SUCCESSIVE DIVISION ALGORITHM (SDA; Dresher 2009) are principles for understanding contrast in phonological inventories, and Oxford (2015) extends these to understanding diachronic change. The MCS represents the minimal difference between phonemes from the application of the SDA that occurs when a child acquires a phonological inventory. Because the SDA models an acquisition procedure, it is committed neither to any particular distinctive feature set nor to any position on binary vs. privative features. An informal version of the SDA (Dresher 2009:16), modified by us for clarity, is presented in 1. The resultant hierarchy is represented only by distinctive features that actively provide a contrastive ‘branch’. Features may not be identical with successive divisions down any one arm of the hierarchy, compared to another arm. Furthermore, the hierarchy cannot consist of vacuous divisions of features; only the features that do work are represented on the tree (and presumably learned by the child acquiring the language).

(1) Successive division algorithm (clarified)

- a. Begin with NO feature specifications: assume all sounds are allophones of a single undifferentiated phoneme.
- b. If the phone set is found to consist of more than one contrasting member, select a distinctive feature to characterize the contrast and divide the set into as many subsets as the features allow.
- c. Repeat step 1b in each subset: keep dividing the inventory into sets, applying successive features in turn, until every set has only one member.

We adopt Oxford’s (2015:311) pursuit of a privative model of distinctive features that ‘makes the model maximally restrictive, since it predicts that only the marked values of contrastive features will be phonologically active’ (see §3). This use of privative features magnifies the importance of identifying which phonological processes are active in a language. The SDA is able to produce multiple distinct featural encodings for any given inventory. For example, 2 presents some ways a three-vowel inventory of /i, u, a/ could be featurally coded by the SDA.

(2) Three-vowel system /i, u, a/

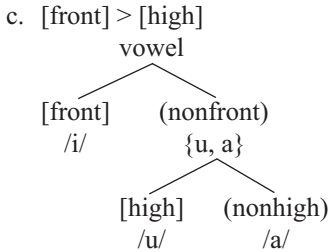
- a. [low] > [round]


```

graph TD
    vowel --> low["[low]  
/a/"]
    vowel --> nonlow["(nonlow)  
{i, u}"]
    nonlow --> round["[round]  
/u/"]
    nonlow --> nonround["(nonround)  
/i/"]
      
```
- b. [high] > [front]


```

graph TD
    vowel --> high["[high]  
{i, u}"]
    vowel --> nonhigh["(nonhigh)  
/a/"]
    high --> front["[front]  
/i/"]
    high --> nonfront["(nonfront)  
/u/"]
      
```



d.

VOWEL	HIERARCHY A	HIERARCHY B	HIERARCHY C
/i/	(nonlow), (nonround)	[high], [front]	[front]
/u/	(nonlow), [round]	[high], (nonfront)	(nonfront), [high]
/a/	[low]	(nonhigh)	(nonfront), (nonhigh)

In 2, we illustrate how a single inventory such as /i, u, a/ can be phonologically represented in many different ways while satisfying the assumptions of MCS and following the SDA operation. Since there are three vowels in this inventory, two features will be needed to encode the phonological contrasts in this system. First, 2a shows how the SDA would encode the inventory if first the feature [low] was selected and then [round]. We use the notation [low] > [round] to indicate the features and hierarchical order of selection by the SDA and refer to this arrangement of features as a ‘contrastive ranking’. The tree structure graphic below this contrastive ranking in 2 visualizes the division steps from the application of the SDA. We refer to this graphic as a ‘contrastive hierarchy’, and the tree graphic in 2a shows that choosing the feature [low] first separates /a/ from the set of {i, u}. /a/ in 2a is now uniquely specified so does not require any more phonological features. The set of {i, u} remains undifferentiated. After using the feature [round], both /u/ and /i/ are uniquely specified, so the SDA stops.

With 2b we show that a phonetically identical inventory can have a different phonological specification depending on what features are chosen by the SDA. The contrastive ranking is [high] > [front], which produces the contrastive hierarchy where first the set of {i, u} is marked with [high], and then [front] marks /i/ as distinct from /u/. Finally, 2c demonstrates that the order of the features in a contrastive ranking matters. Here, [front] is used first, which uniquely specifies /i/, and then [high] distinguishes between /u/ and /a/.

We give in 2d a third way of conceptualizing how the SDA creates phonological specification by showing what phonological features are assigned to each phoneme by the three different contrastive rankings in 2a–c. Each column gives the phonological features assigned to each phoneme based on the particular contrastive ranking. Looking across the row for /i/, this vowel can be specified as ‘nonlow and nonround’ or [high] and [front] or just [front]. The other vowels, /u/ and /a/, both show similar types of variability in features. This variability in phonological specification is the main tenet of MSC (Dresher et al. 1994, Dresher 2009), where contrast drives the phonological specification of phoneme inventories. Contrastive rankings and hierarchies produced by the SDA create different language-specific representations for phonemes needed to account for a language’s phonology. Avery and Rice (1989:179) take the strong position on the importance of representations: ‘rules must follow from representations rather than *vice versa*. This is consistent with the position that the burden of explanation in phonology should be in the representational component rather than the rule component’. This point of view in phonology places a premium on phonological activity or processes as data that determines what the particular contrastive ranking is for a particular language. Given our examples in 2, a language that has a lowering process triggered by /a/ would

require the representations produced by 2a, because /a/ is specified with [low]. The other two contrastive rankings do not specify /a/ with phonological substance that would explain a lowering rule. The presence of a lowering process indicates that the feature [low] must be specified in the phonology. Accordingly, a palatalization process triggered by /i/ would require the representations from either 2b or 2c because /i/ is specified for [front] in each of them (or only 2b if both [high] and [front] are required for the palatalization process). The idea is that the linguist and the learner can use phonological activity to determine both the features selected by the SDA and the order in which the features are chosen.

The contrastivist hypothesis states that ‘the phonological component of language L operates only on those features which are necessary to distinguish the phonemes of L from one another’ (Hall 2007:20). We adopt the contrastivist hypothesis because it provides the strongest version of the connection between phonological activity and specification. Consequently, different contrastive hierarchies/rankings can be compared and evaluated based on whether they produce segmental representations accounting for the phonological processes in a particular language. Not all possible contrastive rankings or hierarchies produce representations that match a specific language’s phonology, so part of the acquisition of phonology is a learner determining which contrastive ranking or hierarchy matches the ambient language.

A final point before we turn to OE and its diachrony is the question of binary or privative distinctive features. The way we present the features in 2 is noncommittal on this fundamental question for now. We adopt this approach from Dresher (2014), who provides additional examples and discussion about the representation of an /i, u, a/ inventory. Later, we abandon this noncommittal position and argue for privative features. On the one hand, the noncommittal approach to distinctive features in 2 suggests a privative stance because there is a clear marked feature (e.g. [low]) without any +/- and then an unspecified feature in parentheses (e.g. ‘(nonlow)’). But the presence of the ‘(nonlow)’ specification, in fact, suggests a binary [+low] vs. [-low] approach because both sides have biased phonological content (i.e. one side is [low] while the other side is ‘not low’). The nature of phonological features is important because it directly affects how phonological activity is defined and accounted for. The pursuit of privative features is not unproblematic. Nevins (2015) discusses the benefits and liabilities of privative features in detail, as we return to in §5 where we develop a novel proposal that derives and formalizes representations similar to the ones in 2d.

The vowel inventory for West Germanic, the precursor to OE (Lass 1994), is presented in 3a. Dresher 2015 and Purnell & Raimy 2015 agree on the contrastive ranking in 3b, the corresponding contrastive hierarchy in 3c, and the contrastive segments in 3d.

(3) West Germanic

a. i i: u u:

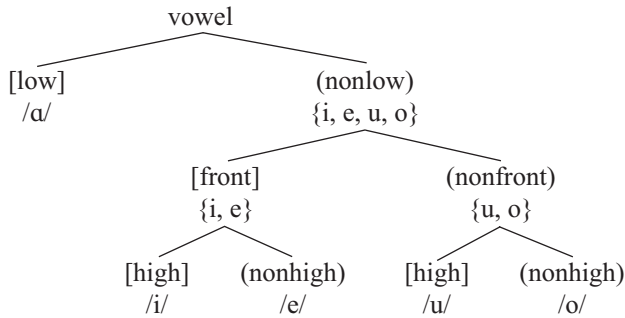
 e e: o o:

 ɑ ɑ:

b. contrastive ranking

 [low] > [back] > [high]

c. contrastive hierarchy



d. contrastive segments

/a/	/i/	/e/	/u/	/o/
[low]	(nonlow)	(nonlow)	(nonlow)	(nonlow)
	[front]	[front]	(nonfront)	(nonfront)
	[high]	(nonhigh)	[high]	(nonhigh)

We do not include vowel length in 3 because it is generally symmetrical in early English and does not affect our arguments. Moreover, we represent segmental length as a structural and nonfeatural aspect of representations (Odden 2011, Ringen & Vago 2011).¹

Additionally, Dresher 2015 and Purnell & Raimy 2015 agree on the representations in 3 based on patterns of phonological activity in West Germanic. First, /a/ is recognized as lacking contrastiveness in backness and thus is phonetically variable along this parameter (Lass 1994:28, n. 9). This justifies [low] as the first feature to separate /a/ from the remaining vowels because it is underspecified for backness and height. Second, the marked aspect of backness is recognized as [front] due to /i/ and /e/ palatalizing consonants (Lass 1994:53–59). Third, [high] is invoked to provide height contrasts. Purnell and Raimy (2015:537–39) discuss how this contrastive ranking/hierarchy and consequent contrastive segments capture Lass’s (1994) main observations about West Germanic phonology.

Although both Dresher 2015 and Purnell & Raimy 2015 are fully committed to the MCS and SDA, they disagree on the model of distinctive features and the contrastive ranking for OE. Both begin with the inventory of vowels in OE (Lass 1994:61) in 4.

(4) OE vowels

i	y	u
e	ø	o
æ		ɑ

The OE vowel inventory reflects the addition of three new vowels (/æ, y, ø/) to the West Germanic system. These vowels require at least one more feature. An additional issue is whether the vowels necessitate a change to the contrastive ranking.

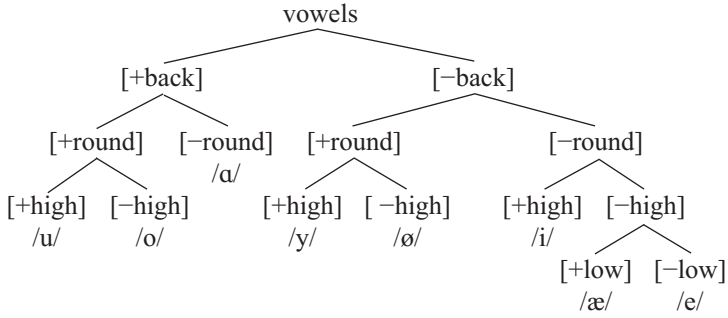
Dresher (2015:520) proposes the contrastive ranking in 5a, the contrastive hierarchy in 5b, and the contrastive feature specification in 5c; he assumes binary features. One upshot of binarity is that while vowels may be underspecified for some features, when a vowel is specified it is either positive or negative: /a/ is specified for only two features, [+back, –round], and /e/ is specified for four features, [–back, –round, –high, –low]. Binary features provide both sides of a contrast with featural substance, potentially allow-

¹ On length in a contrastivist approach, see Spahr 2016 for how nonsegmental features (length and pitch) are encoded in their own distinct contrastive rankings/hierarchies.

ing either side to be phonologically active. A downside of binarity is that the asymmetry in phonological activity arising from only one of the values (either + or -) must be encoded in some fashion other than by the representations themselves (see Chomsky & Halle 1968:Ch. 9, Kean 1975, and Calabrese 2005, 2009). That is, if [-back] triggers an operation or leads to some effect, then the specification of that valence or value (+ or -) must involve another grammatical statement or piece of machinery.

(5) Dresher’s 2015 analysis of OE vowel features

- a. [back] > [round] > [high] > [low]
- b.



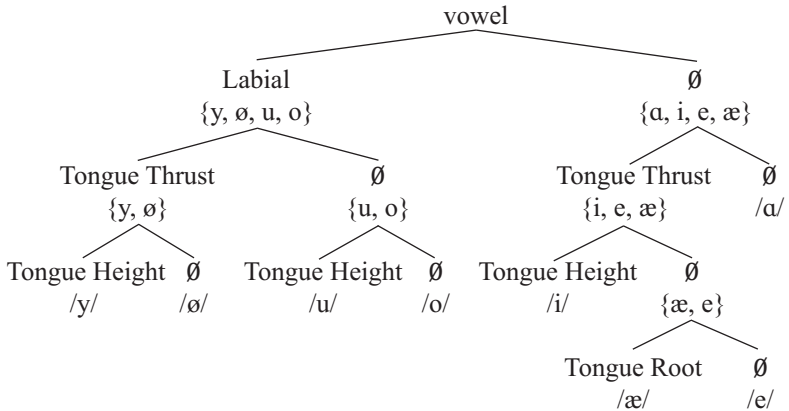
c. contrastive segments

/i/	/e/	/y/	/ø/
[-back]	[-back]	[-back]	[-back]
[-round]	[-round]	[+round]	[+round]
[+high]	[-high]	[+high]	[-high]
	[-low]		
/u/	/o/	/æ/	/a/
[+back]	[+back]	[-back]	[+back]
[+round]	[+round]	[-round]	[-round]
[+high]	[-high]	[-high]	
		[+low]	

Using agnostic featural terms like backness facilitates comparison of Dresher’s proposal to Purnell and Raimy’s, which follows the distinctive feature system proposed by Avery and Idsardi (2001). Purnell and Raimy’s contrastive ranking is given in 6a, while 6b,c present the resulting contrastive hierarchy and contrastive segments. We discuss Avery and Idsardi’s proposals in more detail in §5, but for comparison of Dresher’s proposal in 5 with that in 6, Labial = [+round], Tongue Thrust = [-back], Tongue Height = [+high], and Tongue Root = [+low].

(6) Purnell and Raimy’s (2015:539) analysis of OE vowel features

- a. [round] > [back] > [high] > [low]
- b.



c. Contrastive segments

/i/	/e/	/y/	/ø/
vowel	vowel	vowel	vowel
Tongue Thrust	Tongue Thrust	Labial	Labial
Tongue Height		Tongue Thrust	Tongue Thrust
		Tongue Height	
/u/	/o/	/æ/	/ɑ/
vowel	vowel	vowel	vowel
Labial	Labial	Tongue Thrust	
Tongue Height		Tongue Root	

The main difference between the two analyses, restated in 7 along with the starting West Germanic contrastive ranking, is the order of ‘roundness’ ([±round] for Dresher and Labial for Purnell and Raimy). Purnell and Raimy rank Labial first in the hierarchy, while Dresher ranks it second. Both analyses represent change by demoting [low] from first to last in the ranking.

(7) Change in contrastive rankings

- a. West Germanic (= 3b): [low] > [back] > [high]
- b. OE, Dresher (= 5a): [back] > **[round]** > [high] > [low]
- c. OE, P&R (= 6a): **[round]** > [back] > [high] > [low]

Although both proposals produce the vowel inventory for OE, neither Dresher nor Purnell and Raimy provide in-depth discussion or defense of their proposals. These differences raise the question of which should be preferred and whether either is the best analysis possible. We address these questions in the next section.

3. METRICS FOR CHANGE IN CONTRASTIVE HIERARCHIES. Oxford (2015:317) provides multiple metrics for deciding among possible MCS- and SDA-based analyses of diachronic change. We draw crucially on the CONTRAST SHIFT hypothesis (‘Contrastive hierarchies can change over time’) and the SISTERHOOD MERGER hypothesis (‘Structural mergers apply to “contrastive sisters”’). Both hypotheses place strong constraints on the relationship between diachronic change and contrastive rankings and hierarchies. Neither Dresher’s nor Purnell and Raimy’s proposal fares well against these constraints.

Contrast shift indicates that diachronic change can alter a contrastive ranking by changing the order of features or changing the features themselves. Both reordering distinctive features and adding or subtracting a feature from the hierarchy change the inventory of possible contrasts, as expressed by the emergence of a new contrastive ranking. Small changes to a contrastive ranking might lead to gradual and common diachronic change; large or radical changes to a contrastive ranking would be associated with rare diachronic changes or radical restructurings. To measure a change between contrastive rankings, we propose using JACCARD DISTANCE as a way of measuring the difference between two contrastive rankings. Jaccard distance is a standard method in mathematics to measure the difference between two sets (Matthe et al. 2006, Meyer & Hornik 2009) and can quantify the distance between two contrastive rankings. We hypothesize that diachronic change is more common when the starting and ending contrastive ranking are closer in Jaccard distance and less common when the Jaccard distance is larger. A contrastive ranking can be conceived of as a set of dominance relations where dominance in a ranking is transitive and a-local (akin to c-command). Thus, the contrastive ranking for West Germanic can be represented by the set of dominance relations in 8a, where a dominance relation is a pairing of two features in the contrastive ranking in which the first one dominates the second. Since [low] is the first feature in the West Germanic contrastive ranking in 8a, it dominates the other two features: [low] > [back], and [low] > [high]. The feature [back] is next in the hierarchy and dominates

[high] ([back] > [high]). Finally, [high] is at the bottom so it does not dominate any other feature. In 8a–g we present the dominance sets for relevant contrastive rankings, and in 8h their Jaccard distances from West Germanic.

- (8) Contrast shift in Jaccard distance
- a. West Germanic: [low] > [back] > [high]
{low>back, low>high, back>high}
 - b. Reverse two features: [back] > [low] > [high]
{back>low, back>high, low>high}
 - c. Add a feature to bottom of ranking: [low] > [back] > [high] > [round]
{low>back, low>high, low>round, back>high, back>round, high>round}
 - d. Add a feature to top of ranking: [round] > [low] > [back] > [high]
{round>low, round>back, round>high, low>back, low>high, back>high}
 - e. Add and reverse: [back] > [low] > [high] > [round]
{back>low, back>high, back>round, low>high, low>round, high>round}
 - f. Drescher 2015: [back] > [round] > [high] > [low]
{back>round, back>high, round>high, round>low, high>low}
 - g. Purnell & Raimy 2015: [round] > [back] > [high] > [low]
{round>back, round>high, round>low, back>high, back>low, high>low}
 - h. Jaccard distance from West Germanic

8b ‘reverse two features’	Jaccard distance = 0.500
8c ‘add feature at bottom’	Jaccard distance = 0.500
8d ‘add feature at top’	Jaccard distance = 0.500
8e ‘add and reverse’	Jaccard distance = 0.286
8f Drescher	Jaccard distance = 0.125
8g Purnell & Raimy	Jaccard distance = 0.125

The Jaccard distances in 8h indicate distance between two sets with a value between 1 and 0, with a higher value indicating ‘closer’ (1 is identical) and a smaller value ‘farther’ (0 is completely different). This quantification allows two themes to emerge. First, by considering the Jaccard distance values for 8b,c,d, the switching of two features or the addition of a feature in any position to the hierarchy moves the contrastive hierarchy the same distance away from West Germanic (Jaccard distance = 0.5). This result indicates that principles other than contrast shift are needed to arbitrate among these types of minimal changes in contrastive rankings.

Combining the two separate changes in 8b and 8c as in 8e, where [round] is added to the bottom and [low] and [back] are reversed, produces a ranking still farther away from West Germanic at a Jaccard distance of 0.286, a welcome result. This quantifies the intuition that two changes (i.e. an add and a switch) should produce a ranking farther away than just one change. Importantly, the proposals by Drescher (2015) and Purnell and Raimy (2015), with many additional rerankings, produce rankings that are even farther away from West Germanic, at a Jaccard distance of 0.125 for both. This strongly suggests that we need to consider other possible hierarchies.

Oxford’s second hypothesis, sisterhood merger, provides an additional metric, proposing that segments that merge diachronically must be sisters in the contrastive hierarchy at the time of merger. Where a feature occurs in a contrastive ranking leads to different predictions about diachronic mergers. A feature at the bottom of a hierarchy will undergo mergers on a pairwise basis, so that a feature is gradually lost as pairs of contrasting segments merge. A feature higher up in a contrastive hierarchy will merge

sets of phonemes more broadly. Thus, multiple contrasting pairs of segments will be merged at the same time when the feature is higher in the ranking/hierarchy.

OE obliges us to use sisterhood merger as a metric due to the way that front round vowels created by i-umlaut were later lost through merger. The two front rounded vowels, /y/ and /ø/, disappeared from OE at different times and in different ways in different dialects. Lass (1994:66) indicates three stages of OE, displayed in 9, from an earlier stage on the left to a later stage on the right.

(9) Stages of OE vowels (Lass 1994:66)

pre i-umlaut		post i-umlaut			later West Saxon dialect		
i	u	i	y	u	i	y	u
e	o	e	ø	o	e		o
æ	ɑ	æ		ɑ	æ		ɑ

The mid front rounded vowel /ø/ was lost before the high front rounded vowel /y/. This provides a crucial test for determining where [round] should be added to the contrastive ranking/hierarchy in OE.

Both Dresher and Purnell and Raimy add [round] rather high in the contrastive ranking for OE but in different places, which makes distinct predictions about how contrasts based on roundness will be lost and, consequently, how /y/ and /ø/ merge with other vowels. The basic predictions of Dresher's proposed contrastive ranking for OE are shown in 10 if [round] is lost in OE (based on the contrastive segments representations).

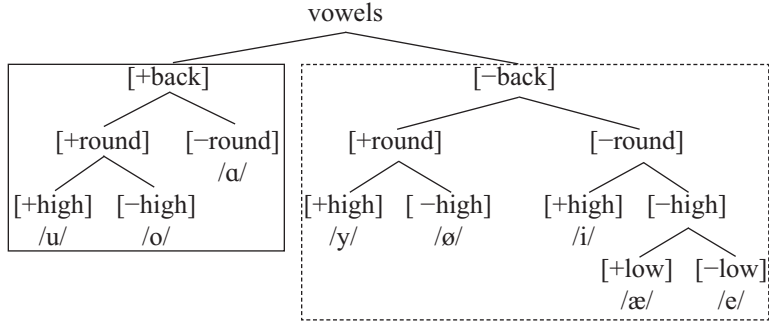
(10) Dresher's predictions on mergers from loss of 'roundness' ([±round])

/i/	/e/	/y/	/ø/	MERGERS
-back	-back	-back	-back	i ~ y
-round	-round	+round	+round	e ~ ø?
+high	-high	+high	-high	æ ~ ø?
	-low		(?low)	u ~ ɑ?
/u/	/o/	/æ/	/ɑ/	
+back	+back	-back	+back	o ~ ɑ?
+round	+round	-round	-round	
+high	-high	-high	(?high)	
		+low		

We indicate loss of 'roundness' by crossing out [±round] and then consider which segments have the same feature specifications. Dresher's analysis clearly predicts that /i ~ y/ should merge (both vowels end up being [-back, +high]), but there are questions about what other related mergers might occur. The resulting specification of /ø/ as [-back, -high], where it is underspecified for [±low], demonstrates this problem. Should it remain distinct from /e/, [-back, -high, -low]? Should it remain distinct from /æ/, [-back, -high, +low]? Should /ø/ stay distinct from both /e/ and /æ/? Similar issues arise with the relationships /ɑ/ (unspecified for [±high]) has with /u/ and /o/.

The problem of predicting which phonemes should merge with the loss of [±round] in Dresher's analysis reflects an interaction between the use of binary distinctive features and the structure of contrasts seen in the hierarchy in 5b. The structures beneath the [±round] contrast in 5b are not symmetrical, which, when combined with binary features, raises the question of how the contrast should change based on the loss of [±round]. In 11 we present a modified 5b.

(11) Contrastive hierarchy from Dresher 2015



Two boxes in 11 show the scope of contrasts affected by removing $[\pm\text{round}]$. The solid box on the left shows that the contrasts under $[\pm\text{round}]$ in the $[\text{+back}]$ limb are asymmetrical because $[\pm\text{high}]$ is contrastive only for the $[\text{+round}]$ vowels, /u/ and /o/. The dotted box on the right identifies an analogous situation in the $[\text{-back}]$ limb. That $[\pm\text{low}]$ is contrastive under the $[\text{-round}, \text{-high}]$ set of vowels /æ/ and /e/ as opposed to the $[\text{+round}, \text{-high}]$ vowel /ø/ causes the problem. Visual inspection of the contrastive hierarchy shows that it is not clear which vowels should merge with each other upon loss of the feature $[\pm\text{round}]$ in Dresher’s proposal.

We conclude that Dresher’s particular contrastive ranking/hierarchy for OE violates the requirements of sisterhood merger. We say ‘violates’ because the actual representations posited do not make clear predictions about what vowels may merge; this prevents a proper evaluation of the sisterhood merger hypothesis. This fragility of predictions about mergers based on binary features and the structure of the hierarchy can be considered an argument for privative features; the underlying ternary nature of binary features $[\text{+}, \text{-}, \emptyset]$ (Stanley 1967) is the ultimate source of this problem. A more optimistic interpretation of this situation is that the lack of symmetry beneath the roundness feature in the proposed hierarchy suggests that $[\pm\text{round}]$ should be a stable feature, thus preventing mergers from happening. Unfortunately, the OE facts do not accommodate this interpretation.

Purnell and Raimy’s proposal places $[\text{round}]$ at the top of the ranking, predicting that all contrastively marked $[\text{round}]$ vowels should merge at once. In 12 we present the basic predictions on mergers that would occur if $[\text{round}]$ was lost in OE, according to Purnell and Raimy’s contrastive segments representations.

(12) Purnell and Raimy’s predictions on mergers from loss of ‘roundness’ (Labial)

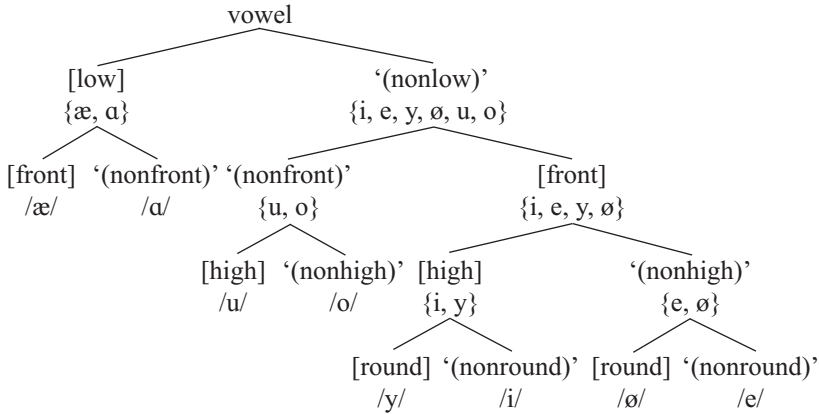
	/i/	/e/	/y/	/ø/	MERGERS
vowel	vowel	vowel	vowel	vowel	i ~ y
Tongue Thrust	Tongue Thrust	Labial	Labial	Labial	e ~ ø
Tongue Height			Tongue Thrust	Tongue Thrust	o ~ a
			Tongue Height		
	/u/	/o/	/æ/	/ɑ/	
vowel	vowel	vowel	vowel		
Labial	Labial		Tongue Thrust		
Tongue Height			Tongue Root		

With 12, we predict that /i ~ y/ (now both specified as vowel, Tongue Thrust, and Tongue Height), /e ~ ø/ (vowel, Tongue Thrust), and /o ~ a/ (simply an unmarked vowel) should all merge pairwise at the same time when ‘roundness’ is lost. As with Dresher’s proposals, Purnell and Raimy’s analysis does not match the diachronic facts of the OE loss of the front round vowels.

To complete this section, we consider the predictions made by the hypothetical contrastive ranking proposed as 8c, which just adds [round] at the bottom of the contrastive ranking for West Germanic. In 13, we present the contrastive hierarchy, the contrastive segments, and the merger predictions.

(13) Hypothetical contrastive ranking from 8c

a. contrastive hierarchy



b. contrastive segments

/i/	/e/	/y/	/ø/	MERGERS
vowel	vowel	vowel	vowel	i ~ y
(nonlow)	(nonlow)	(nonlow)	(nonlow)	e ~ ø
[front]	[front]	[front]	[front]	
[high]	(nonhigh)	[high]	(nonhigh)	
(nonround)	(nonround)	[round]	[round]	
/u/	/o/	/æ/	/a/	
vowel	vowel	vowel	vowel	
(nonlow)	(nonlow)	[low]	[low]	
(nonfront)	(nonfront)	[front]	(nonfront)	
[high]	(nonhigh)			

We show in 13 that this contrastive ranking actually matches the diachronic facts of the loss of the front round vowels through merger. The contrastive hierarchy in 13a has the vowels /i/ and /y/ as sisters along with /e/ and /ø/ as sisters, following Oxford's sisterhood merger hypothesis. Additionally, when we consider the contrastive segments in 13b, we see that losing [round] correctly merges /i/ with /y/ and /e/ with /ø/, again following the sisterhood merger hypothesis. Note that we have returned to the agnostic features used earlier in this report in order to delay the substantive discussion of privative vs. binary features until §5. Oxford's (2015) metrics on diachronic change can identify a contrastive ranking (8c) to prefer over the feature rankings proposed by Dresler (8f) and Purnell and Raimy (8g).

In sum, Oxford 2015 provides clear and explicit metrics for diachronic change: sisterhood merger and contrast shift. We used Jaccard distance to quantify contrast shift, which allowed us to identify contrastive rankings that were minimally different from West Germanic. Both Dresler's and Purnell and Raimy's proposals were more different than the minimal change. Sisterhood merger also provides arguments against both proposals, since neither made correct predictions about how the front round vowels merged and were lost in OE. The minimally changed contrastive ranking in 8c and 13 does make the correct predictions, though. We now turn to diachronic facts about OE phonology to further argue in favor of this minimally changed contrastive ranking.

4. OLD ENGLISH DIACHRONY AND PHONOLOGICAL ACTIVITY. We use Lass's (1994) sketch of developments from West Germanic to OE to further explore whether the contrastive ranking in 13 is adequate to explain the addition of new vowels, the later loss of these added vowels, and aspects of the phonology of OE. We give an overview of the changes in the vowel inventory from West Germanic to OE in 14.

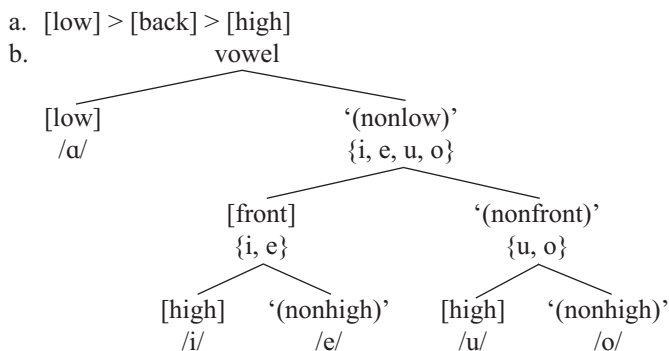
(14) Diachrony of OE vowels

a. West Germanic	b. 'Anglo-Frisian'
i i:	u u:
e e:	o o:
ɑ ɑ:	æ æ:
	ɑ:
	[NB: WGmc ɑ ɑ: > A-F æ æ:; WGmc əj > A-F ə:]
c. Pre i-umlaut OE	d. Post i-umlaut OE
i i:	u u:
e e:	o o:
æ æ:	ɑ ɑ:
	i i:
	y y:
	u u:
	e e:
	(ø ø):
	o o:
	æ æ:
	ɑ ɑ:

West Germanic (14a) had a five-vowel system with a length distinction. Front vowels phonetically contrast with back rounded vowels in high and mid positions, with a single low back unrounded vowel. The first change is BRIGHTENING of /ɑ/ to /æ/. This creates 14b, Lass's 'Anglo-Frisian' (1994:44), and we follow him in treating this as a hypothetical stage rather than a genetic group, avoiding claims about West Germanic subgrouping. Another change is the monophthongization of /ai/ to /a:/, creating a backness contrast for low vowels (at least for the long series). Generalizing the backness contrast to long and short brings us to 14c, the inventory of OE prior to the presence of i-umlaut as a phonological process. Finally, i-umlaut creates two new front rounded vowels, high and mid (/y, ø/), to produce 14d in OE. Mid /ø/ is given in parentheses in 14d in order to indicate that it is lost before /y/, as discussed earlier.

In §3, we concluded that the contrastive hierarchy for OE is 13, but we did not demonstrate how this hierarchy and implicit ranking changed from West Germanic (3) to OE. We repeat the West Germanic contrastive ranking and hierarchy in 15 to facilitate discussion of the diachronic changes, again relying heavily on Lass 1994.

(15) West Germanic



The contrastive hierarchy in 15b starts with 'vowel', representing the contrastive cut separating vowels from nonvowels. The initial contrast between vowels and consonants may be a universal initial contrast. The rationale for West Germanic (15) is that the first contrast is based on [low] in order to separate /a/ out from the other vowels. /a/ is not contrastively paired with another vowel for backness. The lack of backness marking for this vowel allows great phonetic latitude in its implementation. From this point on we use the

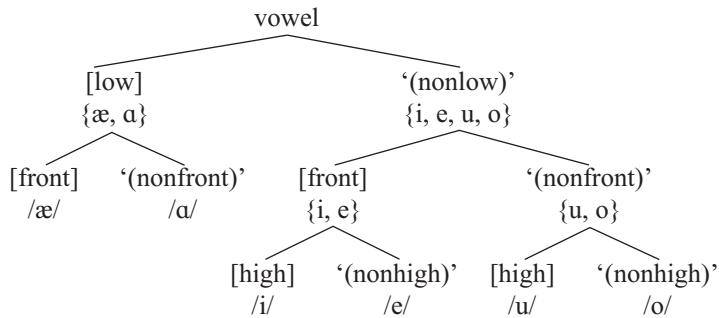
term ‘mark’ (following Dresher et al. 1994), where ‘marking’ indicates that a distinctive feature is memorized to create a contrast. Marking a contrast builds phonological structure and provides some, but not all, phonetic substance to the segment. Consequently, a more ‘marked’ segment has more phonetic substance and less variation in implementation. Relatively ‘unmarked’ segments will not have as much phonetic substance, so they can be more variable in implementation. This interpretation of ‘markedness’ is remarkably narrow (see Rice 1999a,b on markedness in phonology). In order to establish a parallel backness contrast in the remaining vowels, [front] is the next feature in the ranking. Marking front vowels rather than back ones is justified by palatalization facts in West Germanic (Lass 1994:53–55) and will be useful when we discuss i-umlaut in OE later. Finally, [high] is used to mark the height distinction between the mid and high vowels, and we mark high vowels in line with common intuitions that mid vowels are unmarked for both [high] and [low].

The changes that create the ‘Anglo-Frisian’ inventory in 14b follow both the letter and the spirit of Oxford’s views on diachronic change in contrastive hierarchies. The main change between 14a and 14b is the ‘brightening’ of /a/ via phonetic fronting (Lass 1994:43). In the West Germanic hierarchy, /a/ is not specified or contrastive for ‘backness’. Due to this, /a/ is free to move forward in phonetic space without changing any aspect of the contrastive hierarchy in 15a. It is the later monophthongization of /ai/ to /a:/ (Lass 1994:39–40) that requires changing the hierarchy through the addition of a [front] contrast under the branch marked [low]. This produces the hierarchy in 16b.

(16) ‘Anglo-Frisian’ ~ Pre i-umlaut OE

a. [low] > [back] > [high]

b.



The contrastive ranking does not need to change to incorporate the new backness contrast in low vowels (NB: 16a is identical to 15a). We interpret the addition of contrasts and phonemes from 15b to 16b as the inverse of Oxford’s sisterhood merger hypothesis. Newly contrastive phonemes will appear in the tree in a ‘sisterhood’ relationship with the other newly contrastive segment. The hierarchy in 16b represents this with the addition of /æ/ marked with [front] as a sister of /a/ under [low].

This expansion of the sisterhood merger hypothesis is a natural interpretation of Oxford 2015. If we further assume that two of Oxford’s hypotheses—contrast shift and SEGMENTAL REANALYSIS (‘A segment may be reanalyzed as having a different contrastive status’; Oxford 2015:317)—are conservative, the easiest way to add a segment to an inventory is to create a new sister at a terminal node using the distinctive feature at the relevant part of the hierarchy (e.g. [front] is under [low] in 16b). One can identify which distinctive feature is ‘relevant’ by considering the terminal node, which feature it is associated to, and the contrastive hierarchy that provides the daughter of the feature associated with the node. Here, the feature associated with the terminal node (in 16b) is

[low] and its daughter is [front] (an implementation of [back]). This addition requires no change to the contrastive ranking.

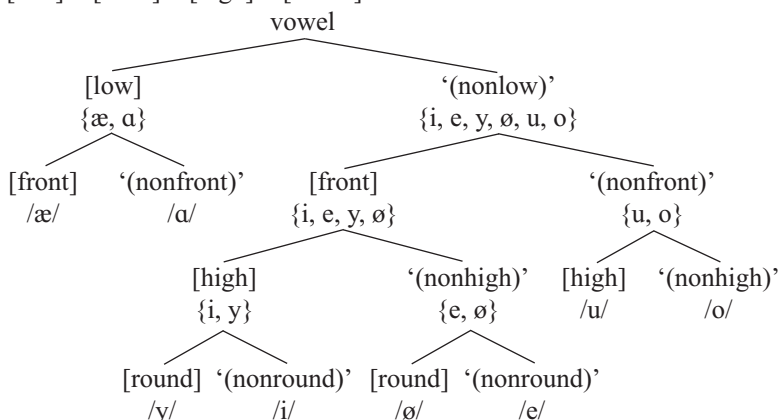
The contrastive hierarchy in 16 is labeled ‘Pre i-umlaut OE’ because nothing about the hierarchy or distinctive features needs to change to account for lengthening of the low back vowel in 14c. The only modification is the introduction of short /a/, making the long and short vowel inventories parallel.

The final change from 14c to 14d is the addition of contrastive front rounded vowels (/y, ø/). We add [round] to the bottom of the hierarchy in 16b following the results of all metrics discussed in §3 to produce the ranking and hierarchy in 17.

(17) Old English

a. [low] > [back] > [high] > [round]

b.



The contrastive ranking and hierarchy we concluded was the proper one in §3 is given in 17. The diachronic changes from West Germanic to OE are clearly minimal. No change to the contrastive ranking of West Germanic is necessary except for the addition of [round] at the bottom in order to produce the vowel inventory of OE. The process we propose can be stated thusly: the front round vowels begin as ‘allophones’ at the end of the phonology, by way of dimensions being completed and by some fill rules and enhancement processes taking place. At that stage, /u/ and /o/ will be specified with [round], and the /y/ will have the [front] added by these latter rules that are necessary to account for the process, but unnecessary and even undesirable at earlier phonological stages.

While these diachronic changes help us understand how the OE inventory arose, phonological rules provide direct evidence for the actual features used to mark the contrasts. Phonological activity not only is helpful in determining the contrastive hierarchy in a language, but it also provides evidence on the privative or binary distinctive features. Four phonological processes in OE provide evidence about the nature of distinctive features: i-UMLAUT, RETRACTION, a-RESTORATION, and BACK MUTATION.

We begin with OE i-umlaut, which provides evidence that the front vowels must be marked so they can act as phonological triggers. Examples showing diachronic change in protoforms to attested words in OE due to i-umlaut are given in 18 (from Hogg 1992). Phonological representations are presented between slashes, and orthographic forms are in italics.

(18) i-umlaut in OE (Hogg 1992:123–24)

- a. */trummjan/ > /trymman/ = *trymman*
‘strengthen’
- b. */gaduling/ > */gadyling/ > */gædyling/ > /gædøling/ = *gædøling*
‘companion’

- c. *brūdi- > brȳd ‘bride’
 d. *ovusti > œfest ‘haste’

These examples from Hogg show that there was a process active in OE where a high front segment (either /j/ or /i/ in these examples) caused preceding back rounded vowels to become front rounded vowels. For *trymman* ‘strengthen’ (18a), the older form /trummjan/ has the /u/ in the first syllable fronted to /y/ and a loss of /j/, producing the corresponding OE form /trymman/. In 18b we give an example where i-umlaut first fronts the word-medial /u/ to /y/, and then the low back vowel /a/ fronts to /æ/. The word-medial /y/ is then later reduced to schwa, creating the attested OE word *gædeling*. Similarly, 18c presents another example of /u/ changing to /y/, with the additional aspects of i-umlaut affecting long vowels and another example of the final opaque grammaticalization of the front rounded vowels in OE. Finally, 18d presents an example of the mid back rounded vowel /o/ undergoing i-umlaut, again with a final opaque grammaticalized OE form.

The changes that i-umlaut caused in OE, creating new front rounded vowels, allow insight into the phonological representation (and activity) of the pre i-umlaut OE vowel inventory in 14c. The presence of /i/ or /j/ causes the fronting of back rounded vowels, which means that they must be phonologically specified as [front] in order to be the trigger and serve as the source of a new [front] specification on the back vowels. West Germanic already marked front vowels, so this is another case where OE did not change. Consequently, adding [front] to back rounded vowels like /u/ and /o/ will create new front rounded vowels, /y/ and /œ/, respectively. See Purnell & Raimy 2015:540–42 for discussion of the gradual grammaticalization of this process. To sharpen this point, 19 presents i-umlaut in a semiformal SPE-type rewrite rule (Chomsky & Halle 1968; this rule is based on Hogg 1992:122).

(19) i-umlaut: Back vowels front before /i/ or /j/ in the next syllable.

$$\emptyset \rightarrow [\text{front}] / \underset{\text{[sonorant]}}{\text{C}_0} \left[\begin{array}{c} \text{sonorant} \\ \text{front} \\ \text{high} \end{array} \right]$$

The SPE rewrite rule provides a convenient way to call attention to the important aspects of the representations for this process. Such rules have the format of $A \rightarrow B / C_D$, where we interpret A as the changeable aspect of the target, B as the new, changed aspect, and C_D as the entire triggering environment except for the changed aspect of the target (i.e. A). In 19 the feature [front] is added to a vowel ([sonorant]) prior to a high front approximant (/i/ or /j/). Thus, according to Lass (1994:60), i-umlaut requires both /i/ and /j/ to be specified for [front] and [high] in order to condition the fronting effects. There are additional raising effects on mid and low front vowels that we leave aside here.² The change specification of the rule must contain [front] because it is only the addition of this feature that will cause the pairwise alternation of /u ~ y/, /o ~ œ/, and /a ~ æ/. Finally, the triggering environment must also contain [front] and [high] as a specification in order to identify that only /i/ and /j/ (and not /e/ or /æ/) trigger this process. It should be easy to see how to convert this logic to constraint-based and other approaches to alternations in phonology.

Although we are discarding Dresher’s (2015) proposed contrastive ranking, we follow him closely for the understanding of retraction, a-restoration, and back mutation rules in OE. Both retraction and a-restoration back /æ/. The former, again, ‘applies before *w* and back *l* (*l* that is followed by a consonant or a back vowel)’ (Dresher 2015:

² See Hogg 1992:122 for discussion of whether multiple rules are needed for these other raising effects.

508), while α -restoration backs /æ/ before a single consonant followed by a back vowel. In the same environment, back mutation adds a ‘schwa-like’ vowel to front vowels. One important thing to understand about these rules in OE is that they did not all occur at the same point in time (Dresher 2015). In 20 the behavior of retraction and α -restoration is shown.

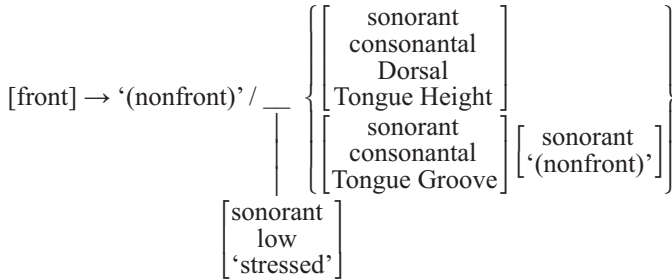
(20) OE grammar with retraction and α -restoration (Dresher 2015:509)

UNDERLYING	/æld/	/hwæł+as/	/fæt+um/	/hæfuc/
RETRACTION	ald	hwalas	—	—
α -RESTORATION	—	—	fatum	hafuc
SURFACE	[ald]	[hwalas]	[fatum]	[hafuc]
GLOSS	‘old’	‘whales.NOM.PL’	‘vats.DAT.PL’	‘hawk.NOM.SG’

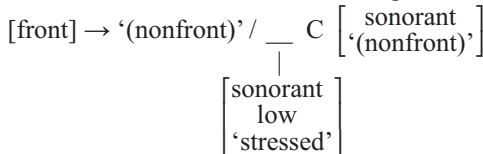
The effect of both retraction and α -restoration is to change a stressed low front /æ/ to a low back /a/ when it is followed by a ‘(nonfront)’ segment. Using the logic we applied to understand i-umlaut, the [back] segments must be specified in order to trigger the rule and to explain the effect. In 21, SPE-type rules demonstrate this.

(21) Retraction and α -restoration

a. Retraction: /æ/ backs before /w/ or /l/ + back vowel.



b. α -restoration: /æ/ backs before a single consonant + back vowel.



The main question raised by the rules in 21 is what to make of the status of ‘(nonfront)’ as a distinctive feature. This is a complicated question, and we ask for the reader’s patience until §5 where we discuss this issue in detail.

The last rule in OE that we consider is back mutation. We present a diachronic summary of its effects in 22.

(22) Back mutation (adapted from Dresher 2015:513)

PRE-OE	*stæðul	*stæðul-es
BACK MUTATION	stææðul	*stææðules
V-REDUCTION	—	stææðeles
ORTHOGRAPHY	<steaðul>	<steaðeles>
GLOSS	‘foundation.ACC.SG’	‘foundations.ACC.PL’

As noted above, back mutation causes the insertion of a schwa immediately after a stressed front vowel, including now /æ/ if it is followed by a back vowel in the following syllable. This effect can be seen in the two examples for ‘foundations’. Back mutation agrees with retraction and α -restoration in requiring ‘(nonfront)’, referenced in 23.

- (23) Back mutation: Front vowels develop a schwa-like element before a back vowel.

$$\emptyset \rightarrow [\text{sonorant}] / \left[\begin{array}{c} \text{sonorant} \\ \text{front} \end{array} \right] _ C \left[\begin{array}{c} \text{sonorant} \\ \text{(nonfront)} \end{array} \right]$$

NB: a plain ‘vowel’ specification is an unspecified vowel, which we assume is schwa in OE.

The four rules we focus on—i-umlaut, retraction, α -restoration, and back mutation—highlight the importance of the nature of distinctive features in understanding phonological activity. The rule of i-umlaut is easily understood as a process in OE due to the fact that [front] was marked in West Germanic and it can remain so. A marked feature such as [front] can and should be phonologically active and participate in phonological processes. The other three rules require ‘(nonfront)’ as a distinctive feature in order to formulate their rules.

The question about ‘(non-X)’-type distinctive features in a contrastive ranking brings a decision point. Nevins (2015) argues that this decision point demonstrates the limits of a purely contrast-based distinctive feature system. Binary features are required to capture phonological activity when two values of a single distinctive feature are required in a single language. This is his ‘oops, I needed that’ observation (2015:58–62). From a surface perspective, the phonological activity exhibited by i-umlaut, retraction, α -restoration, and back mutation in OE provide a strong argument for binary features over privative ones. On that view, one can then do a ‘find and replace’ in this section and insert binary features for the agnostic ones we have used. This would confirm that the core contrastive hierarchy of OE has not changed since West Germanic except for the addition of [±round] at the bottom of the ranking.

In response to Nevins 2015, we treat this question about ‘(non-X)’ features in a privative system as a puzzle. We turn to this puzzle in the next section, distinguishing between two types of ‘(non-X)’ features.

5. SUPERORDINATE NULL MARKING WITH PRIVATIVE DISTINCTIVE FEATURES. The idea of phonological activity is intimately related to the question of whether privative distinctive features are viable. As already noted, Oxford argues (2015:311) that ‘[p]rivativity ... makes the model [of phonology and distinctive features—*PR&S*] maximally restrictive, since it predicts that only the marked values of contrastive features will be phonologically active’. Following Avery and Rice (1989), the strong connection between privative features, markedness, and phonological activity is desirable. If markedness is read directly off of the phonological representations, then the phonology is transparent. Binary systems are opaque because markedness cannot be directly read off the representations themselves. The main differences between binary and privative approaches to distinctive features are (i) whether an unmarked feature can be phonologically active or inert, and (ii) whether the unmarked feature is predictable from the representations.

Since at least Chomsky & Halle 1968:Ch. 9, naive binary distinctive feature systems have been problematic in claiming equal status with respect to feature valency, and, consequently, phonological activity is possible for both sides of a contrast. Various solutions have been pursued. One is Calabrese’s (2005, 2009:285–87) proposal that phonological rules are parameterized along the lines of ‘marked’/‘contrastive’/‘all’, with an independent component of grammar doing the bookkeeping of which distinctive features are of which type. In contrast, the traditional privative solution offers only the ‘marked’ option of Calabrese’s triad. A more nuanced version of marking in a priv-

ative system closes this gap between the Calabrese and naive systems. Moreover, a nuanced approach provides a representation-internal method to produce ‘marked features only’ and ‘contrastive features only’ phonological rules. Lastly, segmental representations may change over the course of a derivation, with more features or information added by rule. Consequently, an organic ordering restriction sequences rules that work on ‘marked’ representations prior to ones that work on ‘contrastive’ representations, which, in turn, precede the ‘all features’ representations.

We identify the ‘unmarked’ side of a contrast with the superordinate node or feature in the geometry of the system. We adopt Avery and Idsardi’s (2001) feature system, shown in Figure 1. This combines strict privativity and feature geometry in a unique way. The feature geometry in Fig. 1 is related to the revised articulator model of Halle et al. 2000, from which it receives aspects of the hierarchical feature geometry. This type of geometry, with dependency relations among features and abstract organizational nodes, has been common in generative grammar since Sagey 1986, with Ni Chiosáin & Padgett 1993, Clements & Hume 1995, and Halle et al. 2000 representing different motifs on the general theme.

The privative aspect of Avery and Idsardi is reflected in positing GESTURES instead of features and the introduction of DIMENSIONS as organizational nodes of gestures. These gestures and dimensions are inherently connected, based on the fact that muscles in the human body are organized in a complementary yet antagonistic fashion (Avery & Idsardi 2001:44). A dimension is then an organizational node that dominates two antagonistic gestures. For example, the Tongue Height dimension dominates the gestures [high] and [low]. From a biophysical gestural perspective, tongue height cannot be both [high] and [low] at the same time. Avery and Idsardi introduce the DIMENSION HYPOTHESIS: ‘phonological representations provide specification for the phonetic DIMENSIONS and not for phonetic FEATURES [gestures—*PR&S*]’ (2001:41), where underlying phonological representations consist only of dimensions, a hypothesis we follow. The conversion of dimensions to terminal gestures is referred to as COMPLETION (Avery & Idsardi 2001:46–48), and in modeling speech behavior it is one step in translating phonological representations into phonetic ones. The completion rules for early English vowels are: Tongue Root > [RTR] (i.e. [+low]), Tongue Thrust > [front] (i.e. [–back]), Tongue Height > [high] (i.e. [+high]), and Labial > [round] (i.e. [+round]).

Avery and Idsardi’s is not the only distinctive feature system that is strictly privative in representation. DEPENDENCY PHONOLOGY (Anderson & Jones 1977), PARTICLE PHONOLOGY (Schane 1984), and GOVERNMENT PHONOLOGY (Kaye et al. 1985) all have nonbinary phonological features. These approaches do not, however, adopt feature geometry, which prevents them from taking advantage of the superordinate node marking proposals.

The geometry in Fig. 1 provides the subordinate and superordinate relationships among organizational nodes, articulators, dimensions, and gestures. Following this, we revise the OE contrastive hierarchy in 17b as in 24 below where the ‘(non-X)’ values have been replaced, not with the null symbol, but with the relevant superordinate category. Crucially, the use of the superordinate needs to be unmarked with respect to the subordinate; thus, instead of marking a plain Dorsal articulator for ‘(non-X)’, we mark DorsalTH (unspecified Dorsal with respect to Tongue Height) or Dorsal^{TT} (unspecified Dorsal with respect to Tongue Thrust). At each step in 24 the ‘null’ side of the branch is marked with the superordinate node, as indicated in Fig. 1. Since Tongue Root is the

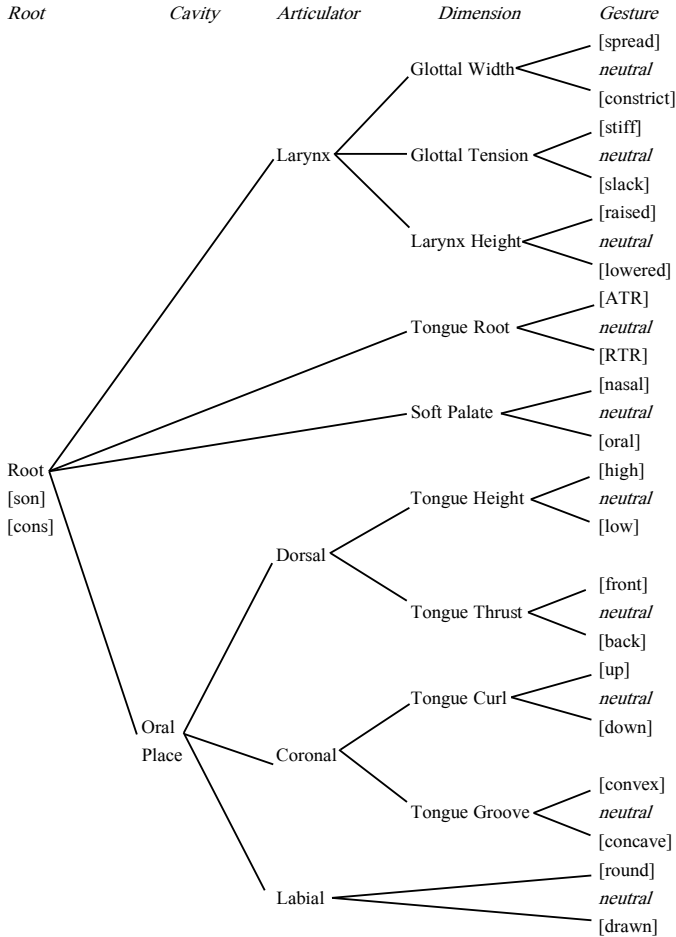
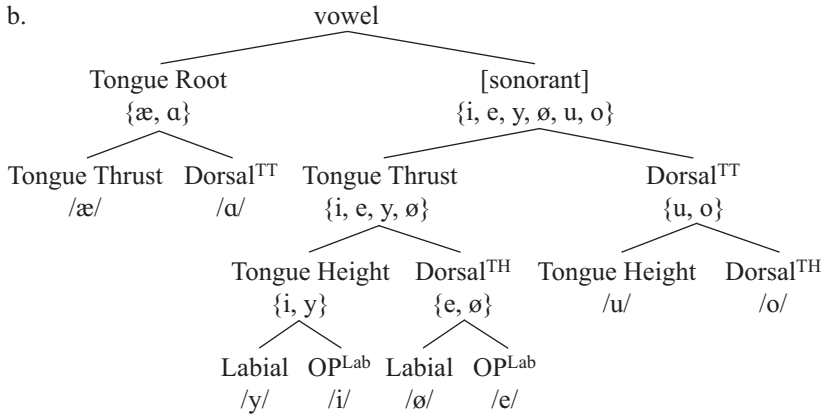


FIGURE 1. Modified feature system from Avery & Idsardi 2001:66 (see also Purnell & Raimy 2015:526).

first feature in the hierarchy in 24, the marked side receives Tongue Root, and the ‘null’ side is marked with [sonorant] since it dominates Tongue Root in the hierarchy (given that vowels are not [consonantal]). The next feature is Tongue Thrust, so the ‘null’ side is marked with Dorsal^{TT}. Tongue Height is next and also marks the ‘null’ side with DorsalTH. Dorsal without either reference to a specific subordinate would be unmarked for both. (We return the status of superordination of a feature in §6.) Labial is the last feature in the hierarchy, and the ‘null’ side is consequently marked with Oral Place. Use of superordinates is not necessarily introducing new machinery, which would make this analysis less simple in the technical sense. Rather, use of superordinates makes explicit an implicit phonological assumption, namely, that when a feature deletes, such as [front], that feature and only that feature is deleted and not all features or some random set of features. Thus, the actual interpretation of null in past phonological analyses has a real semantic interpretation of a feature disappearing and the node above that feature being ‘empty’ for that feature; we represent this phonological emptiness with the superordinate.

(24) OE with superordinate marking

- a. Tongue Root > Tongue Thrust > Tongue Height > Labial
- b.



Although a feature geometry provides informational structure to distinctive features, we agree with Halle (2005) in rejecting the purely structural interpretation of feature geometry. Instead, feature geometry is a calculus for how distinctive features and organizational nodes can be manipulated, as developed in §6. Thus, completion is a natural interpretation of the ‘bottle brush’ organization (Halle 2005). We follow Spahr (2016: 65–66) in seeing distinctive features as organized in a FEATURE CHAIN, which encodes the order of features in the contrastive hierarchy as the order of features in the chain. In 25 we present the updated OE contrastive segments following these assumptions.

(25) Featural representation for OE vowels

/i/	/y/	/e/	/ø/
1. vowel	1. vowel	1. vowel	1. vowel
2. [sonorant]	2. [sonorant]	2. [sonorant]	2. [sonorant]
3. Tongue Thrust	3. Tongue Thrust	3. Tongue Thrust	3. Tongue Thrust
4. Tongue Height	4. Tongue Height	4. Dorsal TH	4. Dorsal TH
5. Oral Place ^{Labial}	5. Labial	5. Oral Place ^{Labial}	5. Labial
/u/	/o/	/æ/	/ɑ/
1. vowel	1. vowel	1. vowel	1. vowel
2. [sonorant]	2. [sonorant]	2. Tongue Root	2. Tongue Root
3. Dorsal ^{TT}	3. Dorsal ^{TT}	3. Tongue Thrust	3. Dorsal ^{TT}
4. Tongue Height	4. Dorsal TH		

Encoding various aspects of phonology directly, the representations in 25 provide a novel synthesis of proposals on distinctive features. We implement Spahr’s feature-chain proposal by providing row numbers in the representations and leave full discussion of these new types of representation to §6.

Importantly, the representations in 25 provide the resources to account for the four OE rules, given the contrastive segments posited. The process of i-umlaut is a straightforward spreading rule where the Tongue Thrust feature spreads from a following /i/ to the preceding Dorsal-specified vowel, as in 26a. Retraction is also a straightforward linear rule: the structural description of the rule is that a ‘back sonorant’ immediately follows a stressed /æ/, as in 26b. A ‘back sonorant’ can now be specified with the contrastive superordinately marked Dorsal on sonorants (i.e. ‘/w/ and back [Dorsal for us—PR&S] /l/ (l followed by a consonant or back vowel)’; Drescher 2015:508–9). The reanalysis is that the rule deletes Tongue Thrust instead of spreading a [back] feature.

Tongue Thrust is contrastively opposed with Dorsal in vowels, where Dorsal is the supervening node; consequently, a sonorant segment specified as Dorsal without any Tongue Thrust feature can be used as the environment specification for retraction. Again, we use an SPE-esque rule format, and we leave the translation into any chosen phonological approach to the reader.

(26) Updated i-umlaut and retraction rules

a. i-umlaut

$$\text{Dorsal}^{\text{TT}} \rightarrow \text{Tongue Thrust} / \underset{\substack{| \\ [\text{sonorant}]}}{\text{C}_0} \left[\begin{array}{l} \text{sonorant} \\ \text{Tongue Thrust} \\ \text{Tongue Height} \end{array} \right]$$

b. Retraction

$$\text{Tongue Thrust} \rightarrow \text{Dorsal}^{\text{TT}} / \underset{\substack{| \\ \left[\begin{array}{l} \text{sonorant} \\ \text{Tongue Root} \\ \text{'stressed'} \end{array} \right]}}{\text{C}_0}$$

Analyses of α -restoration and back mutation can be developed drawing on Lombardi's proposals (1995a,b) about phonological computation with privative laryngeal features. She proposes that privative systems require two fundamentally different ways of causing phonological activity. Marked features are privative and present in representations. As target and trigger of spreading and insertion processes, these marked features also cause phonological activity. Unmarked features, by contrast, are absent from representations and thus cannot spread since there is no specification to be referenced by rules. Where unmarked features appear to spread (as in breaking and α -restoration), there must be a structural constraint deleting marked features. In other words, alternations that appear to be spreading an unmarked feature are actually the deletion of the marked feature based on a structural environment of some sort (as we posited for retraction in the previous paragraph). Lombardi concludes this from the interaction of privative laryngeal features and the LARYNGEAL CONSTRAINT, which says that laryngeal features are licensed only if they precede a sonorant (Lombardi 1995b: 42). A laryngeal feature that does not immediately precede a sonorant is then deleted. This provides a template for analyses of α -restoration and back mutation, with the main question being what the structural constraint for each rule is and what the response is.

A HEAD-DRIVEN ASYMMETRY analysis (HDA; Dresher & van der Hulst 1998) provides structural configurations to implement a distinctive-feature-licensing effect as opposed to the prosodic structure invoked by Lombardi's work on laryngeal features. HDAs are a head-dependent relationship where dependents can only be equal to or lesser in complexity compared to the head. Marked features are more complex via explicitness than their superordinately marked counterparts. For the rules in question then, Tongue Thrust (which marks [front]) is explicit, and thus the segment is more complex than a comparable segment with Dorsal^{TT} (which is the superordinate unmarked value). For α -restoration and back mutation, the complexity relationship is based on the features in row 3 of the feature chain for each vowel in 25. The domain in which the HDA applies for these two rules is based on each vowel acting as a head and building a domain with a preceding vowel, that is, with this preceding vowel being the dependent vowel. We illustrate the domains built (marked with curly brackets) for each of these rules in 27.

(27) Head-dependent asymmetries in OE

	a. α -restoration	b. back mutation
GLOSS	‘vats.DAT.PL’	‘men.NOM.PL’
UNDERLYING	/fæt+um/	/wer+as/
DOMAIN	...{ TT C Dor }...	...{TT C Dor }...
EQUAL COMPLEXITY	...{ TT C Dor}... = ...{Dor C Dor}...	...TT{ \emptyset C Dor}...
SURFACE	[fatum]	[wæras]

Note: TT = Tongue Thrust. Boldface indicates the head of the domain. Double strikethrough means that the dependent is too complex and must be altered. The alteration for α -restoration is the following line.

For α -restoration (27a), the HDA requires the dependent vowel (on the left) to be less complex than the head. The repair, if this relationship does not hold, is for Tongue Thrust to delete, thereby matching the head vowel in complexity (or lack thereof). Consequently, the Tongue Thrust vowel becomes Dorsal^{TT}. Back mutation (27b) enforces the complexity relation in a different manner by inserting a schwa to remove the offending more complex Tongue Thrust vowel out of the domain in question. Presumably, domains are not rebuilt after this repair. The reanalysis of these two rules as HDA effects demonstrates one way ‘unmarked’ features (e.g. Dorsal) in a privative distinctive feature system can trigger phonological activity. The key to this account is the combination of HDAs providing structures that can evaluate the presence/absence of a marked feature compared to an unmarked feature.

An immediate benefit of this approach is that phonological activity can be understood in a pluralistic manner. Phonological activity is not only the spreading and/or insertion of marked phonological material. Phonological activity should also be understood through Lombardi’s work on licensing laryngeal features. Unmarked features can cause phonological activity through licensing restrictions on marked features. Here we implement this distinction in phonological activity through spreading of marked features for i-umlaut and then deletion of marked features through the failure of licensing structures for retraction, α -restoration, and back mutation. We see this bifurcation of phonological activity based on marked vs. unmarked features paralleled in the discussion of positional-markedness vs. positional-licensing approaches in OPTIMALITY THEORY as discussed by Krämer (2018:45–52), who suggests that this debate is a ‘current issue’ in optimality theory, which needs to be resolved in a parsimonious fashion by eliminating one of the approaches. We suggest that the two approaches (and possibly the variations on them discussed by Krämer) are needed to reflect the privative nature of distinctive features.

To conclude, marking of the unmarked side of contrast with a superordinate category as determined by the feature geometry provides a way for a privative feature system to account for surface phenomena that appear to require binary features. This makes the analysis proposed in §3 clearly preferable. It follows Oxford 2015 more closely than rival analyses while capturing the same processes.

6. FEATURE GEOMETRY, PHONOLOGICAL SUBSTANCE, AND PHONOLOGICAL ACTIVITY.

The segmental representations adopted above have useful characteristics that need to be unpacked. A general strength of these new representations is that many topics of phonological theory—markedness, tiers, feature geometry, and so forth—are encoded directly in representations.

Central to these new types of segments is Spahr's feature-chain proposal (2016). This essentially returns to certain aspects of SPE feature matrices specifically in allowing for ordering and structure within the matrix to encode information. By ordering the features in a feature chain, the information contained in contrastive hierarchies is directly encoded, which answers a metaphysical question about the nature of contrastive hierarchies. From our perspective, the hierarchies exist only in the segmental representations themselves.

Also, directly encoded in these representations are contrasts among segments. Since a feature will be added to a chain only if it is contrastive (as a result of the SDA; Drescher 2009), then whether a segment is contrastive with another segment for some feature can be computed directly off representations. We repeat OE contrastive segment representations from 25 as 28.

(28) Contrastive segments for OE vowels

/i/	/y/	/e/	/ø/
1. vowel	1. vowel	1. vowel	1. vowel
2. [sonorant]	2. [sonorant]	2. [sonorant]	2. [sonorant]
3. Tongue Thrust	3. Tongue Thrust	3. Tongue Thrust	3. Tongue Thrust
4. Tongue Height	4. Tongue Height	4. Dorsal TH	4. Dorsal TH
5. Oral Place ^{Labial}	5. Labial	5. Oral Place ^{Labial}	5. Labial
/u/	/o/	/æ/	/ɑ/
1. vowel	1. vowel	1. vowel	1. vowel
2. [sonorant]	2. [sonorant]	2. Tongue Root	2. Tongue Root
3. Dorsal ^{TT}	3. Dorsal ^{TT}	3. Tongue Thrust	3. Dorsal ^{TT}
4. Tongue Height	4. Dorsal TH		

By considering the numbered tiers (on which more below), one can see how contrast is distributed over different features for different sets of vowels. From tier 2, /æ, ɑ/ are contrastively [low] (marked with Tongue Root) compared to the rest of the vowels marked [sonorant]. From tier 3, all front vowels (/i y e ø æ/, marked with Tongue Thrust) contrast with nonfront vowels (marked with Dorsal^{TT}). Continuing to tier 4, we see that some segments are not contrastive with others: /i, e/ contrast with /y, ø/ based on rounding (tier 5, Labial); /u, o/ are not in a contrastive relationship with these vowels. Similarly, /æ, ɑ/ are not contrastive for [high] (tier 4, Tongue Height) with the other vowels in the inventory. To reiterate, the proposal emphasizes the nature of dimensions as being marked (Tongue Height for /u/), unmarked but latent (DorsalTH for /o/), and unmarked as irrelevant (/ɑ/ is associated with neither Tongue Height nor DorsalTH). As with the contrastive ranking, the contrastive hierarchy is directly encoded in the representations. Both hierarchies provide different perspectives on a phonemic inventory. In the end, all of the information of contrasts resides solely within the segments themselves, following Spahr's feature chains.

The adoption of the feature-chain organization of distinctive features does not conflict with feature geometries such as Clements 1985, Sagey 1986, or others. In fact, feature chains further refine the Halle 2005 'bottlebrush' position, which treats the hierarchical organization of distinctive features as implicational set-based knowledge as opposed to actual constituents. The difference can be seen in the implementation of terminal spreading by Halle, Vaux, and Wolfe (2000:395, 430–32). In essence, feature geometry provides the semantics of where an organizational node activates and manipulates all of the features or terminal elements dominated by it. This applies to the Avery and Idsardi model in Fig. 1, as indicated in 29.

(29) Feature geometry referential semantics

CATEGORY	REFERS TO
x-slot	ALL
Oral Place	Dorsal, Tongue Height, [high], [low], Tongue Thrust, [front], [back] ... , Labial, [round] ... , etc.
Dorsal	Tongue Height, [high], [low], Tongue Thrust, [front], [back]
Tongue Thrust	[front], [back]
[front]	[front]

We show in 29 how general or specific a category is based on feature geometry. Selecting an x-slot causes total copying (or deletion) because it can include all of the subcategories in feature geometry. Oral Place selects only categories of the feature geometry that it dominates, including Dorsal, Coronal, and so forth, but excludes Laryngeal features. Dorsal is a subset of Oral Place, so it selects only its dependent dimensions and gestures: Tongue Thrust, [front], [back], Tongue Height, [high], and [low]. The dimension Tongue Thrust will select only its dependent gestures [high] and [low] but no others. Finally, a single gesture such as [front] selects only itself.

The choice of the word ‘semantics’ above is not accidental. Many aspects of these phonological representations are deeply related to (and possibly derived from) general aspects of the human mind/brain. Our starting point is the position of Murray, Wise, and Graham (2017) that MEMORY and REPRESENTATION are the same thing in the human mind. Their main thrust (2017:6) is that:

we have a much better idea about what memory means. Computer scientists and information theorists have explained that memory means stored information: nothing more and nothing less As behaviorism gave way to a more cognitive approach to psychology in the 1970’s, it became possible to identify different forms of memory in both humans and animals. It has become fashionable to refer to each kind of memory as a system, and much of this book explores what the term *memory system* means. In particular, we ask why memory seems to be organized in systems. Our answer is that new representational systems evolved at certain times and places—in particular ancestral species—and that they augmented existing representations when they did.

We adopt their idea of MEMORY SYSTEMS and treat phonology as one memory system. Consequently, if representation and memory are the same thing, memory can tell us about the nature of representations. Following this, we expect phonological representations to have identifiable aspects of human memory, and they do. Wicklegren (1981:22) states that ‘one of the most important successes of cognitive psychology is that we can confidently assert that this nonassociative theory of LTM [long term memory—PR&S] is false’ and that ‘the most critical defining feature of an associative memory is the capacity for direct access retrieval’. The use of the term *direct access* is the bridge between cognitive psychology and computer science that Murray et al. (2017) recognize to confidently assert that memory = representation. Kohonen (1987:1) writes:

It seems that there exist two common views of associative memory. One of them, popular in computer engineering, refers to the principle of organization and/or management of memories which is also named *content-addressing*, or searching the data on the basis of their contents rather than by their location. The coverage of this book in broad outline coincides with it. The other view is more abstract: memory is thereby understood as a semantic representation of knowledge, usually in terms of *relational structures*.

The core aspects of §5 on numbered tiers in feature chains and superordinate marking result directly from content-addressing and relational structures. The ad hoc numbering and presence of tiers in feature chains is a consequence of the content-addressable nature of phonological representations. Each feature tier can be named/defined/referred to by the marked feature directly through invocation of the feature as a result of the asso-

ciative/content-addressable nature of human memory. Returning to 29, ‘tier 1’ is actually the VOWEL tier, ‘tier 2’ is the TONGUE ROOT tier, ‘tier 3’ is the TONGUE THRUST tier, and ‘tier 4’ is the TONGUE HEIGHT tier. Because of the associationist nature of human memory, the tier-based organization of phonological representations is likely a result of innate general human cognition rather than language-specific.

We go one step beyond using tiers and their ad hoc numbering by marking the superordinate with respect to the subordinate, as mentioned above. Thus, DorsalTH and Dorsal^{TT} are shorthand notations for a superordinate Dorsal that appears on a particular tier in 25 and 28. The use of the superscript is for ease of discourse and representation.

The hierarchical organization of feature geometry can also be related to aspects of human cognition not specific to phonology. Lyons’s (1977:291–304) discussion of hyponymy in lexical semantics provides a way to derive the hierarchical structure of feature geometry and the basis of contrast in phonology. ‘Hyponymy is definable in terms of unilateral implication’ (1977:292), and when this is combined with the grounded nature of an articulatory-based model of distinctive features, the general structure of feature geometry is derived. The basic implicational structure inherent in the gestures of the Avery and Idsardi system in Fig. 1 was presented as if the feature geometry imposed these relations, but in fact, ‘[t]he relation of hyponymy imposes a hierarchical structure upon the vocabulary and upon particular fields within the vocabulary’ (Lyons 1977:295). Furthermore, ‘co-hyponyms of the same superordinate will contrast in sense ... and the nature of contrast can be explicated in terms of a difference in the encapsulated syntagmatic modification of the superordinate’ (Lyons 1977:294), and this seems to express the type of contrast needed to understand phonological representations along the lines in 29. The addition of dimension-level representations from Avery and Idsardi is understood as phonological antonyms where the two hyponymic gestures of a dimension are opposite and incompatible with each other. The similarities in the structures governing distinctive features in phonology and lexemes in semantics are support for the proposals in §5. Phonology and semantics clearly remain different due to the nature of the domain-relevant primitives, phonology being more analytic and lacking other types of relations that semantics requires.

7. CONCLUSION. Our goals have been to (i) present a new analysis of diachronic changes in early English vowels and (ii) develop proposals from Oxford 2015 about how contrastive hierarchies capture diachronic change and the role of privative features. The present analysis improves on proposals by Drescher (2015) and Purnell and Raimy (2015), revealing striking stability over history in the English contrastive hierarchy. We develop more explicit methods to understand Oxford’s proposals on diachronic changes in hierarchies. Sections 4 and 5 defend privative features and offer a new approach to understanding ‘null’ marking of the unmarked side of a contrast. This addresses Nevins’s (2015) concerns about privative-contrast-only approaches. Parallels from semantics help justify aspects of feature geometry in general and particular aspects of Avery & Idsardi 2001. These provide general solutions to the issues discussed and underscore the value of diachronic evidence for formal phonology.

REFERENCES

- ANDERSON, JOHN, and CHARLES JONES. 1977. *Phonological structure and the history of English*. Amsterdam: North-Holland.
- AVERY, PETER, and WILLIAM IDSARDI. 2001. Laryngeal dimensions, completion and enhancement. *Distinctive feature theory*, ed. by T. Alan Hall, 41–70. Berlin: Mouton de Gruyter.

- AVERY, PETER, and KEREN RICE. 1989. Segment structure and coronal underspecification. *Phonology* 6(2).179–200. DOI: 10.1017/S0952675700001007.
- CALABRESE, ANDREA. 2005. *Markedness and economy in a derivational model of phonology*. Berlin: Mouton de Gruyter.
- CALABRESE, ANDREA. 2009. Markedness theory versus phonological idiosyncrasies in a realistic model of language. *Contemporary views on architecture and representation in phonology*, ed. by Eric Raimy and Charles Cairns, 261–304. Cambridge, MA: MIT Press.
- CHOMSKY, NOAM, and MORRIS HALLE. 1968. *The sound pattern of English*. New York: Harper & Row.
- CLEMENTS, G. N. 1985. The geometry of phonological features. *Phonology Yearbook* 2. 225–52. DOI: 10.1017/S0952675700000440.
- CLEMENTS, G. N., and ELIZABETH HUME. 1995. The internal organization of speech sounds. *The handbook of phonological theory*, ed. by John Goldsmith, 245–306. Oxford: Blackwell.
- DRESHER, B. ELAN. 2009. *The contrastive hierarchy in phonology*. Cambridge: Cambridge University Press.
- DRESHER, B. ELAN. 2014. The arch not the stones: Universal feature theory without universal features. *Nordlyd* 41.165–81.
- DRESHER, B. ELAN. 2015. Rule-based generative historical phonology. In Honeybone & Salmons, 501–21. DOI: 10.1093/oxfordhb/9780199232819.013.026.
- DRESHER, B. ELAN. 2018. Contrastive feature hierarchies in Old English diachronic phonology. *Transactions of the Philological Society* 116.1–29. DOI: 10.1111/1467-968X.12105.
- DRESHER, B. ELAN; GLYNN PIGGOTT; and KEREN RICE. 1994. Contrast in phonology: Overview. *Toronto Working Papers in Linguistics* 13.iii–xvii. Online: <https://twpl.library.utoronto.ca/index.php/twpl/article/view/6331>.
- DRESHER, B. ELAN, and HARRY VAN DER HULST. 1998. Head-dependent asymmetries in phonology: Complexity and visibility. *Phonology* 15.317–52. DOI: 10.1017/S0952675799003644.
- HALL, DANIEL CURRIE. 2007. *The role and representation of contrast in phonological theory*. Toronto: University of Toronto dissertation. Online: <https://twpl.library.utoronto.ca/index.php/twpl/article/view/6497>.
- HALLE, MORRIS. 2005. Palatalization/velar softening: What it is and what it tells us about the nature of language. *Linguistic Inquiry* 36.23–41. DOI: 10.1162/0024389052993673.
- HALLE, MORRIS; BERT VAUX; and ANDREW WOLFE. 2000. On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31.387–444. DOI: 10.1162/002438900554398.
- HOGG, RICHARD M. 1992. *A grammar of Old English, vol. 1: Phonology*. Oxford: Blackwell.
- HONEYBONE, PATRICK, and JOSEPH SALMONS (eds.) 2015. *The Oxford handbook of historical phonology*. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199232819.001.0001.
- KAYE, JONATHAN D.; JEAN LOWENSTAMM; and JEAN-ROGER VERGNAUD. 1985. The internal structure of phonological elements: A theory of charm and government. *Phonology Yearbook* 2.305–28. DOI: 10.1017/S0952675700000476.
- KEAN, MARY LOUISE. 1975. *The theory of markedness in generative grammar*. Cambridge, MA: MIT dissertation.
- KOHONEN, TEUVO. 1987. *Content-addressable memories*. 2nd edn. Berlin: Springer.
- KRÄMER, MARTIN. 2018. Current issues and directions in optimality theory: Constraints and their interactions. *The Routledge handbook of phonological theory*, ed. by S. J. Hannahs and Anna R. Bosch, 37–67. Abingdon: Routledge.
- LASS, ROGER. 1994. *Old English: A historical linguistic companion*. Cambridge: Cambridge University Press.
- LOMBARDI, LINDA. 1995a. Laryngeal features and privativity. *The Linguistic Review* 12.35–59. DOI: 10.1515/tlir.1995.12.1.35.
- LOMBARDI, LINDA. 1995b. Laryngeal neutralization and syllable wellformedness. *Natural Language and Linguistic Theory* 13.39–74. DOI: 10.1007/BF00992778.
- LYONS, JOHN. 1977. *Semantics*, vol. 1. Cambridge: Cambridge University Press.
- MATTHE, TOM; RITA DE CALUWE; GUY DE TRÉ; AXEL HALLEZ; JÖRG VERSTRAETE; MARC LEMAN; OLMO CORNELIS; DIRK MOELANTS; and JOS GANSEMANS. 2006. Similarity be-

- tween multi-valued thesaurus attributes: Theory and application in multimedia systems. *Flexible Query Answering Systems: FQAS 2006* (Lecture notes in computer science 4027), 331–42. Heidelberg: Springer.
- MEYER, DAVOD, and KURT HORNIK. 2009. Generalized and customizable sets in R. *Journal of Statistical Software* 31(2). DOI: 10.18637/jss.v031.i02.
- MURRAY, ELISABETH A.; STEVEN P. WISE; and KIM S. GRAHAM. 2017. *The evolution of memory systems: Ancestors, anatomy, and adaptations*. Oxford: Oxford University Press.
- NEVINS, ANDREW. 2015. Triumphs and limits of the contrastive-only hypothesis. *Linguistic Variation* 15.41–68. DOI: 10.1075/lv.15.1.02nev.
- NÍ CHIOSÁIN, MÁIRE, and JAYE PADGETT. 1993. Inherent V-place. Report LRC-93-09. Santa Cruz: Linguistics Research Center, University of California, Santa Cruz.
- ODDEN, DAVID. 2011. The representation of vowel length. *The Blackwell companion to phonology*, ed. by Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, Ch. 20. Chichester: John Wiley & Sons. DOI: 10.1002/9781444335262.wbctp0020.
- OXFORD, WILL. 2015. Patterns of contrast in phonological change: Evidence from Algonquian vowel systems. *Language* 91.308–57. DOI: 10.1353/lan.2015.0028.
- PURNELL, THOMAS, and ERIC RAIMY. 2015. Distinctive features, levels of representation, and historical phonology. In Honeybone & Salmons, 522–44. DOI: 10.1093/oxfordhb/9780199232819.013.002.
- RICE, KEREN. 1999a. Featural markedness in phonology: Variation, part I. *GLOT International* 4(7).3–6.
- RICE, KEREN. 1999b. Featural markedness in phonology: Variation, part II. *GLOT International* 4(8).3–8.
- RINGEN, CATHERINE O., and ROBERT M. VAGO. 2011. Geminates: Heavy or long? *Handbook of the syllable*, ed. by Charles E. Cairns and Eric Raimy, 155–69. Leiden: Brill. DOI: 10.1163/ej.9789004187405.i-464.47.
- SAGEY, ELIZABETH. 1986. *The representations of features and relations in non-linear phonology*. Cambridge, MA: MIT dissertation.
- SCHANE, SANFORD A. 1984. The fundamentals of particle phonology. *Phonology Yearbook* 1.129–55. DOI: 10.1017/S0952675700000324.
- SPAHR, CHRISTOPHER. 2016. *Contrastive representations in non-segmental phonology*. Toronto: University of Toronto dissertation. Online: <http://hdl.handle.net/1807/73177>.
- STANLEY, RICHARD. 1967. Redundancy rules in phonology. *Language* 43.393–436. DOI: 10.2307/411542.
- WICKLEGREN, WAYNE A. 1981. Human learning and memory. *Annual Review of Psychology* 32.21–52. DOI: 10.1146/annurev.ps.32.020181.000321.

Purnell & Raimy
 Department of English
 600 North Park Street
 University of Wisconsin – Madison
 Madison, WI 53706
 [tcpurnell@wisc.edu]
 [raimy@wisc.edu]

[Received 8 December 2018;
 revision invited 11 January 2019;
 revision received 21 April 2019;
 revision invited 17 May 2019;
 revision received 24 June 2019;
 accepted 14 July 2019]

Salmons
 Language Sciences
 1220 Linden Drive
 University of Wisconsin – Madison
 Madison, WI 53706
 [jsalmons@wisc.edu]