18. Diphthongization in Particle Phonology

SANFORD A. SCHANE

Subject	Theoretical Linguistics » Pholonogy
DOI:	10.1111/b.9780631201267.1996.00020.x

0 Introduction

At the heart of particle phonology is a set of unary features: the aperture particle |a| and the tonality particles |i| and |u|. These three particles encompass the traits of openness, palatality, and labiality, respectively. The different combinations of single occurrences of the particles (including the null set) yield the eight vowels in (1).[1]

(1)

[ɨ]/[ə]	[a]	[i]	[u]	[e]	[o]	[ü]	[ö]
—	a	i	u	a	a	i	a
				i	u	u	i
							u

The initial impetus for the theory of particle phonology stemmed from limitations of the binary distinctive features of generative phonology in expressing interrelationships between diphthongs and monophthongs. The standard notation, as presented in Chomsky and Halle (1968), did not reveal in an enlightening manner processes such as the diphthongization of [üː] or the monophthongization of [ai̯] to [eː].[2]

In the earliest work on particle phonology (Schane 1984a, 1984b), diphthongization was viewed as the splitting apart (fission) of a particle complex and a realignment of the component particles into the two halves of the diphthong. Conversely, monophthongization was the fusion of a sequence of particles into a simultaneously occurring configuration. Fission and fusion acted directly on the particles. In (2) we illustrate, for that early framework, the fission of [ü

ː

] into [i

u̯

] and the fusion of [a

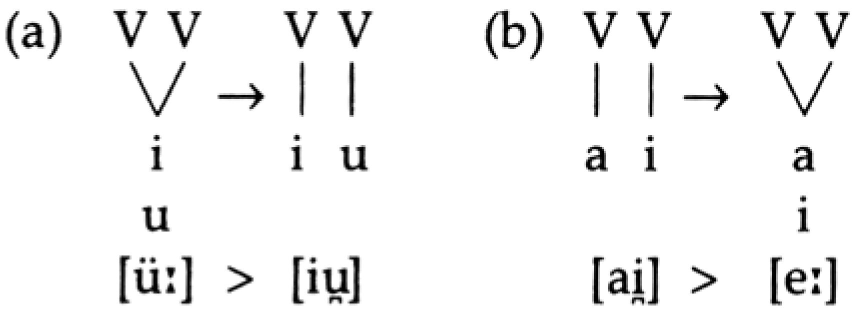
i̯

] into [e

ː

].[3]

(2)



Phonological representation has become considerably richer, and we shall see that the fission/fusion that characterizes diphthongization/monophthongization no longer operates on individual particles but rather on the root nodes that dominate those particles.

In this paper, we examine the phonological structure of vowels and diphthongs and the processes of diphthongization and monophthongization. We establish the following points: (1) In phonological representation, long vowels and falling diphthongs have two timing slots; short vowels and rising diphthongs have one. Diphthongs (whether falling or rising) have two root nodes; monophthongs (whether long or short) have one. (2) Fusion, which merges two root nodes into one, is the formal mechanism that underlies monophthongizations; fission, which splits apart one root node into two, accommodates diphthongizations. (3) Two kinds of OCP constraints govern the representation of diphthongs: the two halves of a diphthong have neither separate occurrences of the same particle nor all of their particles in common. (4) The notation for representing diphthongs refers to three parameters: particles shared by both halves of a diphthong, particles unique to the first half, and particles unique to the second half. The various monophthongization and diphthongization processes will manipulate one or more of these parameters. (5) Differences in vowel height are characterized by the number of aperture particles, and one-step shifts in height by the addition or deletion of an aperture particle. (6) One half of a diphthong can change without affecting the other half. Consequently, when a diphthong has multiple aperture particles, they may behave independently.

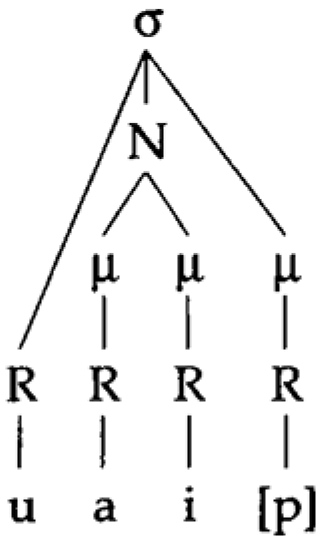
1 The Representation of Diphthongs and Monophthongs

We adopt a multitiered representation with five distinct levels: a syllable tier, a nucleus tier, a timing tier, a root tier, and a particle tier. Example (3) depicts the representation of English *wipe* [wa

i̯

p].[4]

(3)



The syllable node (σ) encompasses the various constituents of a syllable: an optional onset, an obligatory nucleus, and an optional coda. The nucleus node (N) dominates the different types of nuclear material: long and short vowels, syllabic nasals and liquids, and the peak and nonpeak components of diphthongs and triphthongs. Because there is no separate rhyme node, both onset and coda consonants are linked directly to the syllable tier. This notation is able to differentiate between those glides that function as onset or coda consonants and those that are elements of diphthongs. In (3) the initial consonantal glide is attached directly to the syllable tier, whereas the diphthongal glide is connected through the intermediary of the nucleus node.[5]

The μ nodes represent syllable weight: within the nucleus, a single μ node dominates one-mora entities (i.e., short vowels and rising diphthongs such as [

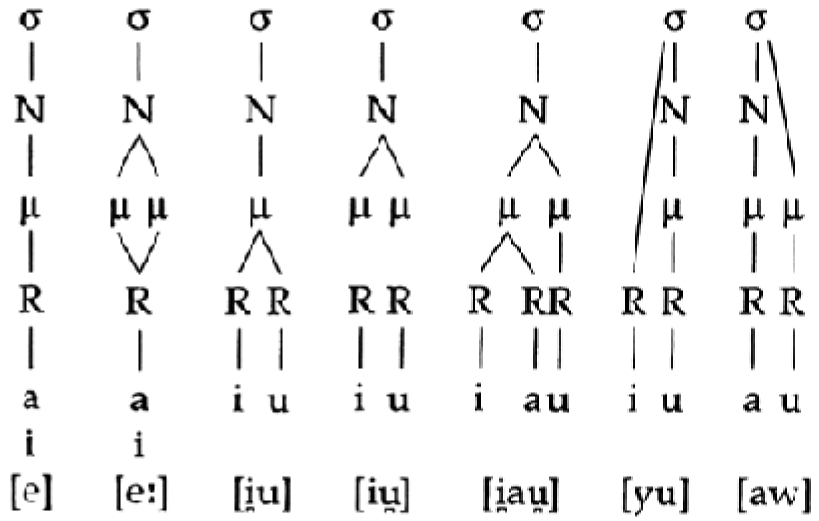
i̯

a]), whereas two μ nodes dominate entities that count as two moras (i.e., long vowels and falling diphthongs such as [a

i̯

].[6] The root node (R) organizes segmental content – that is, in the case of vowels and diphthongs it dominates the particles |a|, |i|, and |u|. Moreover, within the nucleus, the number of root nodes is what distinguishes monophthongs from dipthongs: a monophthongal vowel (whether long or short) always has a single R node, whereas each half of a diphthong must have its own R node. It is the R node that most closely embodies the notion of segment. A monophthong, even when long, intuitively is one segment; a diphthong corresponds to a sequence of two nonidentical vowels (segments) within a single nucleus. The representations of (4) depict various kinds of syllable structures.

(4)

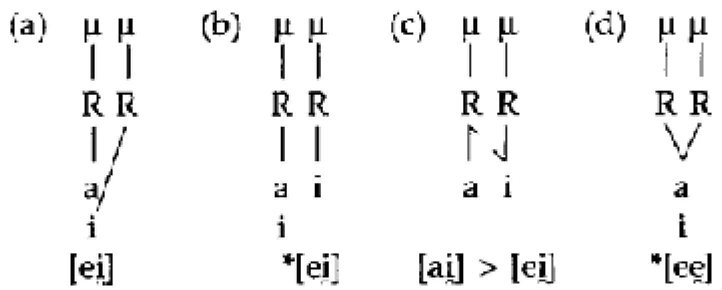


What kinds of connections are permitted between R nodes and the particles that they dominate? Within the syllable nucleus N, two contiguous R nodes may not dominate separate occurrences of the same particle. This constraint can be viewed as an OCP violation, a prohibition against sequential occurrences of the same autosegment at the melodic (particle) level.[7] For example, the diphthong [e

i

] must have the representation of (5a), where the two halves of the diphthong share the |i| particle, and not that of (5b), where each half would have its own |i| particle.[8]

(5)



The structure of (5a) finds corroboration from autosegmental spreading, the analog of assimilation. In (5c), we illustrate a partial assimilation whereby [a

i

] becomes [e

i

]. Here the |i| particle from the nonpeak element of the diphthong has spread to the root of the peak. (Spreading is indicated by the dotted line.) The resulting structure corresponds to that of (5a).

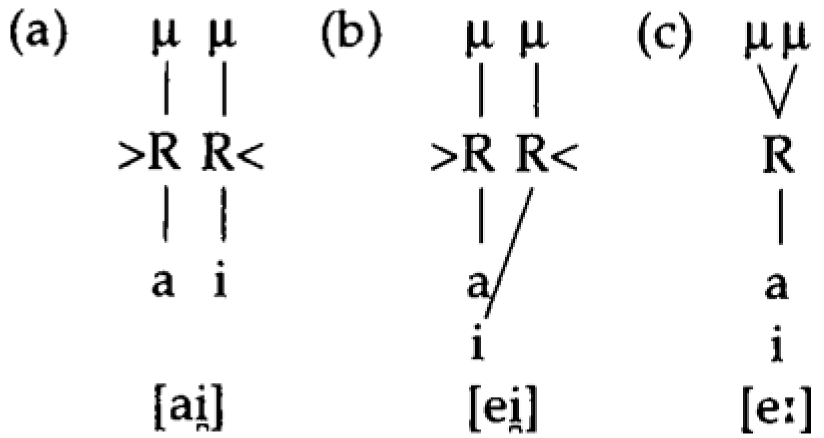
There is a second kind of OCP effect governing the representation of diphthongs. The two halves of a diphthong may not have all of their particles in common – that is, all the particles may not be doubly-linked to the two R nodes. The structure in (5d) represents this type of illegal configuration. A root node and its constituent particles constitute the formal analog of a segment. Consequently, adjacent root nodes dominating identical features would represent the same segment, and a sequential representation of R nodes would count as a violation of the OCP.[9] This representational constraint is in accord with a fundamental fact about the nature of diphthongs, what I have called *diphthongal differentiation* (Schane 1984a, 1989): The two halves of a diphthong may not be identical. That is, at least one of the halves must have a particle that it does not share with the other. This is not to deny the occurrence of phonetic sequences such as [ii̯] or [uu̯], but when such sequences are found, they function as noncontrastive variants of the corresponding long vowels [iː] and [uː]. Thus, whereas a language may have among its contrastive nuclei both [eː] and [ei̯] (viz., Early Modern English immediately after the Great Vowel Shift), no language will have a phonemic contrast between [iː] and [ii̯].[10]

The two kinds of OCP constraints find strong support in conversions of diphthongs and monophthongs. We turn now to these phenomena.

1.1 Monophthongization (Fusion)

Different diphthongs may monophthongize to the same vowel. For example, [e:] is a frequent outcome of the monophthongizations of both [ai̯] and [ei̯]. These two diphthongs and the resulting monophthong are represented in (6). The converging arrow heads around the R nodes in (6a–b) symbolize the fusion operation that will result in (6c).

(6)



For [a

i̯

] there is no sharing of particles between the two halves of the diphthong, whereas for [e

i̯

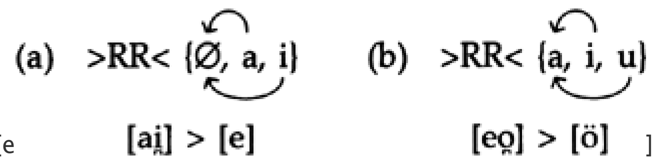
] there is sharing of the |i| particle. What is important, of course, is that both of these diphthongs (regardless of sharing) contain only the particles |a| and |i|, and the combining of these two particles into a single segment can result only in [e:]. Because it is the number of R nodes (i.e., two vs. one) that distinguishes diphthongs from monophthongs, we can define the process of monophthongization as the fusion (or closure) of the R nodes of a diphthong, the result being their complete overlap. As a consequence of fusion, any separately occurring particle from either half of a diphthong as well as any originally shared particle will find itself under the single R node of the monophthongal vowel.

There are more dramatic cases of different diphthongs fusing to the same monophthong. In Old French [

u̯

]and [e

u̯

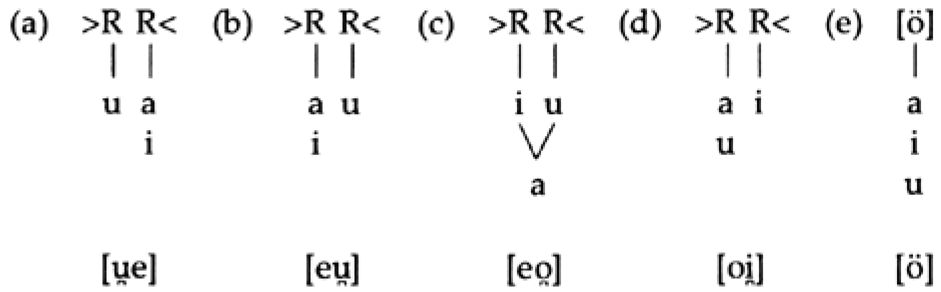


] both became [ö], in Old English [e] changed to [ö], and in some Greek dialects [o

i̯

] fused to [ö]. All of these diphthongs exemplify different distributions of the particles |a|, |i|, and |u| (see (7a–d)), and the monophthong [ö] is composed of the same particles (see (7e)).

(7)



1.1.1 Diphthong Notation

Because any given particle may be shared by both halves of a diphthong, may belong only to the first half, or may belong only to the second half, every diphthong can be specified by an ordered set of three parameters: {s, h_1, h_2}. In this formula *s* stands for the particle(s) shared in the diphthong, h_1 for the particle(s) unique to the first half, and h_2 for the particle(s) unique to the second half. One or two of the parameters may be null.[11]

The formulas of (8a–b) represent the diphthongs of (6a–b). (RR signifies the two R nodes that characterize a diphthong.) Because the two halves of the diphthong [a

i̯

] of (6a) share no particles, the *s* parameter is null for the representation of this diphthong in (8). On the other hand, for the diphthong [e

u̯

] of (6b) there is sharing of the particle |i|; consequently, this particle occupies the *s* position in the representation of (8b). But because the second half of this diphthong has no unshared particle, it is the h_2 position that is null. Additional examples of the notation are found in (8c–f), which represent the diphthongs of (7a–d).

(8)

(a) RR {Ø, a, i} [ai̯]
(b) RR {i, a, Ø} [ei̯]
(c) RR {Ø, u, ai} [u̯e]
(d) RR {Ø, ai, u} [eu̯]
(e) RR {a, i, u} [eo̯]
(f) RR {Ø, au, i} [oi̯]

When a diphthong turns into a monophthong there is fusion of the R nodes and, consequently, all particles will become shared. The notation readily lends itself to a representation of the resulting monophthong. In the formulas, as shown in (9a) for [e] and (9b) for [ö], a monophthong has one R and both *h* positions can only be null.

(9)

(a) R {ai, Ø, Ø} [e]
(b) R {aiu, Ø, Ø} [ö]

1.1.2 Extending the Notation: Fusion

The notation for diphthongs will allow us to express directly fusion processes. We need only add converging arrowheads around the R nodes of the original diphthong to indicate fusion. Arrows (from particles in the *h* positions) pointing to the *s* position indicate, for the resulting monophthong, the sharing of those particles that were unique to each half of the diphthong.[12] The monophthongizations of [a

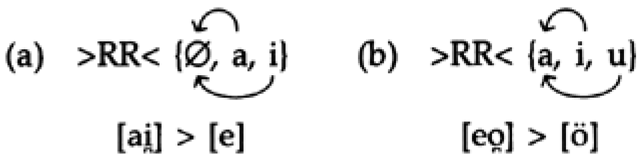
i̯

] to [e] and of [e

u̯

] to [ö] are depicted in (10). These notations function as rules that will convert the structures of (6a) and (7c) to (6c) and (7e), respectively.

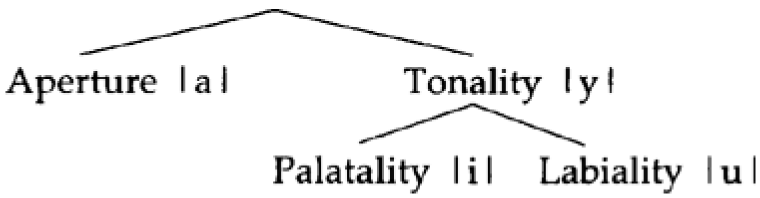
(10)



1.1.3 Extending the Notation: Tonality Vowels

It is traditional to portray vowel systems in a two-dimensional articulatory array, where the vertical dimension corresponds to tongue height, and the horizontal to tongue placement and lip configuration. Particle phonology makes a similar dichotomy. (11) shows a main division of aperture and tonality, and for the latter, a further division of palatality and labiality.

(11)



Front unrounded and back rounded vowels form a natural class – viz., the widespread symmetry in their distribution and in the phonological processes they undergo. It is convenient, then, that there be some way of referring, within rules, to any vowel with tonality. Following Ewen and van der Hulst (1988), we shall adopt |y| as a notational convention for either |i| or |u|. For example, the following rule of Sanskrit states that [a

i̯

] and [a

u̯

] monophthongize to [e

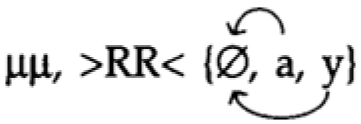
ː

] and [o

ː

], respectively:[13]

(12)



1.2 Diphthongization (Fission)

Fusion, as we have described it, entails that the particles that are distributed throughout the diphthong will occur intact in the resulting monophthong. It would seem that fission should be the converse operation – i.e., the fused particles of the monophthong should find themselves distributed between the two halves of the diphthong. Moreover, if fusion, as an operation, is the closure of two R nodes into one, then fission should involve the opening of a single R node into two. In the conversion of a monophthong to a diphthong, this splitting apart of the R node will indeed happen, but this operation alone would lead to an OCP violation, since the resulting R nodes would come to dominate identical particles. Because of the requirement of diphthongal differentiation, there will have to be a redistribution of the particles within each half of the diphthong. Consider, for example, the diphthongization of [ü

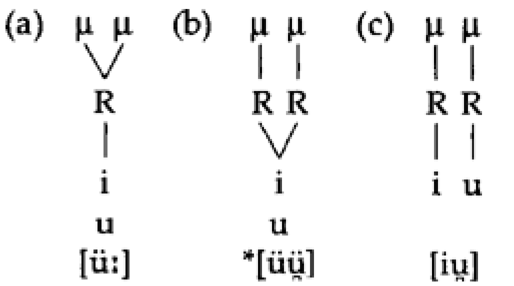
ː

] to [i

u̯

], a change that occurred in Faroese (Rischel 1968). The fission of the R node of the monophthong in (13a) yields the ill-formed intermediate structure of (13b), where both tonality particles are doubly-linked to the R nodes. This OCP violation has been resolved in (13c) by the delinking of |u| from the left R and of |i| from the right R. Note that both tonality particles end up unshared.

(13)



Because diphthongal differentiation requires that the two halves of a diphthong not have all particles in common, shared particles will be permitted just so long as there is also at least one unshared particle. The common change of [e

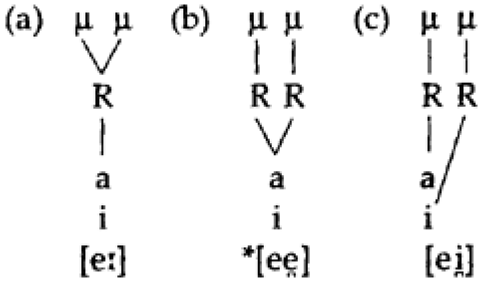
i̯

] to [e

i̯

] illustrates this kind of partial sharing. In the diphthong of (14c), the |a| particle has been delinked from the right R (i.e., it has become unshared), but both halves of the diphthong continue to share the |i| particle.

(14)



There is another means for diphthongal differentiation: all the particles from the monophthong may remain shared in the diphthong on condition that one of the halves acquire a particle not present in the original monophthongal configuration. This change is *de rigueur* where the monophthong contains a single particle, but it is not limited to this environment. The diphthongization of [i

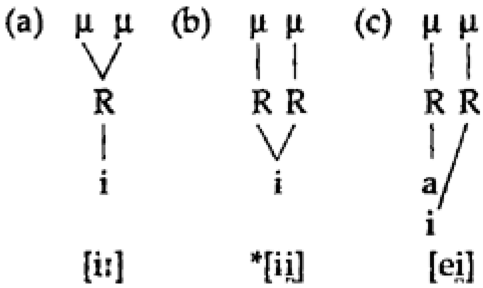
ː

] to [e

i̯

], which was part of the English Great Vowel Shift, provides an obvious example of this change, as shown in (15). The |i| particle continues to be shared in the resulting diphthong of (15c), but to comply with diphthongal differentiation, the left half has acquired an |a| particle.

(15)

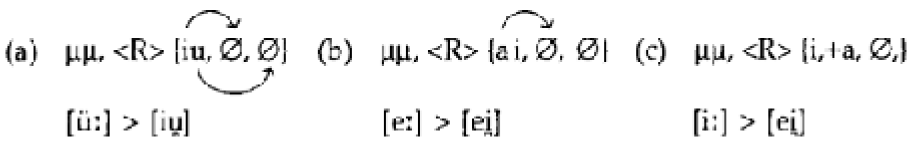


The addition of an aperture particle is a common way of resolving diphthongal differentiation for vowels of all heights.

1.2.1 Extending the Notation: Fission

The notation for diphthongization can be adapted to the representation of fission: (1) A single R node and the particles in the *s* position denote the original monophthong; diverging arrowheads around the R node signify its opening into two nodes. (2) An arrow (from a particle in the *s* position) pointing to one of the *h* positions indicates, for the resulting diphthong, the unsharing of that particle – i.e., it becomes unique to that half of the diphthong. (3) A particle preceded by + in one of the *h* positions is added to that position of the diphthong. The changes of [üː] to [iu̯], of [eː] to [ei̯], and of [ii̯] to [ei̯] are characterized in (16).[14]

(16)



1.3 Other Diphthongal Changes

Not only do monophthongs turn into diphthongs, and diphthongs fuse to monophthongs, but diphthongs change into other diphthongs. There are three avenues of change: (1) the two halves of the diphthong may become less alike (dissimilation); (2) the two halves may become more alike (assimilation); (3) the two halves may retain their degree of similarity but undergo feature exchange (i.e., corresponding to the delinking of a particle from one of the halves and its reattachment to the other).[15]

The halves of a diphthong will become less alike whenever a particle shared by both halves is delinked from one of them (i.e., unsharing) and/or one of the halves acquires a new particle. The change of [ou̯] to [au̯] in (17a) illustrates unsharing of the |u| particle, that of [au̯] to [e

u̯

] in (17b) the acquisition of an |i| particle, and that of [u̯o] to [u̯e] in (17c) unsharing of |u| and the acquisition of |i|. In the conversion of one diphthong to another, the R nodes are not affected (i.e., neither fission nor fusion has taken place). Arrows show the movement of particles (i.e., delinking and/or sharing) from the *s* and *h* positions of the original diphthong to those of the derived one.

(17)

(a) μμ, RR {u, a, ∅}
[ou̯] > [au̯]

(b) μμ, RR {∅, a+i, u}
[au̯] > [eu̯]

(c) μ, RR {u, ∅, a+i}
[u̯o] > [u̯e]

The halves of a diphthong will become more alike whenever a particle *unique to one half* becomes shared or else is deleted. In the notation, a particle preceded by – in one of the *h* positions is deleted from that position of the diphthong. The change of [a

u̯

] to [e

u̯

] in (18a) illustrates the sharing of |i|, and that of [e

u̯

] to [i

u̯

] in (18b) the deletion (i.e., complete delinking) of |a|.

(18)

(a) μμ, RR {∅, a, i}
[ai̯] > [ei̯]

(b) μμ, RR {∅, i–a, u}
[eu̯] > [iu̯]

Feature exchange involves delinking and reattachment of a particle from one half of a diphthong to the other. This process often accompanies the shift of a falling diphthong to a rising one. The change of [e

i̯

] to [

u̯

o] is an example of this kind of particle exchange. Here there have been a delinking and a reattachment of the |a| particle. In the notation, a particle has been shifted from one of the *h* positions to the other. The fusion of the μ nodes in (19) is what converts a falling diphthong of two moras into a rising diphthong that counts as one mora.[16]

(19)

>μμ<, RR {∅, ia, u}
[eu̯] > [i̯o]

1.3.1 Chain Shifts

Within the evolution of a language there may be a string of diphthongal changes. Romance and Germanic richly exemplify such chains. Vulgar Latin [e

i̯

] changed to Old French [e

i̯

] (diphthongization), then to Middle French [o

i̯

] (dissimilation), then to [

u̯

e] (feature exchange), and then finally to Modern French [

u̯

a] (assimilation). This sequence of events is depicted in (20).

(20)

[eː] > [ei̯]	μμ, <R> {ia, ∅, ∅}	(fission of R; unsharing of |a|)
[ei̯] > [oi̯]	μμ, RR {i, a+u, ∅}	(unsharing of |i|; addition of |u|)
[oi̯] > [u̯e]	>μμ<, RR {∅, ua, i}	(fusion of μ; exchange of |a|)
[u̯e] > [u̯a]	μ, RR {∅, u, a–i}	(deletion of |i|)

Old High German [e

ː

] took an entirely different route. It diphthongized to [e

i̯

], which became [i

i̯

] (dissimilation), then [i

u̯

] (assimilation), and then Modern German [i

u̯

] (assimilation), which is equivalent to [i

u̯

]. These Germanic changes are shown in (21).

(21)

[e:] > [ea̯]	μμ, <R> {ai, ∅, ∅}	(fission of R; unsharing of |i|)
[ea̯] > [ia̯]	μμ, RR {a, i, ∅}	(unsharing of |a|)
[ia̯] > [ie̯]	μμ, RR {∅, i, a}	(sharing of |i|)
[ie̯] > [ii̯]	μμ, RR {i, ∅, −a}	(deletion of |a|; OCP violation)
[ii̯] > [i:]	μμ, >RR< {i, ∅, ∅}	(resolution of OCP violation; fusion of R)

Note that from these notations one can easily construct the corresponding tree structures.

2 The Representation of Height

The combination of single particles (including the null set) yields the eight vowels that were shown in (1). With the exclusion of [a], this system can accommodate at most two vowel heights. In particle phonology, additional occurrences of the aperture particle specify vowels of lower height. The common seven-vowel system, for example, has the particle structure shown in (22).[17]

(22)

[i]	[e]	[ɛ]	[a]	[ɔ]	[o]	[u]
i	a	a	a	a	a	u
	i	a	a	a	u	
		i		u		

The representation for [a] is a function of the other vowels in a language system and of the number of aperture particles for characterizing them. The vowel [a] will have the same number of aperture particles as the lowest tonality vowel(s). In Schane (1984a), I refer to this phenomenon as the Law of Maximum Aperture, because whenever [a] interacts with a tonality vowel, the interaction most typically involves the lowest one.[18] Moreover, the vowel [a] will have more than one aperture particle whenever there are several central vowels. For the series [i ə a], there will be zero, one, and two aperture particles, respectively.[19]

2.1 Extending the Notation: Vowel Height

In rules, one needs to be able to refer to natural classes of vowel height. Consider the three-height system: [i e ɛ]. All three vowels as well as any pair of adjacent vowels will form a natural class. Within the notation we need a way of specifying minimum and maximum number of |a| particles. We shall use the standard device of subscript and superscript numerals, or, alternatively, of parenthesized elements. (23) illustrates these notational conventions for the natural classes derived from [i e ɛ].

(23)

[i e]	$\{ia_0^1\} = \{i(a)\}$
[e ɛ]	$\{ia_1\} = \{ia_1^2\} = \{ia(a)\}$
[i e ɛ]	$\{ia_0\} = \{ia_0^2\} = \{i(a)(a)\}$

Note that this notational system correctly excludes classes composed of noncontiguous vowel heights.

On occasion, one may need to refer to the absence of a particle. We shall use a superscript ^{0} (with no subscript) to indicate no occurrences of a particle. For example, a^0 would represent the natural class of high vowels.

2.2 Justification for Multiple Aperture Particles

The occurrence of multiple aperture particles has been criticized by some of the adherents of dependency phonology (see den Dikken and van der Hulst (1988), but see also Hayes 1990 for a defense). The original justification came from multiple one-step shifts in vowel height. For example, in the Great Vowel Shift, [eːoː] shifted to [iːuː], and [ɛː] to [eːoː]. This upward one-step movement can be succinctly summarized as loss of an aperture particle from any long vowel having a tonality particle and one or two aperture particles.[20]

(24) μμ, R {y(a) – a}

I shall now adduce another argument for multiple aperture particles: They may function independently. This independence is most striking for diphthongs. As shown in (25), a diphthong, such as [eɛ], with two |a| particles, has one of them shared by both of its halves but the other is unique to only the second half.

(25)

Each aperture particle should be able to undergo some change independently of the other. This phenomenon, which often interacts with movement along the height scale, does indeed exist, and it provides compelling confirmation for the multiple occurrence of aperture particles. In support of this claim, we shall look at some data from Frisian and from Malmö Swedish.[21]

2.2.1 West Frisian Breaking

Standard West Frisian has nine vowels – [i e ε u o ε ü ö a] – that occur both long and short. A tenth vowel [e] (schwa) occurs only short.[22] Because the front unrounded and back rounded vowels exhibit three heights, the particle representation for [a], in conformity with the Law of Maximum Aperture, will require two |a| particles. The schwa vowel has a root node with no particles under it. (See n. 1.)

(26)

[i]	[e]	[ε]	[u]	[o]	[ɔ]	[ü]	[ö]	[a]	[ə]
i	a	a	u	a	a	i	a	a	—
	i	a		u	a	u	i	a	
		i			u		u		

Now the higher and lower mid vowels of the standard dialect have two different diphthongal reflexes in the “breaking” dialects. The latter exhibit alternations between falling and rising diphthongs, as illustrated by the following pairs (de Graaf and Tiersma 1980, pp. 110, 112).

(27)

	Sg.	Pl.	Sg.	Pl.	
e(:)	iə̯	i̯e	stiən	stienən	stone(s)
			viət	vietə	wet(ness)
ε(:)	eə̯	i̯ε	beəm	biεmən	tree(s)
o(:)	uə̯	u̯o	stuəl	stuolən	stool(s)
			guət	guodlək	good(ly)
ɔ(:)	oə̯	u̯a	doər	duarən	door(s)
			moəj	muajər	pretty/ier

In these same environments, standard West Frisian shows alternations between monophthongal long and short vowels (de Graaf and Tiersma 1980, p. 113).

(28)

Sg.	Pl.	Sg.	Pl.	
iː	i	viːf	vifkə	(small) woman
eː	e	heːx	hextə	high/height
ɛː	ɛ	bɛːt	bɛtsjə	(small) bed
aː	a	baːrx	bargən	pig(s)
üː	ü	slüːf	slüfkə	(small) envelope
öː	ʌ	gröːt	grʌtər	larg(er)
uː	u	muːs	muzən	mouse/mice
oː	o	knoːp	knopkə	(small) button
ɔː	ɔ	lɔːn	lɔnən	land(s)

Historically, all dialects had both long and short vowels. In standard Frisian, the long vowels were shortened when followed by another syllable or by a cluster of two or more consonants. In these same environments the dialects with breaking have rising diphthongs, but they exhibit falling diphthongs where standard Frisian has retained long vowels. To account for the evolution of original long vowels in the breaking dialects, van der Meer (1977) claims that the four long mid vowels became falling diphthongs whose first halves were one step higher in height than the original monophthongs. These diphthongs then changed in two different ways: in the "shortening" environments, they became rising diphthongs with high first components; otherwise, their second half changed to schwa.

(29)

Monophthong	Diphthongization	Shortening environment	Elsewhere
eː	ie̯	i̯e	iə̯
ɛː	eɛ̯	i̯ɛ	eə̯
oː	uo̯	u̯o	uə̯
ɔː	oɔ̯	u̯ɔ > u̯a	oə̯

The notational representation for this chain of events is provided in (30)

(30)

(a) μμ, R {ya(a), Ø, Ø} (original monophthongs)

(b) μμ, <R> {ya(a), Ø, Ø} (diphthongization)

(c) >μμ<, RR, {y(a), Ø, a} (shortening environment)

(d) μ, RR {u, Ø, aa} (u̯ɔ > u̯a)

(e) μμ, RR, {y(a), Ø, –a} (elsewhere)

We see in (30a) the representation for the class of vowels that undergoes breaking – i.e., each of the four vowels possesses a tonality particle and one or two aperture particles. In the diphthongization of (30b), the first half of each diphthong must become one dgree higher in height. In addition to the fission of the R node, there is an unsharing of an aperture particle (i.e., it is delinked form the first half of the diphthong); the other aperture particle, if present, continues to be shared. Here we see the first example of different behaviors manifested by aperture particle. In (30c), there is conversion of a falling diphthong to a rising. Recall that this change for diphthongs in the breaking dialects is equivalent to the shortening of monophthongs in standard Frisian. The equivalence has an identical notational representation: for both kinds of changes, there is fusion of μμ. In addition, the first half of the rising diphthong must become high (if it is not already high). A shared [a] particle, if present, must be delinked from the first half. The unshared aperture particle of the second half is unaffected. Here is another example of multiple aperture particles be having differently. In (30d), the rising diphthong [

u̯

ɔ

] is adjusted to [ua]. Its tonality particle becomes unshared. The second half of the diphthong will retain its two [a] particles, the representation for [a] in Frisian. In (30e), the second half of those falling diphthongs that did not change to rising will become [e]. Recall that the representation of Frisian schwa has no particles. Consequently, all particles must be delinked from the second half–i.e., any shared particle(s) will become associated only with the first half of the resulting diphthong, and the aperture particles unique to the second half will be deleted. Here too aperture particles are affected differently. An originally shared aperture particle becomes unshared, and an originally unshared one gets deleted.

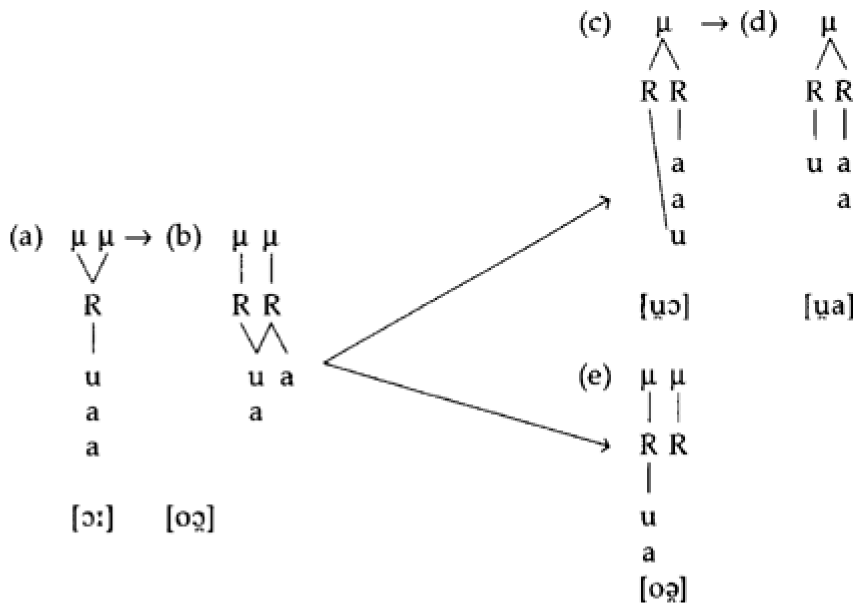
To illustrate the application of the rules in (30), we provide in (31) derivations for the vowel [

ɔ

ː

].

(31)



2.2.2 Malmö Swedish Height Shift

Diphthongizations of aperture can take one of two forms: delinking of an aperture particle or acquisition of one. Delinking causes a one-step upward shift in height, and acquisition a one-step downward shift. The first stage of diphthongization in Frisian exemplifies upward shift. As an example of down-ward shift, we consider the multiple diphthongizations of the Malmo dialect of Swedish.

In Malmö Swedish, all long vowels diphthongize (Bruce 1970; Lindau 1978). The first halves of the resulting diphthongs take a one-step downward shift in height, and those emanating from back rounded vowels also become front unrounded. (See(32).)

(32)

/iː/ > [ei]	/üː/ > [öü]	/uː/ > [eu]
/eː/ > [ɛe]	/ʉː/ > [öʉ]	/oː/ > [ɛo]
/ɛː/ > [æɛ]	/öː/ > [œö]	/a/ > [æa]

The particle representations for the Malmo vowels are presented in (33)[23].

(33)

[i]	[e]	[ɛ]	[æ]	[ü]/[ʉ]	[ö]	[u]	[o]	[a]
i	a	a	a	i	i	u	u	u
	i	a	a	u	u		a	a
		i	a		a			a
			i					

Two rules are required: a rule that diphthongizes long vowels by a one-step downward shift in height in the first half of the diphthong, and a rule that makes that first half front unrounded when the original vowel was back rounded. The rules are given in (34a-b)[24].

(34)

(a) $\mu\mu,\ \langle R\rangle\ \{\alpha, +a, \varnothing\}$

(b) $\mu\mu,\ RR\ \{a_0 i^0 u,\ a+i,\ \varnothing\}$

The diphthongization rule in (34a) adds an aperture particle to the first half of a diphthong. All original particles (tonality and/or aperture), which are represnted in the rule by α remain shared. The tonality rule in (34b) applies to the class of back rounded vowels (i.e., those with any number of aperture particles, no palatal particle, and a labial particle). The rule adds an [i] particle and delinks an [u] particle in the first half of the diphthong. The Malmo shift provides another illustration of the independence of aperture particles. Any aperture particle(s) from an original monophthoong will remain intact and will continue to be shared by both halves of the derived diphthong. The added aperture particle, needed for the one-step downward shift in height, will be associated with the first half only.

3 The Diphthongization Paradox

Hayes (1990) contains a discussion of the notational problems of diphthong representation and a proposal for the expression of diphthongization processes. The “diphthongization paradox” to which Hayes alludes can be stated as follows: If monophthongs require one root node in their representations but diphthongs need two, then, in conversions between these two kinds of vocalic entities, how is this discrepency in number of root nodes to be handled? Hayes critiques three other proposals and he presents his own solution to the problem.

Clements (1985a) supposes that long segments (i.e., those with two timing units) have two root nodes at the outset. Each featural daughter node is doubly-linked to the R nodes. Diphthongization, then is simply the addition or delinking of one or more daughter features from either R node. Hayes finds this solution to be flawed because "it contradicts the basic principle of prosodic theory that long segments are doubly-linked single units" (Hayes 1990, p. 36). That is, long segments should have one R node.

Steriade (1987b) suggests eliminating altogether root nodes from phonological representations and, instead, linking certain feature nodes directly to timing units. Since feature nodes, under this approach, would be doubly-linked to the two timing nodes of a long segment, for dipthongization processes, they could be individually delinked from one of the timing nodes or a new feature could be added to it. Hayes points out that the absence of root nodes would play havoc with total assimilation, which in autosegmental treatments, has been viewed as the spreading of root nodes[25]. He argues also that Steriade's proposal would require low-level features to have direct access to the timing nodes, a maneuver that would essentially wipe out the higher-level feature nodes advocated in feature geometries.

Selkirk (1990) sets up two root nodes for long segments. For vowels, the root nodes would also bear the feature value [- cons]. Daughter nodes are permitted to be doubly-linked only when they are connected to identical root nodes, which will be the case for long vowels. Dipthongization is the relabeling of one of the root nodes-the one to become the nonpeak of the diphthong-as [+cons]. Because the root nodes are no longer identical, by general convention, there will be delinking of all features from the [+cons] root, Feature values that are to remain identical in the two halves of the diphthong will be spread from the [-cons] root back to the [+cons] root, and any feature values unique to one of the halves will be added. Hayes notes that this theory erroneously assumes that the glide elements of diphthongs are always [+cons]. The [+cons] specification may be appropriate for those glides that function as onset or coda consonants, but not for those that are in the nucleus. Moreover, the mechanism is overly cumbersome. Instead of delinking only those features that will not be shared by both halves of the diphthong, Selkirk is forced to delink all features from one of the halves, only to spread many of them back again.

Hayes (1990) claims that the diphthongization paradox arises because of ambiguity in the interpretation of association lines in autosegmental representations containing feature trees. The line between a higher-level node in a feature tree (e.g., Root) and a lower-level node (e.g., the feature [high]) indicates category membership, whereas the line between a timing unit (e.g., μ, X, or V) and the element it dominates signifies simultaneous realization in time. Hayes proposes to separate these two functions by getting rid of autosegmental tree structure. Instead, he develops a notation with labeled brackets for indicating category membership and with indices for temporal relations. Long vowels continue to have two adjacent timing units (labeled V_1 and V_2, respectively). A single root element and single feature values are coindexed to the timing units (e.g., R_{12}, $[-high]_{12}$. Diphthongization becomes the deletion of one of the indices of a feature specification and its replacement by the opposite specification (e.g., R_{12}, $[-high]_1$, $[+high]_2$)[26].

In Hayes's representation of long vowels and diphthongs there are two timing units. But how many root entities are there? Is R_{12} equivalent to one root node or to two? Perhaps ths question is unfair, since it is Hayes's intention to distance his notation from a conventional tree structure, where one would have no difficulty in counting number of nodes. In any case, what is important for us to notice is that in Hayes's notation, diphthongs and long vowels have an identical specification at the root, namely R_{12}. Hence, the distinction between diphthongs and monophthongs does not reside in the specification of the root, but is to be found at a lower level of the representation, where some feature (s) will bear one or two indices, respectively. Hayes too, like Clements, Steriade, and Selkirk, ends up positing an identical root con-figuration (or its complete absence for Steriade) for both long vowels and diphthongs. The notion that monophthongs correspond to one segment and diphthongs to two is not an aspect of any of these representations.

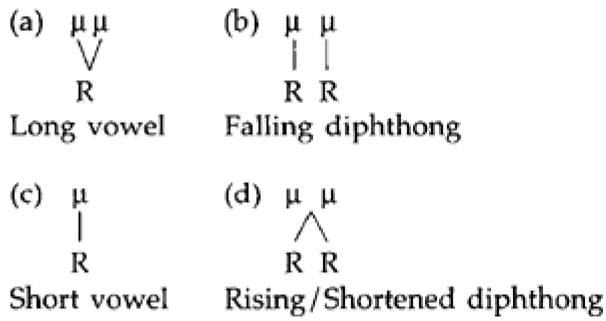
The frameworks of all four of these researchers permit diphthongs to be derived only from long vowels, because of the nature of double linkages. Both Clements and Selkirk allow a long vowel to have two R nodes, one corrsponding to each timing unit. It is dubious, though, that they would permit two R nodes to emanate from the single timing unit of a short vowel, which would be required in order to derive a diphthong from it. For Steriade, features can be doubly-linked to the two timing nodes of a long vowel; it is the delinking of a doubly-linked feature from one of the timing slots that converts a monophthong to a diphthong. But there is no way for features to be doubly-linked to the single timing node of a short vowel, the structure that would be needed in order for a short vowel to diphthongize. Hayes requires that multiple indices be associated with the same number of timing units. The diphthongization of a short vowel would necessitate the unlikely structure of features with two indices (so that one of them could be deleted) that are associated to a vowel with only one timing unit.

Although diphthongization, as a process, gnerally applies to long vowels, there are languages, such as the Romance family, that have synchronic alter-nations of diphthongs and short vowels. Moreover, there are "shortened" diphthongs, such as those found in some of the Scandinavian languages, which occur in the same environments as short vowels, and then there are rising diphthongs, such as [ia] and [ui] in Sanskrit, that count as one mora. An adequate theory of diphthongization must allow for the representation of shortened and rising diphthongs and for the possibility of deriving them from short vowels.

3.1 The Particle Answer to the Diphthongization Paradox

I claim that, in their phonological representations, shortened diphthongs and those rising ones that count as one mora, have one timing unit but two root nodes. Falling diphthongs that are not in shortening environments always count as two moras and, consequently, they have two timing units and two root nodes. In (35) we see configurations of timing and root nodes for long vowels, falling diphthongs, short vowels, and rising and shortened diphthongs.

(35)



Standard autosegmental tree structure has no difficulty in representing these four types of vowels and diphthongs. The problem for autosegmental phonology has been the inability to convert a monophthongal representation to a diphthongal one and vice versa. The so-called diphthongization paradox, in reality, stems from an inadequacy of the standard operations of autosegmental phonology (i.e., insertion, deletion, spreading, delinking) for increasing or decreasing the number of root nodes in a tree without, at the same time, disrupting connections with other nodes.

The particle phonology operations of fission and fusion *can* increase and decrease the number of nodes without any other disruptions. For diphthongization, the fission or splitting apart of the R node into two R nodes retains the connections to the timing and feature (particle) tiers. The features, which were singly-linked for the monophthong, automatically will become doubly-linked in the dephthong. With all features doubly-linked (shared), the two halves of the diphthong would be identical. Hence, diphthongal differentiation has to kick in: one or more features may be delinked from one of the halves, or else a new feature value may be adjoined to one of the halves.

For monophthongization, the fusion of two R nodes into one has interesting consequences for the features dominated by those nodes. The features of the original diphthong, whether doubly-linked to both R nodes or linked to only one of them, will end up dominated by the single R node of the monophthong. Fusion is similar to the set operation of union: If R_1 dominates the particle set {ai} and R_2 {au} (where [a] is shared in the tree), then fusion yields R, which dominates the particle set {aiu}.

This set theoretic union works only with unary features, and it is this property of fusion that was the original inspiration for particle phonology. Note what would happen with binary values: If R_1 dominates the feature set {+low, −back} and R_2 {+low, +back} (where [+low] is shared in the tree), the fused set R containing {+low, −back, +back} has contradictory specifications for the feature [back] [27].

The solution to the diphthongization paradox does not require that one abandon the tree structures of autosegmental representation, as Hayes was compelled to do. It does require, though, phonological operations and a featural representation that hitherto have not been standard characteristics of autosegmental phonology. The solution to the diphthongization paradox requires the recognition of fission and fusion as autosegmental operations and the utilization of unary features for segmental specification. There is, however, one aspect of diphthongization on which we are in complete agreement with Hayes.He argues that diphthongizations involving height shifts demonstrate the appropriateness of multiple occurrences of the aperture particle for characterizing vowel height differences. Only in this way is one able to account for the independence of height chages that can take place in each half of a diphthong.

4 Conclusion

Because this paper deals with diphthongization and monophthongization, the processes of fission and fusion have acted primarily on root nodes. These processes, however, turn out to be much more general. They are not restricted just to roots. Any node above the root is a candidate for fission/fusion[28]. We have already seen in (19), (20), and (31c) some instances where the timing node is affected. The fission of a μ node will convert a short vowel to a long one, or a rising diphthong to a falling; conversely, the fusion of two μ nodes will convert a long vowel to a short one, or a falling diphthong to either a rising or a shortened one (Schane 1989). The fusion and fission of μ nodes can replace the operations of deletion and insertion of a μ node, which are the only devices otherwise available in autosegmental phonology for handling vowel shortenings and noncompensatory lengthenings.

The fusion of adjacent syllable nodes (σ) and their daughter nucleus nodes (N) is a common way of dealing with vowel hiatus (Schane 1987). For example, in Sanskrit vowel sandhi, the disyllabic sequences [i]+[a] and [a]+[i] become monosyllabic [

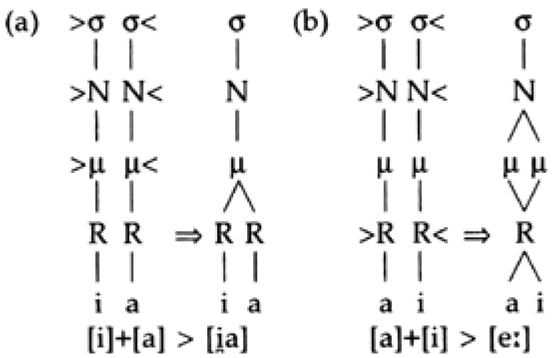
i̯

a] and [e

ː

], respectively. In (36), we illustrate these operations. In both cases, there is fusion of the σ and N nodes as well as one of the lower nodes: the μ nodes for [i]+[a], which yields a rising diphthong; and the R nodes for [a]+[i], which produces a long monophthong.

(36)



1 Languages with [ə] and no [i] have the schwa vowel as the null particle set. Within a set there is no ordering of particles. Hence, {ia} and {ai} both represent [e]. In dependency phonology (see chap. 17, this volume), which also makes use of unary vowel components, hierarchic ordering plays a role: [i] dominating [a] represents a higher mid vowel; [a] dominating [i], a lower mid. In particle phonology, vowels of lower height have additional occurrences of the [a] particle; see section.

2 Because a sequence of segments was involved, processes affecting diphthongs necessitated a transformational format. For example, the monophthongization of [ai] to [e:] (ignoring the length of the resulting vowel) required conversion of the first element of the diphthong to [e] and deletion of the second:

$$\begin{bmatrix} \text{V} \\ +\text{low} \\ +\text{back} \\ -\text{round} \\ 1 \end{bmatrix}\begin{bmatrix} \text{V} \\ +\text{high} \\ -\text{back} \\ -\text{round} \\ 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ -\text{low} \\ -\text{back} \end{bmatrix} \varnothing$$

3 In (2), V represents a vocalic timing unit: long vowels and falling diphthongs have two V slots.

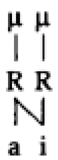
4 As I am not concerned with the representation of consonants, other than the glides, I have given no feature analysis for the segment [p] in (3).

5 Language-internal criteria determine how glides will function. In English, the variant a (*an) of the indefinite article before word initial [Y]and [w] demonstrates the functioning of these glides as onset consonants, and morphophonemic alternations of vowels and diphthongs (e.g., *divinity:divine, profundity:profound*) establish post-vocalic glides as belonging to the nucleus. For the transcriptions enclosed within square brackets, [Y] and [w] represent glides functioning as onset or coda consonants; [i] and [u], glides functioning as nonpeaks of diphthongs. This distinction between consonantal and nuclear glides is phonological; their phonetic manifestations may often be indistinguishable.

6 In many languages a coda consonant may also count as a mora (see chaps. 5 and 6, this volume).

7 Contiguous R nodes may dominate separate occurrences of a particle whenever their vowels belong to separate morphemes or are in hiatus (i.e., belong to different syllables). In such cases, each R node will be dominated by its own N node.

8 From here on, in order to economize on space, I shall omit the syllable and nucleus nodes. All of the structures to be illustrated are dominated by one occurrence of each of these nodes. In the representation of (5a), the vertical alignment of the particles is for ease of deciphering and has no other import. The horizontal representation below would be another way of representing [ei], and I shall make use of that mode on occasion.



The vertical and horizontal representations are equivalent, for both portray the sharing of the [i] particle by the two halves of the diphthong and the possession of the [a] particle by the left half only.

9 William Morris first brought to my attention the characterization of this kind of OCP violation.

10 A language may contrast [i:] and [iy]; however, the latter sequence is not a diphthong (i.e., with both elements in the nucleus), but a vowel followed by a coda consonant.

11 Two of the parameters can be null only where one of the halves is [i]. Note that both h_1 and h_2 may not be null. There would then be a total sharing of all particles- i.e., a monophthong (see (9)).

12 The arrows are redundant, in actuality, for the sharing of all particles is deducible from the very nature of fusion. I have included them, though, because they are essential for the representation of fission (see sec.

1.2.1) and of other kinds of diphthongal changes (see sec. 1.3).

13 From here on, we shall indicate in the formulas the number of moras. For the Sanskrit example, this tells us that the resulting vowels are long and that they arise from falling diphthongs.

14 Recall that μμ represents a long vowel where there is one R node, and a falling diphthong after fission of that node.

15 See Donegan (1978) for a thorough treatment of these different kinds of diphthongal changes.

16 In Japanese, [e

u̯

] became [

i̯

o

ː

] (Poser 1986, p. 181). The lengthened peak of the rising diphthong preserves the mora count of the original falling diphthong. For this change, there is no fusion of μ nodes; rather, concomitant with the particle exchange, the R node emanating from the second μ spreads back onto the first μ. The following tree structures depict [e

u̯

] > [

i̯

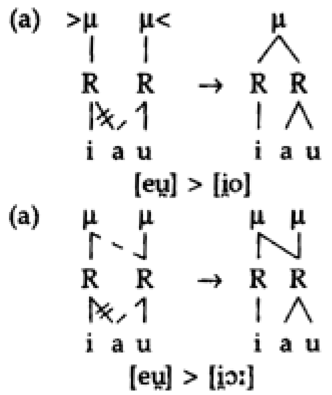
] > [

i̯

o

ː

]. (Dotted lines represent spreading.)



17 Since vowel systems with more than four heights are rare, in general there can be at most a maximum of three aperture particles – e.g., for a system with the four front unrounded vowels [ie ε æ], the vowel [æ]

would have three [a] particles.

18 The pairing of [a] with a particular vowel in vowel-harmony systems illustrates the Law of Maximum Aperture. Turkish and Hungarian have two front unrounded vowels – [i] and [e], but Finnish has three- [i], [e], and [æ]. In conformity with the Law of Maximum Aperture, the representation for [a] will have one [a] particle in the former two languages, but two of them in the latter. These representations find corroboration in palatal harmony, which is the addition of an [i] particle to an underlying nonfront vowel. In Turkish and Hungarian, [a] alternates with [e]; in Finnish, with [æ]. See the discussion in Goldsmith (1987c).

19 For languages with [ə a], (but without [i]), the schwa vowel will generally be without any aperture particle. See (1).

20 There is no need to indicate the null h_1 and h_2 positions when one monophthongal vowel is converted to another; *h* positions are needed for the representations of diphthongs and of conversions between diphthongs and monophthongs. In Chomsky and Halle (1968), combinations of the features [high], [low], and [tense] specify vowel height. Because three binary features are attempting to describe what is essentially a continuum, in order to express shifts in height, the notation necessitates a complex use of Greek-letter variables (see Yip 1980a).

21 Hayes (1990) analyzes similar data from Eastern Finnish, Lund Swedish, and Quebec French. In the analyses of these languages, Hayes makes use of aperture particles, and he argues that the data provide convincing evidence for multiple occurrences of |a|. However, his mechanism for diphthongization is radically different from mine. He proposes a notation for indexing all nodes and for percolating indices from higher nodes downward. I evaluate Hayes's framework in section 3.

22 The data on Frisian breaking are taken from de Graaf and Tiersma (1980). The vowel [ö

ː

] has [ʌ] as its short counterpart. I have represented the short high mid vowel as [e] (instead of using their [I]).

23 The vowels [ü] and [u] are both high front rounded: The rounding for the former is made with protruded lips, and that for the latter with compressed lips (Lindau 1978, p. 548). I shall not venture on what the particle distinction should be, as that determination is not relevant for the shifts treated here. Bruce (1970, p. 13) describes [a] as a back rounded vowel whose feature composition is equivalent to [

ɔ

].

24 In (34a), alpha stands for any set of particles.

25 In total assimilation, it is the root node of the source segment that spreads to the μ (timing) node of the target segment with concomitant detachment of all featural nodes that were linked to the target. In this way, the target will come to share all the segmental content of the host's R node.

26 For example, the conversion of [e: o:] to [e

i̯

o

u̯

] would have these changes as part of the representation. Other features would be required, of course, for the representation of this particular diphthongization process. See Hayes (1990, p. 47) for the complete specification required for the diphthongization of [o:] to [o

u̯

].

27 Fusion could be made to operate with underspecified binary features on condition that default and redundant specifications are filled in only after the monophthong has been derived. However, de Haas (1988, pp. 46–47) has demonstrated that radical underspecification (Archangeli 1988a), which allows one of the vowels to be totally unspecified, may give the wrong results. For example, if [u] is specified as [+round], [a] as [−high], but [i] is unspecified, the fusion of [a] + [u] would correctly yield [o] (=[+round, −high]), whereas that of [a] + [i] would erroneously give [a] (=[−high]). In order for [a] + [i] to yield [e] (=[−back, −high]), the vowel [i] needs to be specified minimally as [−back].

28 Fission/fusion will not apply below the root node, i.e., it will not apply to any feature node or particle.

Cite this article

SCHANE, SANFORD A. "Diphthongization in Particle Phonology." *The Handbook of Phonological Theory*. Glodsmith, John A. Blackwell Publishing, 1996. Blackwell Reference Online. 31 December 2007 <http://www.blackwellreference.com/subscriber/tocnode?id=g9780631201267_chunk_g978063120126720>

Bibliographic Details

The Handbook of Phonological Theory

Edited by: John A. Glodsmith
eISBN: 9780631201267
Print publication date: 1996