20. Phonological Derivation in Early Generative Phonology

Michael J. Kenstowicz and Charles W. Kisseberth

20.1 Introduction

As articulated in *The Sound Pattern of English* (*SPE*, Chomsky & Halle 1968), a generative grammar is a formal system of rules and constraints that assigns a phonetic realization to an unbounded number of structured expressions built by the syntax from a stock of lexical items. The grammar is also a characterization of the tacit knowledge of the native speaker and thus has psychological reality. As discussed in Kenstowicz (this volume), the early generative model represents sounds at both the underlying and the surface phonetic levels by a universal set of distinctive features that define the inventory of segments encoding the lexical representations and the natural classes that appear in the phonological rules and constraints. The features also characterize the grammatically determined aspects of the phonetic realization of an expression in speech and are not responsible for speaker-specific properties such as gender and age or modifications reflecting speech rate, style, etc. In the *SPE* model, the surface form of a word is derived from its underlying form by extrinsically ordered rules without any explicit recognition of an intermediate level of representation such as the post-Bloomfieldian phoneme (see Ladd, this volume) that was intended to record surface segmental contrasts such as reflected by minimal pairs. An evaluation metric based on formal economy favors grammars in which the surface alternants of a lexical item are derived from a single underlying form.

As this methodology was applied to the analysis of more and more languages, certain fundamental questions arose that will be reviewed in this chapter. First, are there any limitations on the ways in which the underlying representation of a lexical item can diverge from its various surface phonetic realizations? Given that the grammar has psychological reality, what counts as a valid descriptive generalization that could plausibly be learned by a child acquiring the language? Second, can and should a distinction be drawn between morphophonemic rules that map between phonemic segments and are often restricted by the morphological or lexical context vis-à-vis more superficial rules that introduce allophones and typically enforce restrictions on what the native speaker can pronounce? Third, given that the grammar contains multiple rules, how do they interact? Do they apply simultaneously, or in sequence? Are some ordering relations more optimal? Fourth, in the *SPE* model each rule applies based solely on information contained in the output of the immediately preceding rule. Are there linguistically significant generalizations that require a rule to access more remote stages of the derivation such as the input? Fifth, does it make any sense to say that a rule applies, either singly or in concert with other rules, to achieve a particular output target? Finally, is the application of a rule dependent on information contained in another (related) word?

In this chapter, we review some of the major discussions of these questions in the decade following the publication of *SPE*.

20.2 The issue of abstractness

In a general sense, any grammatical statement is an abstraction over the observable input data and any sound segment (phoneme) is a category that abstracts over the factors of co-

articulation and the gradience of speech. The issue discussed here is a more technical one and was first raised by Paul Kiparsky (1968) in an unpublished but widely read and debated paper. Kiparsky called into question the practice found in many applications of the *SPE* model of what he termed the 'diacritic use of phonological features' employing rules of 'absolute neutralization' (see Halle 2019: 4–5 for a recent example). This state of affairs is typically the synchronic residue of an alternation that has been obscured by the merger of a contrast present at an earlier stage of the language. For example, in the development of Sanskrit, velars palatalized before front vowels, including [e]. But at a later stage, [e] was merged to [a] resulting in a synchronic situation in which some surface [a]'s trigger palatalization while others (those deriving from earlier *[a]) do not. In the *SPE* model, it is possible to set up an underlying phonological contrast between /e/ and /a/, define the palatalization rule over the abstract /e/, and then posit a later context-free rule that merges the /e/ to /a/ everywhere, as in the schematic derivation in (1).

(1)

/k-e/	/k-a/	
č-e	—	Palatalization
č-a	—	[–high] → [+low]

The alternative 'concrete' analysis marks the [a]'s that fail to trigger the rule with an exception feature [–Palatalization]. In the *SPE* model, where rules are the basic engine of the grammar, arbitrary lexical exceptions are inherently disfavored and so the evaluation measure would select the abstract analysis. Based on his study of linguistic change, Kiparsky claimed that rules of absolute neutralization such as the /e/ to [a] merger of Sanskrit introduce an inherent instability into the grammar and later stages of the

language will tend to either curtail or overgeneralize the alternation. This regularization can be understood as following from a presumed bias to minimize lexical exceptions and hence presupposes that rule features are less highly valued than phonological features.

Another relevant point is that selection of the feature to distinguish the merged segments is often arbitrary. Analysts typically resolved the choice based on their knowledge of the history of the language. For example, in Spanish some mid vowels diphthongize under stress while others are stable: cf. infinitive and 1 sg. present tense forms *neˈgar*, *ˈniego* 'deny' vs. *deˈber*, *ˈdebo* 'owe'. In an abstract analysis, a phonological difference between the two types of mid vowels would need to be postulated. We could appeal to a feature like [nasal]—a feature that is not employed for phonemic vowel contrasts in Spanish and hence is available to differentiate the two classes of mid vowels. But why would Spanish speakers deduce that nasality is at work here? Length/tenseness might seem a better choice on markedness grounds, leading to a rule that would diphthongize long or tense vowels. But in actuality, the diphthongizing vowels derive from Proto-Romance short lax vowels.

The argument against absolute neutralization was mounted chiefly on largely untested assumptions about learnability and an assumed bias to posit URs that depart from the surface form only if the evidence warrants. For example, the German data in (2) from Wiese (1996) motivate a rule of Final Devoicing (FD) for obstruents.

(2)	a.	Lo[p]	Lo[b]-es	'praise'
		Ra[t]	Ra[d]-es	'wheel'
		Ta[k]	Ta[g]-es	'day'

b.	Bla[t]	Bla[t]-er	‘leaf’
	Sac[k]	Sac[k]-es	‘bag’

The FD rule allows the voicing alternation in each row of (2a) to be generated from a single lexical form /lob/, /Rad/, and /tag/, next to /blat/ and /zak/ in (2b). Under an alternative analysis that lists both alternants (/lop, lob/, /Rat, Rad/, /tak, tag/), we still need a rule to choose between them in the appropriate contexts; so, for example, the computation of output [k] for *Ta*[k] is stated twice: once in the listed alternant and a second time in the allomorph selection rule. Furthermore, a rule selecting among listed allomorphs fails to generalize to new words such as ‘orange’ [oːRaŋʃ], [oːRaŋʒə]. The analysis that posits /tag/ and FD is simpler in terms of lexical storage, states the devoicing just once, and generalizes to novel words. But what about adverbs such as [avek] ‘away’ that lack an inflection? They must have a surface voiceless stop, given FD. But there are two possible underlying forms: /avek/ or /aveg/. Kiparsky points to the dialect of Yiddish (Sapir 1915) that lost the final devoicing rule so that earlier *Ta*[k] appears as [tog] while *Sac*[k] surfaces with a voiceless final consonant [zak]. Crucially, nonalternating [avek] has a voiceless consonant [avek] in Yiddish. Kiparsky interprets the Yiddish innovation as evidence for an Alternation Condition that warrants positing an underlying representation different from the surface form only if motivated by alternations. In the absence of such evidence, the learner will posit a UR that is identical with the surface form (while allowing for the insertion of redundant features by automatic allophonic rules). It should be noted that Yiddish has a related word *Vek*, *Veg-e* ‘way’ that does exhibit the voicing alternation and hence poses the interesting question of how close two lexical items must be in order to activate the Alternation Condition (King 1980). The

choice of underlying /avek/ could also be motivated by an economy principle that strives to minimize the length of derivations—a line of reasoning that anticipates the OT notion of Faithfulness (Prince & Smolensky 2004).

20.3 The challenge from Yawelmani (Yowlumne)

Kiparsky's proposal that rules of absolute neutralization introduce instability and a more complex state of affairs that should be disfavored by the evaluation measure was immediately challenged by studies demonstrating that unlike in the Sanskrit case, a language might have multiple lines of evidence that converge on the same underlying abstract distinction. The best known and widely cited example is Kisseberth's (1970) discussion of the Yawelmani dialect of Yokuts. (Hyman 1970 and Brame 1972 voiced similar objections to Kiparsky's Alternation Condition as did various later publications such as Kaye 1980 and Halle & Clements 1983).

Building on Kuroda's (1968) analysis of the materials in Newman's (1944) grammar, Kisseberth focused on the vowel system of the language. Glossing over certain complications (see below), the surface inventory of Yawelmani is shown in (3).

(3)

i	u			
(e)	o		e:	o:
	a			a:

The [e] arises from a general rule shortening long vowels in closed syllables and hence can be removed from the phonemic inventory; this leaves four short vowels and three nonhigh long vowels. The language has a vowel harmony system in which the rounding

value for suffixes depends on the root; moreover, rounding from the stem is transmitted to the suffix only if the vowels agree in height (4a).

(4)

	nonfuture	dubitative	
a.	max-hin	max-al	‘procure’
	gop-hin	gop-ol	‘care for an infant’
	xil-hin	xil-al	‘tangle’
	hud-hun	hud-al	‘recognize’
b.	lan-hin	laːn-al	‘hear’
	dos-hin	doːs-ol	‘report’
	mek’-hin	meːk’-al	‘swallow’
	c’om-hun	c’oːm-al	‘destroy’
	ʔoṭ’-hun	ʔoːṭ’-al	‘steal’

The abstractness issue arises in the behavior of roots with [oː] such as ‘destroy’ and ‘steal’ in (4b). They are not mere exceptions to rounding harmony but condition the alternation in an ‘upside-down’ fashion where the surface nonhigh stem vowel is associated with the rounding of a high suffixal vowel (*c’om-hun*) and with no rounding of a nonhigh suffixal vowel (*c’oːm-al*) . Hence, positing a rule feature is not a straightforward option. A simpler and more insightful analysis postulates an underlying distinction in height and derives the roots in ‘destroy’ and ‘steal’ from underlying high vowels based on a rule of absolute neutralization that lowers [+high, +long] vowels to [–high] in a context-free manner, as shown in (5).

(5)	/doːs-hin/	/doːs-al/	/c'uːm-hin/	/c'uːm-al/	
	——	doːs-ol	c'uːm-hun	——	Vowel Harmony
	——	——	c'oːm-hun	c'oːm-al	Vowel Lowering
	dos-hin	——	c'om-hun	——	Closed Syllable Shortening

As Kisseberth observes, there is considerable independent evidence for this analysis. First, the lowering rule can be generalized to derive surface [eː] from /iː/. This makes the underlying vowel inventory symmetrical with three binary features distinguishing eight vowels (Clements' 1992 feature economy). Second, more evidence is provided by a class of forms with the canonical shape CVCVːC that Newman termed 'echo roots', illustrated in (6).

(6)	p'axat'-hin	p'axaːt'-al	'mourn'
	hiwet-hin	hiweːt-al	'walk'
	yolow-hin	yoloːw-ol	'assemble'
	sudok-hun	sudoːk-al	'remove'

In these stems, the [oː] that triggers rounding in low vowels is paired with short [o] while the [oː] that triggers rounding in high vowels is paired with [u]. Finally, [eː] is paired with [i], and [aː] with [a]. If [eː] and the harmonically 'exceptional' [oː] are derived from high vowels then the underlying structure of this stem class is simple: the vowels are identical in quality. Thus, on grounds of formal economy as well as descriptive insight, the best analysis is the abstract one that entails a rule of absolute neutralization. In later discussion, Kiparsky (1973) concedes that the abstract analysis is the better one for Yawelmani and states that his proposal to ban absolute neutralization was intended to

exclude cases where the evidence for the underlying contrast is more meager, as in Sanskrit.

The Yawelmani data made their way into several textbooks to illustrate the abstractness question (Kenstowicz & Kisseberth 1979, Kenstowicz 1994, Odden 2003). They were also discussed by the American Structuralist Charles Hockett (1967, 1972),[1] who proposed an analysis with ordered rules that starts with bases containing long high vowels and a vowel lowering rule (his rule 5) that does much of the same work as the analyses of Kuroda and Kisseberth. At the end of his 1967 article, Hockett contrasts phonemic vs. morphophonemic phonology. The phonemic transcription represents 'articulatory targets' while the morphophonemics has no such 'direct tie to reality' (indeed, in Hockett's analysis the underlying representations contain arbitrary symbols to trigger phonological changes) and is proposed for 'descriptive convenience'. Hockett suggests that a more psychologically realistic analysis generates outputs by analogy to a stored paradigm of one prototype verb for each prosodic stem shape. While acknowledging that this alternative description occupies much more space than the system with underlying forms and morphophonemic rules, he argues in its favor on the grounds that there would be 'a net gain in realism'. Here we see laid bare the skepticism shared by many linguists concerning abstractness and the intuition that phonological derivations are not all fashioned from the same cloth, as in *SPE*, but comprise two distinct systems—an intuition that runs through the field from Baudouin de Courtenay (Radwanska-Williams, this volume) until the current day (cf. Liberman 2017). In her

[1] See Ladd (this volume) for discussion of Hockett's approach.

incisive commentary on Yawelmani, Blevins (2004) notes that the abstract analysis of Kuroda-Kisseberth oversimplifies the situation in that there are certain verbal forms with long high vowels that are realized on the surface (reflecting compensatory lengthening from loss of a glottal stop). In addition, Spanish loanwords contain them: *mušgiːda'* 'mosquito' and *'uːbaš* 'grape' (< Sp. *uvas*). In view of these considerations, she endorses the more concrete analyses of Newman and Hockett's 'realistic' paradigmatic alternative against the Kuroda-Kisseberth (and morphophonemic Hockett) ones with long high vowels that are always mapped to either [–long] or [+high]. Combining these features in the same underlying phoneme is an unjustified abstraction, in Blevins' view.

This point was also taken up by Hansson & Sprouse (2002). Working with two of the last surviving native speakers of the language, they find that even in these challenging circumstances with few available interlocutors the opaque alternations largely survive with no wholesale simplification. In particular, high vowel suffixes such as the aorist *-hIn* retain the opaque harmony alternations with rounding after an [o:] that derives from [u:], as in /ʔuːṭ'-hIn/ > *ʔoṭ-hun* 'steal', versus no rounding after an underlying nonhigh vowel, as in /won-hIn/ > *won-hin* 'hide'.[2] Low vowel suffixes retain the opaque alternation when they also impose a particular prosodic template on the stem. So for example, the durative passive aorist /-ʔAt/ requires the preceding verb stem to take an 'iambic' CVCVV(C) shape. This suffix also distinguishes the two types of surface low round vowels: /loʒoːx-ʔAt/ > *loʒox-ʔot* 'frighten' vs. /ṣuxuːṭ-ʔAt/ > *ṣuxoṭ-ʔat* 'wash'. However, when a low

[2] In these examples *I* represents a high vowel that alternates between *i* and *u*, and *A* represents a low vowel that can appear as *a* or *o*.

vowel suffix is not associated with a prosodic subcategorization of the stem, then Hansson & Sprouse report that depending on the particular speaker the a≈o alternation has either been curtailed or is governed by the surface form of the stem and does not distinguish [o:] from /u:/ vs. [o:] from /o:/. The latter behavior is shown by the dubitative suffix *–Al*: /lɔʒo:x-Al/ > *lɔʒo:x-ol* 'might frighten' and /hulu:ʂ-Al/ > *huloo*ʂ*-ol* 'might sit down'. In sum, it appears that when the opacity induced by the lowering of high vowels involves a counter-bleeding relation with harmony (as it does with the high vowel suffixes) then the alternation is readily retained. But when lowering entails a counter-feeding relation with harmony then the alternation has been reanalyzed in terms of the surface form unless it is also supported by a morphological allomorphy selection process. These data suggest that a counter-feeding relation is more challenging to learn compared to counter-bleeding. The former state of affairs presents a surface contradiction to the postulated rule or constraint while the latter does not directly contradict the hypothesis but requires the learner to extend the range of factors that trigger the rule. See Prickett (2019) for an artificial grammar learning experiment that also finds counter-feeding to be more challenging than counter-bleeding.

In conclusion, it appears that even in a state of extreme duress with few native speakers and a largely unsupportive multilingual environment, opaque analyses can still be learned. A fortiori the same point holds for the Vowel Shift and Velar Softening alternations of English and other well-known examples such as the Slavic yers, alternations which have remained largely stable for centuries in terms of both their structural relationships with other alternations as well as in terms of the particular lexical items that display them. It is worth pondering the question of why the alternations

underlying such abstract analyses persist in the intricate morphophonemic systems of these and countless other languages?

20.4 Two-stage models

Two distinct schools of thought developed in reaction to the *SPE* model's basic premise that the underlying representation is mapped to the surface phonetic form with no intermediate level of representation comparable to the structuralist morphophonemic vs. allophonic break. Both tried to restore this distinction and adopted the moniker 'natural', seeing the lower-level allophonic rules as fundamentally distinct from the deeper morphophonology.[3] But as in *SPE*, sounds are represented as feature matrixes in both modules and eschew the nonphonetic arbitrary morphophonemes found in many structuralist analyses.

20.4.1 Natural Phonology

Stampe (1973) and especially Donegan & Stampe (1979) mounted an aggressive critique of both structuralist and early generative approaches to phonology and advocated a return to the conceptions of sound structure espoused by Maurice Grammont, Edward Sapir, and other linguists who flourished at the beginning of the 20th century. In the theory of Natural Phonology, David Stampe and his followers draw a sharp distinction between

---

[3] However, in view of Halle's (1959) famous argument from Russian voicing assimilation against American Structuralism (Dresher & Hall, this volume), the distinction between morphophonology and phonology was drawn differently.

natural ‘processes’, which substitute one sound for another in the interests of downstream phonetics, versus more arbitrary sound changes (which they call ‘rules’) such as the Velar Softening of *electric-electricity*.[4] The latter are ‘cognitive’ operations that speakers are required to learn in order to speak their language properly but have the status of arbitrary conventions bequeathed by the history of the language. They are often utilized to express paradigmatic grammatical distinctions, as in German umlaut. The failure to differentiate the two kinds of sound substitutions is a fundamental category error. For Stampe, the natural processes are innate; what must be learned for a given language is their presence or absence, their scope, and their interaction. The natural processes enforce limitations on what the native speaker can pronounce and are largely unconscious while rules may have exceptions and idiosyncratic lexical restrictions reflecting their diachronic origin. The natural processes are revealed in the allophony of a language as well as in alternations between casual and emphatic speech. Their productivity is evident in that they govern the output of language games, judgments of rhyme, loanword adaptations, and foreign accents. From this naturalistic perspective, *SPE*’s concern with formal economy is misplaced just as it would be in trying to understand the properties of other bodily mechanisms such as the visual or digestive systems.

For Stampe, a phoneme is a phonetically natural object. He critiques the structuralists’ archiphoneme (see Battistella, this volume), emphasizing that in a context of neutralization such as for the voicedness of stops in the word-initial *s*C clusters of English, what appears is a fully specified segment (what many would characterize as the

[4] For more on Natural Phonology, see Kenstowicz, Scheer, and Calabrese, all this volume.

maximally unmarked voiceless unaspirated stop) rather than a sound unspecified for voicing. This is the sound that emerges when the cluster is simplified, as in French *école* 'school' (cf. Spanish *escuela*). Whereas *SPE* chose to represent sounds underlyingly in the most economical way consistent with being able to predict the correct surface realization (hence the appeal to unspecified feature values), Stampe chooses the most 'natural' specification. While this proposal works well for many features, there are others such as [±back] or [±high] where it is unclear which values underlie the suffixal vowels of Turkish and Hungarian or, for that matter, the schwa of English. Another problem is that the sound substitutions the child confronts do not come labelled as 'process' vs. 'rule'. What information resolves the choice? Why would the [ɨ] ≈ Ø alternation seen in the English plural (as in *bushes* [bʊšɨz] ≈ *bulls* [bʊlz]) be treated as a process rather than a rule? This is especially problematic since most rules have their origins in earlier processes.

While the teleology of natural processes is rooted in considerations of ease of articulation and perceptual saliency, the discussion in Donegan & Stampe (1979) takes an introspective 'bird's eye' view of phonetics: rarely are results from experimental phonetics cited in support of the analyses, and the natural processes themselves are never formalized. Many of the examples described as segment substitutions would now be more properly treated as articulatory shortcuts which, while planned and hence 'mental', operate at the level of gradient phonetic implementation rather than segment-level feature changes.

In spite of these limitations, the Natural Phonology critique of the *SPE* model attracted considerable attention and admiration and motivated empirical research on the

phonetic foundations of phonology in the Laboratory Phonology framework that developed in the following decade. And with its conception of the lower-level phonology of a language as the resolution of the conflicting forces of articulatory ease and paradigmatic contrast, Natural Phonology set the stage for the formal expression of this intuition in the OT model of Prince & Smolensky (2004).

20.4.2 Natural Generative Phonology

A distinct critique of the *SPE* program was mounted by Theo Vennemann, Joan Hooper, and their followers under the label Natural Generative Phonology (NGP).[5] Like Stampe, NGP also distinguishes between the phonetically motivated allophonic rules vs. morphophonemic rules. But unlike Stampe, Hooper and Vennemann professed a deep skepticism of grammatical statements that did not state generalizations over 'overt' phonetic representations. Hence, a Stampean natural process converting /plænt/ to [plæ̃t] would be excluded since the nasal consonant that is the source of nasality in the vowel is not present in the phonetic surface form. Taking the point of view of an inductive learner, rules of both types were required to state generalizations over a language's surface representations (Hooper's True Generalization Condition). The descriptive generalizations expressed by opaque rule ordering (see section 20.5) in the *SPE* model were recast as statements that take into account the morphological and lexical contexts in which the sounds were assumed to be embedded.

[5] For further discussion see Scheer, this volume.

The role of the True Generalization Condition is nicely illustrated by Hooper's (1976) discussion of the Granada dialect of Castilian Spanish (based on material in Alonso et al. 1950) where vowels are laxed in closed syllables and /s/ is lenited to aspiration or deleted in the syllable coda. The rules apply in counter-bleeding fashion (see section 20.5) in other Castilian dialects: *voz* 'voice', singular [bo̞^{h}], plural [bose̞^{h}]. The laxing has given rise to a harmony in Granada seen in *pedazo* 'piece', singular [peðaθo] vs. plural [pe̞ða̞θo̞]. Hooper takes this alternation as evidence that the laxing rule has been reanalyzed into a morphological process marking the plural as distinct from the singular rather than remaining as a phonological rule conditioned by an underlying /s/. This reanalysis would be required by the True Generalization Condition when the /s/ was no longer reflected in aspiration. The rule would then assign [−tense] in the morphological context X____Y#]$_{\text{noun, adj [+plural]}}$. Some independent support for this concrete analysis derives from Spanish words designating days of the week that have invariant /s/: *el martes* 'Tuesday', *los martes* 'on Tuesdays'. In the Granada materials from Alonso et al. (1950), these expressions are distinguished by laxing as singular [marte] vs. plural [ma̞rte̞]. Under a purely phonological analysis with underlying /s/ such a distinction is mysterious. But it follows nicely if the laxing rule is stated in terms of the plural grammatical context without regard to the presence of /s/ or /h/. Hooper does not indicate the status of (historically) *s*-final words like *voz* in this dialect. If the laxing functions to mark singular vs. plural, we would expect the laxing to be lost in the singular, just as in *el martes*. Another factor cited in support of the concrete analysis is that according to the source, laxing in Granada is greater in magnitude compared to Castilian. This would

make sense if the tense-lax contrast has taken on semantic weight compared to its subconscious allophonic status in Castilian.

A recent experimental phonetic investigation of the Salamanca and Granada dialects of Castilian Spanish by Henriksen (2017) fails to support the major claims of the concrete analysis. Paradigms were collected from twelve speakers of each dialect that involved four word conditions: (1) singular, vowel-final (e.g., *nene* 'boy'); (2) bare plural, no article (e.g., *nene-s* 'boys'); (3) plural, with articles (e.g., *los nene-s* 'the boys'); and (4) singular, /s/-final (e.g., *jueves* 'Thursday'). Granada speakers evidenced laxing in conditions (2)–(4) (*nene-s*, *los nenes*, and *jueves*) compared to *nene* while Salamanca speakers did not, confirming the dialect difference. However, for the Granada speakers there was no significant difference in laxing among the three conditions (2)–(4). Indeed, in cases like *jueves* F2 shows that there is greater laxing. Henriksen concludes that at least for his Granada speakers 'laxing is not motivated by functional considerations'.

20.5 Natural rule order

In the *SPE* model the derivation works much like an assembly line in an automobile factory. A given rule applies just once per cycle and its application depends solely on the information passed on from the immediately preceding rule. In the post-*SPE* period, various aspects of this list-like conception were called into question based on both empirical and theoretical considerations. Most of the subsequent discussion assumed that the *SPE* list could be broken down into binary precedence relations between pairs of rules: either A precedes B, B precedes A, or both orders are possible with no material effect on the output. Questions arose as to whether the ordering relation was totally free

or could be predicted based on the form of the rules. And if not totally predictable, then are certain ordering relations more expected than others? We review some of the major discussions in this section.

Kiparsky (1968) introduced the concepts of 'feeding' and 'bleeding' to describe functional relations between a pair of rules. Rule A (potentially) *feeds* B if the application of A creates novel inputs to B. If A is ordered before B then both rules will apply and A is said to (actually) feed B. If B is ordered before A then B will not apply, given that the change making B applicable only occurs later in the derivation and that B cannot anticipate this change. This is *counter-feeding* order. In a (potentially) *bleeding* relation, the application of a rule A removes inputs to B. If A applies before B, then B will not apply and A is said to (actually) bleed B. But if B is ordered first, then both rules apply (*counter-bleeding*).

Kiparsky (1965, 1968) identified a number of diachronic changes that could be characterized as a switch in the order of rules (following Halle 1962). The directions of change in these cases were from counter-feeding to feeding or from bleeding to counter-bleeding. In each case the innovative order involved the application of both rules and so a principle of Maximal Utilization was proposed as a clause in the evaluation measure to enforce these orderings as optimal and hence motivate the diachronic innovations. Kenstowicz & Kisseberth (1971) documented a number of examples where a bleeding relation between rules was imposed and on these grounds questioned the validity of Maximal Utilization. One of their general types is vowel epenthesis into consonant clusters that are also subject to a local assimilation. For example, Maximal Utilization predicts that the rule of progressive voicing assimilation seen in English *cat-s*, with [s]

derived from /z/ (compare *dog-s* with [z]), should apply in the derivation of stems ending in a sibilant that trigger epenthesis so that the plural of *bush* /bʊš-z/ should map to [bʊšis] instead of the correct [bʊšiz]. But from a functional perspective, epenthesis separates the two consonants and thus removes the motivation for voicing assimilation. Derivational economy would then favor the minimal change enforced by bleeding.

In a subsequent study, Kiparsky (1971) revised his theory of natural rule ordering. Feeding was still regarded as more optimal than counter-feeding; but bleeding was reinterpreted as unmarked compared to counter-bleeding. The diachronic innovations that seemed to motivate a change from bleeding to counter-bleeding could better be explained as instances of analogical leveling. More importantly, Kiparsky sought the motivation for the natural ordering of the rules in learnability. In a counter-feeding relation, the earlier rule has superficial exceptions created by the later rule; its generality can only be discerned by undoing the effects of the later rule. An example is the interaction of Yawelmani Vowel Harmony and Vowel Lowering in the derivation of *c'oːm-al* shown in (5) above: Vowel Lowering creates a surface exception to Vowel Harmony. Similarly, in a counter-bleeding relation such as *writer-rider* (*wr*[ʌjɾ]*er-r*[aːjɾ]*er*), the Canadian dialect's otherwise unmotivated raised diphthong [ʌj] before a voiced sonorant [ɾ] can only be explained by undoing the effects of the flapping rule that merges the underlying voicing contrast in the medial stops (cf. *write* [t] vs. *ride* [d]).[6] Thus, in both counter-feeding and counter-bleeding, the later rule masks the conditions of application of the

[6] See Ladd (this volume) for discussion of the difficulties that English flapping created for the post-Bloomfieldians.

earlier rule; in Kiparsky's terms, the later rule makes the earlier one *opaque*. Kiparsky reasoned that such 'opaque' relations present a significant challenge to the language learner and should be evaluated as more complex and disfavored.

However, shortly thereafter Kisseberth (1973b) and Kaye (1974) observed that while an opaque rule order may make the masked rule harder to learn, it also provides clues to the underlying representation—an equally important factor in the acquisition of phonological alternations. For example, given that outside of the flapping context the [ʌj] diphthong is a variant of [aːj] before a voiceless consonant (as in *type, tight, bike*), the [ʌj] in *writer* is prima facie evidence that the flap derives from an underlying voiceless consonant. If the rules were to apply in the opposite order, with flapping bleeding raising, then *writer* would be pronounced like *rider*, further obscuring the underlying distinction between /t/ and /d/. The relative rarity of opaque derivations such as *bushes* /bʊš-z/ > *[bʊšis] where the opacity does not aid in the recovery of the underlying form was argued to support this assertion.

The concepts of feeding, bleeding, and opacity introduced by Kiparsky entered the lexicon of generative phonology and are used to describe and analyze phonological patterns even in models of grammar that no longer employ rule ordering or even rules (e.g. Baković 2007). Opacity remains a subject of intense interest since it presents a direct challenge to classical Optimality Theory with its one-step mapping between the input and output, as discussed by Idsardi (1998, 2000), Vaux (2008), and more recently Buccola (2013) and O'Hara (2017).[7]

[7] See van Oostendorp (this volume) for further discussion of this issue.

Based on a survey of some of the earliest generative analyses, Koutsoudas, Sanders, & Noll (KSN; 1971, 1974) suggested that rule ordering was largely predictable and proposed what they felt to be the conceptually simplest mode of derivation: rules freely apply whenever their structural description is met. This theory is compatible with feeding and counter-bleeding orders. It predicts free reapplication and so derives 'anywhere rules' (Chafe 1967), which were claimed to apply at multiple points in a single derivation. A bleeding relation is problematic for simultaneous application, however, since the structural description of both rules is satisfied yet only one applies. To accommodate such cases, KSN proposed a Proper Inclusion Precedence Principle according to which rule A will take precedence over rule B when the structural description of A properly includes the structural description of B. This principle can properly analyze the depalatalization of /lʲ/ and otherwise loss of laterality in Latin American [akel] 'that' ≈ [akejos] 'those' from /akelʲ/ discussed by Saporta (1965), given that depalatalization is restricted to the syllable coda while delateralization is context-free. Thus, only delateralization applies in /akelʲos/, where /lʲ/ is in a syllable onset; and where both rules could potentially apply to /akelʲ/, the more specific one, depalatalization, takes precedence. However, there are many other cases of bleeding that do not fall under this rubric such as the derivation of English *bushes* mentioned above or Donegan & Stampe's (1979) optional vocalization of pretonic sonorants (*pray* > [pɪrɛj]), which bleeds rhotic flapping after tautosyllabic [θ], so that *three* is realized as [θɨri] rather than [θɨɾi].

Anderson's (1969, 1974) theory of Local Ordering took issue with the *SPE* assumption that the order between two rules was fixed for the grammar as a whole. His clearest example is from Icelandic where a rule that umlauts /a/ to [ö] when the next

syllable contains /u/ must follow syncope in forms such as *alin* (/alin/) ‘ell of cloth’ and its dative plural /alin-um/ > [öln-um], but must precede syncope in *böggull, bögg-l-i* ‘parcel’ from /bagg-ul/, /bagg-ul-i/, respectively; cf. *baggi* ‘pack’. Anderson acknowledges that a cyclic analysis of these data would permit a uniform ‘syncope precedes umlaut’ order to be imposed but rejects this alternative on the grounds that cyclic rules are the provenance of stress rules rather than segmental phonology. This rationale was not accepted by most researchers. Indeed, in the Lexical Phonology model developed by Kiparsky (1982) in the following decade (see Scheer, this volume), all morphophonemic rules were regarded as cyclic in virtue of their interleaving with the rules of affixation in a level-ordered morphology and various formal properties of phonological rules were tied to their [±cyclic] status.

Donegan & Stampe (1979) make a number of proposals for how their natural processes interact. First, given their phonetic teleology, the processes necessarily follow the morphophonemic rules. Second, the natural processes do not apply in a fixed order and may freely reapply to create feeding relations. However, Donegan & Stampe also suggest that a counter-feeding order may arise by taking the option of restricting a process to the earliest input form. This proposal is motivated by their judgment that while flapping applies obligatorily in *pat#it*, speakers divide into two groups with respect to the intervocalic /t/ arising from nasal deletion in *plant#it* : ‘noniterating’ speakers fail to flap in the later context while iterating speakers freely flap in both cases. Third, Donegan & Stampe speculate that many bleeding relations fall under a general schema in which fortition processes precede lenitions. Assuming that epenthesis is a fortition, this accounts for the English plural example of *bushes* [bʊšiz] mentioned above. Finally, both ordering

principles might be called on to explain the intricate intervocalic spirantization seen in the following Somali data (Kenstowicz 1994: 129). Underlying /gabḍ-ta/ *girl*-DEF 'the girl' is realized as [gaβaḍa]: epenthesis into the underlying triconsonantal cluster gives /gabaḍ-ta/, followed by lenition to /gaβaḍ-ta/, and then deletion of the suffixal /t/ (perhaps via assimilation and degemination) without reapplication of spirantization. The data would also be compatible with a simultaneous one-step move along a lenition chain: geminate > singleton > spirant. On this scenario singleton /b/ lenites to the spirant [β], while /ḍt/, assimilated to /ḍḍ/, lenites just one step to singleton [ḍ] without proceeding further to a spirant.

20.6 Conspiracies, (derivational and surface-structure) constraints, and global rules

Kiparsky's 1968 paper 'How abstract is phonology?' raised issues that dominated the 1970s but persist to this day. Kisseberth's 1970 paper 'On the functional unity of phonological rules' followed shortly after *SPE* and raised issues that likewise reverberated throughout the 1970s. Yet its impact was not realized fully until the advent of Optimality Theory, more than two decades later.

20.6.1 Conspiracies

The essential justification for each element of the *SPE* model was the expression of 'linguistically significant generalizations'; the failure to express a perceived generalization was viewed as a defect in the model's architecture. The fundamental issue raised by this methodology, of course, is whether some missed generalization is in fact

'linguistically significant'. Linguists can disagree radically on this point (as seen in the earlier discussion of the Yawelmani vowel alternations).

Kisseberth (1970) argued that the *SPE* view of rules (where a rule consists of a structural description, a structural index, and a structural change) and how they interact often fails to express the 'functional unity' underlying disparate 'rules' found in a single language. For example, in Yawelmani multiple rules function to avoid sequences of the shape CC{C, #} (recall that at this point in generative phonology, syllables were not part of the theory, hence the absence of a reference to syllable structure in this example). In Yawelmani, the combination of morphemes often creates these disallowed sound sequences and the language avoids them by means of vowel epenthesis or, in certain morphologically defined contexts, by consonant deletion. Furthermore, Yawelmani has certain vowel deletion rules that are blocked from applying if their output would be one of these unacceptable sequences. Drawing on parallel findings in syntax (cf. the constraints of Ross 1967 and the surface filters of Perlmutter 1968), Kisseberth (1970) referred to such situations as 'conspiracies'.[8]

---

[8] The term 'conspiracy' was used extensively in classes taught by J. R. 'Haj' Ross and George Lakoff at the 1968 LSA Institute at Illinois and during the following academic year (1968-69) that Kisseberth spent in Cambridge attending classes at MIT (Ross), Harvard (Lakoff), and Brandeis (Perlmutter). These researchers saw parallelisms between Kisseberth's work and what they were doing.

Let us refer to the Yawelmani conspiracy as the *CC{C, #} conspiracy. The essential point of Kisseberth's discussion is that the device of abbreviatory conventions that *SPE* relies on to express the 'sameness' of two or more rules by compiling them into a single statement (the evaluation measure based on formal economy) is unable to collapse together rules inserting vowels, deleting consonants, and failing to delete vowels if such a deletion would lead to violations of the *CC{C, #} conspiracy. The 'sameness' that underlies the rules in a conspiracy could not be expressed by means of such notational devices.

20.6.2 Constraints

Kisseberth did not attempt a full solution to the problem posed by the missed generalization he identified. In fairness, the absence of a real solution is not surprising. *SPE* had only just appeared and the problem identified was a fundamental one, albeit one not necessarily recognized as such at the time. The 1970 paper did make one suggestion that had some traction in the 1970s. Specifically, Kisseberth suggested that *CC(C, #) was a *constraint* on derivations that could *interact* with rules deleting vowels such that a vowel deletion rule would be blocked from applying not by virtue of a failure of an input structure to satisfy the structural description of the rule, but rather by virtue of a hypothetical output violating the *CC{C, #} constraint. In this approach, like a move in chess, rules can be required to 'look ahead' to determine whether their output is acceptable in terms of a specific constraint.

Kisseberth himself did not require the constraint underlying a conspiracy such as *CC{C, #} to necessarily hold of surface structure, since he allowed for the possibility of

constraints that 'low-level' or 'fast-speech' rules in the language might actually violate. (Compare consonant clusters that are 'impermissible' in English but nevertheless do occur when unstressed vowels in words such as *potato* are syncopated in casual speech.) On the other hand, Shibatani (1973) took precisely this step with the proposal that phonological grammars contain a set of constraints that define the permissible sound sequences at the surface phonetic level. He allows positive constraints, which require that phonetic form have a certain structure (e.g. $\#(CV)_1\#$, which requires a word to consist of one or more sequences of CV). Shibatani also allows 'implicational' or 'if-then' constraints (e.g. in a language which allows [ma], **[**bã], **[**ba], but not ***[**mã], there is a constraint that says that if a segment precedes or follows a nasal segment, then it must be oral). Finally, he allows 'negative' constraints (e.g. an alternative to characterize the [ma], **[**bã], **[**ba], but not ***[**mã] facts would be a negative constraint barring successive nasal sounds: *[+nasal][+nasal]).

The idea of 'surface structure constraints' had been rejected in the earliest stages of generative phonology largely on the grounds that they are redundant; while such constraints might be correct descriptions of surface structure, they play no role in the actual derivation of surface forms from underlying forms. For example, Postal (1968: 214) writes: 'an independent phonotactics is necessarily and in all cases useless and redundant in its entirety. . . [because] . . . every fact which such a separate phonotactics describes is accounted for . . . by the morpheme-internal restrictions on morphophoneme combinations [MSCs] and the morphophonemic rules which must exist in any event.' And Johnson (1972:15) states: 'there seems to be no need at all for a special component to describe the set of admissible phonetic strings, since this set is determined indirectly

by the morpheme structure component and the phonological rules.' Shibatani's (1973) argument against this position is that the ultimate goal of phonological theory is to characterize a speaker's knowledge of their language and not just to derive surface forms from underlying representations. He proposes that knowledge of such constraints can be observed in the course of language acquisition even before the child has worked out the underlying representations and morphophonemic alternations associated with morphologically complex forms. Knowledge of surface phonetic constraints can also be observed in loanword adaptation, even in the absence of morphophonemic processes that enforce the surface regularity in the native grammar. For example, in many Bantu languages all words end in a vowel, but there is no morphophonemic rule that introduces a vowel at the end of a word (because there are no underlying representations of words where the word ends in a consonant). Nevertheless, a Bantu language like Chimwiini has borrowed countless words from Arabic and Somali that end in a consonant, but regularly introduces a vowel after the final consonant, as in *farasi* < Ar. *faras* 'horse'. A third source of evidence for the psychological reality of surface phonetic constraints is that their effects can be observed in the course of speaking a foreign language.

Sommerstein (1974) broadened the role of constraints. Like Shibatani, he proposed that the grammar should 'contain an exhaustive set of conditions on the output of phonological rules—in fact, a surface phonotactics' (p. 71). But for Sommerstein, these surface phonotactic constraints are not redundant (as claimed by Postal and Johnson) because they can be invoked as a means to simplify the formulation of the phonological rules. Specifically, Sommerstein suggests that a phonotactic constraint can 'motivate' a phonological rule either 'positively' or 'negatively'. To take our Yawelmani example,

vowel epenthesis is positively motivated by the *CC{C, #} constraint (i.e. a vowel is inserted just in case to do so will prevent the banned consonant clusters). Vowel deletion, on the other hand, is negatively motivated by the same constraint (i.e. a vowel in a specified context will delete unless doing so would lead to a violation of the constraint).

Sommerstein focuses on positively motivated phonotactic constraints (since Kisseberth had already argued for the negatively motivated constraints) and shows how they can help to simplify the morphophonemic rules in Latin. For our purposes, we can illustrate Sommerstein's idea using the Yawelmani phonotactic constraint *CC{C, #) and the morphophonemic rules that avoid violations of this constraint. In Sommerstein's proposal, postulation of the *CC{C, #} constraint can be used to simplify the vowel epenthesis rule, which can be formulated as: Ø → *i*. At first glance, the epenthesis appears to massively over-apply since [i] would be inserted without any environmental restriction at all. But Sommerstein's proposal is that a rule like Ø → *i* can be restricted so that it can only apply if to do so would have the effect of making the output of the rule satisfy the phonotactic constraint *CC{C, #}. In other words, Sommerstein proposes that if a rule is not phonotactically motivated, then it applies as long as its structural description is satisfied. When, however, a rule *is* phonotactically motivated, then in order for the rule to apply, an input structure both must satisfy the rule's structural description and also yield an output that has been altered so as to remove a violation of the phonotactic constraint at issue.

Vowel epenthesis in Yawelmani is the most general means of avoiding violations of the *CC{C, #} constraint. But recall that in addition there are morphologically restricted examples where a consonant (specifically, [h] or [ʔ]) is deleted from a triconsonsonantal

structure arising across certain morpheme boundaries. Setting aside the issue of how to restrict the rule to the appropriate morphological structures, Sommerstein's proposal would be that the phonological description of the rule would be that {h, ʔ} → Ø if motivated by *CC{C, #}.

The idea that constraints might interact with phonological rules in some fashion did become a significant aspect of theoretical developments in the 1980s, a decade that otherwise was mainly focused on representational matters (cf. Kisseberth, this volume). Some of this work grew quite directly out of the papers discussed above. For example, Singh (1987) and Paradis (1988) both adopt approaches that recognize constraints and 'repairs' that serve to alleviate violations of the proposed constraints. There are significant differences in the details of their approaches, however. Singh adopts the ideas of Natural Generative Phonology and dismisses many phenomena as aspects of morphology, which do not count as 'phonological' rules. Setting such phenomena aside, he proposes replacing phonological rules by a language-specific set of well-formedness constraints that require surface forms violating these constraints to be adjusted by a universal set of 'repair strategies'. If multiple repairs are possible, then there must be universal principles that predict the correct repair for the language in question. Paradis, on the other hand, retains the more expansive notion of rules familiar from classical generative phonology, and assumes that the well-formedness constraints may be either universal or language-specific. She shares with Singh the idea that repairs are phonological actions triggered by the constraints and that in the event of multiple ways to avoid a constraint violation, some principles must come into play to select the appropriate repair. A third development of the constraint-and-repair approach can be found in

Calabrese (1985, 1988). Calabrese's model and the entire constraint-and-repair approach is discussed in detail in Calabrese (this volume).

20.6.3 Global rules

In this section, we discuss a proposal to grant rules the power (regarded by many as 'excessive') to 'look back' at the origin of particular phonological structures rather than be restricted to properties of the immediate input to the rule. Kisseberth (1973a) referred to rules that looked back to earlier stages of a derivation as 'global rules'.

The core idea in *SPE* is that a rule *R*'s applicability is determined entirely by properties present in the input to *R* whether that input is the underlying representation or a representation that has been modified by a rule or rules preceding *R* in the ordered set of rules. The notion of ordered rules had been utilized in the pre-generative literature (particularly Bloomfield (1939), but pre-generative phonological practice was in fact characterized by a rather different approach. This approach assumed the sequential application of phonological rules, but also assumed that a rule applies whenever conditions for the rule's application arise. At the same time, some rules may have conditions that make reference to the 'derivational history' of the structure that triggers the rule.

Specifically, writers like Boas, Sapir, Swadesh and Newman often restricted the scope of a phonological principle on the basis of whether a sound was 'organic' or 'inorganic' (cf. Kenstowicz 1976 and Silverstein, this volume). It is clear that these writers used the term 'organic' to refer to the 'primary' or (in generative phonology terms) underlying representation, and 'inorganic' to refer to sounds that were not

underlying, but inserted or otherwise represented a ‘secondary’ development. If a rule was stated to affect only organic sound structures and not inorganic ones, this derived a state of affairs that generative phonology captured by means of rule ordering. But if a rule was stated so that it affected inorganic sounds but not organic ones of the same phonological character, that was an effect that generative phonology could not express by rule ordering (since it required distinguishing between two inputs on the basis of their derivational history).

Taking his cue from this pre-generative descriptive practice, Kisseberth (1973a) suggested that the ability to look back found support in the complicated morphophonemics of Klamath. He argued that there are underlying long vowels in Klamath as well as long vowels that derive from vocalized glides. The latter ‘inorganic’ vowels shorten in certain contexts, while the ‘organic’ long vowels do not. He referred to such a shortening rule as a ‘global rule’ in that it could distinguish between sounds in its input on the basis of their derivational history. Additional evidence for globality was developed in Wilbur (1973), who sought to explain a range of ‘identity’ effects in reduplication with this notion.

The proposed concept of ‘global rules’ did not meet with much acceptance by phonologists working in the generative tradition (let alone by proponents of more ‘concrete’ approaches to phonology). The argument based on Klamath was effectively countered by an influential alternative analysis from Clements & Keyser (1983) as part of their development of the so-called *CV* approach to phonological representation. Wilbur’s arguments from reduplication were somewhat obscured in the course of the development

of prosodic morphology in the 1980s, but eventually re-emerged as part of the motivation for Optimality Theory.

The rejection of global rules, however, seemed to be based largely on the argument that they are 'unconstrained' and 'too powerful'. This sort of rejection of a proposal often seems quite reasonable. The difficulty is that there are linguists such as Householder (1965) who would reject the entire generative phonology architecture as 'unconstrained' and 'too powerful' (to say nothing of a later approach like Optimality Theory!). The deeper issue is whether or not global rules are getting at some fundamental truth about phonological systems.

In retrospect, the notion of global rules perhaps should have been expressed in a manner that would have linked them more clearly to their origin in the pre-generative notion of 'inorganic' versus 'organic'. In other words, the potential evidence that was most robustly attested was that underlying phonological elements such as the long vowels of Klamath *may* fail to undergo a phonological rule that identical derived phonological elements do undergo.

We should note Kisseberth's claim that phonological rules might need to distinguish between underlying (organic) and derived (inorganic) sounds was clearly related to Kiparsky's (1973) notion that rules might be restricted to derived environments, i.e. either across a morpheme boundary or within a morpheme just in case the structure is the result of a prior rule.

It should be mentioned that the discussion of derivational constraints and global rules has clear implications for the issue of whether or not rule ordering is required. This point is developed in Kisseberth (1973b). Unlike Natural Generative Phonology, which argued

against the necessity for rule ordering based in large part on restricting phonological rules to surface true phonetic processes, Kisseberth accepts the *SPE* conception that morphophonemic rules can in fact be quite 'opaque' (i.e. the application of a rule may not be transparently observable from the ultimate surface form). In addition, he proposes that rules must be applied in a sequence that minimizes opacity. This means that feeding sequences and bleeding sequences of application are required (cf. section 20.5, where we saw that Kiparsky had concluded that these sequencings were unmarked, albeit not *required*). But since Kisseberth *does* allow rules creating opacity to exist, some grammatical mechanism must also exist to derive the opacity that in rule ordering is created by counter-feeding and counter-bleeding sequences of application. He suggests that allowing rules to look at derivational history can achieve the same results as counter-feeding and counter-bleeding sequences of application.

For example, consider counter-bleeding. In Polish (cf. Bethin 1978, Kenstowicz & Kisseberth 1977, Kenstowicz 1994, to restrict our references to relatively early discussion) the vowel *o* is raised in front of a word-final voiced non-nasal consonant, even if the voiced consonant is devoiced in phonetic representation by a rule that devoices word-final obstruents. Thus underlying /ʒwob/ 'crib' becomes [ʒwup] due to Raising and Final Devoicing (cf. plural [ʒwobi], where the consonant is not word-final) while /sol/ 'salt' becomes [sul] by virtue of Raising and /gruz/ 'rubble' becomes [grus] by virtue of Final Devoicing. Raising does not occur in front of an underlying voiceless consonant. Thus we find [kot] 'cat' (cf. [kot-a]). In a rule ordering solution as shown in (7), Raising applies before Final Devoicing. In Kisseberth's proposal, Raising applies

before final non-nasal consonants that are *underlyingly* voiced. In other words, Raising is a global rule that looks back to underlying structure.

(7)	/ʒwob/	/ ʒwob-i/	/sol/	/gruz/	/kot/	
	ʒwub	———	sul	—	—	Raising
	ʒwup	———	—	grus	—	Final Devoicing

Consider next an example of a counter-feeding rule interaction. In Western Basque (de Rijk 1970), a non-low vowel is raised to a high vowel in front of a vowel (thus /seme+e/ 'son' becomes [semie]). Call this Raising I. There is also a rule that raises a low vowel to a non-low vowel in front of another vowel (thus /alaba+a/ 'daughter' becomes [alabe+a]). Call this Raising II. Notice that in a theory where feeding (the minimization of opacity) is required, we would expect [alabe+a] to undergo Raising I, resulting in the incorrect *[alabi+a]. To prevent this from happening, Kisseberth's proposal is that Raising I is formulated so that it raises only [–low] vowels that are not underlyingly [+low]. Again, opacity is derived by evaluating an input to a rule on the basis of its underlying structure.

Kisseberth does note one complexity where a sequencing statement is required even when global rules are included in one's grammatical arsenal. Specifically, although bleeding minimizes opacity, rules can potentially be 'mutually bleeding'. For example, German dialects (cf. Vennemann 1970, Kiparsky 1971) differ in their treatment of an input /laŋg/ 'long', which meets the structural description of two rules: a rule that devoices final obstruents (see the discussion earlier in this chapter) and a rule that deletes *g* after a nasal. If Devoicing applies before *g*-Deletion, one derives [laŋk], which is correct in some dialects. If *g*-Deletion applies before Devoicing, [laŋ] is derived. This is

correct in other dialects. Neither outcome is opaque, so minimization of opacity does not appear relevant. Globality is also not at issue. A stipulation with respect to rule sequencing seems to be the only viable solution.

20.6.4 Trans-derivational constraints.

One other idea concerning derivations that received some discussion in the 1970s did not evolve out of Kisseberth (1970), though it did—like conspiracies—find its spiritual origins in syntactic proposals flourishing at the time (cf. Lakoff 1973). Specifically, the proposal was advanced that the derivation of one lexical item might be affected by the derivation of a different lexical item. This hypothetical phenomenon was referred to as a 'trans-derivational' rule. Kisseberth & Abasheikh (1974) argued that in the complex morphophonemics of the perfect stem in Chimwiini, the choice between using a suffixed perfect stem and a so-called 'imbricated' perfect stem (i.e. a stem form involving a kind of ablauting of the final vowel of the stem) is governed by the need to prevent two contrasting lexical stems from being merged (see discussion of this example in Kenstowicz 2005).

Trans-derivational constraints did not have a powerful impact on phonological analysis at the time, though it should be noted that ultimately in Optimality Theory, the idea achieved considerable importance in the postulation of so-called 'Output-Output' faithfulness (Kenstowicz 1997, Benua 1997).

20.7 The controversies of the 1970s in the light of subsequent developments

Most of the controversies reviewed in this chapter largely were overshadowed as generative phonology more and more focused on representational issues in the late 1970s and throughout the 1980s. Generative phonologists became committed to the hope that advances in our understanding of representation could substantially shift the role of derivations/rule ordering to the margins of phonological theory.

Developments in terms of representation, however, could not obscure the conceptual simplicity of the original conspiracy argument, i.e. that a single constraint could drive or inhibit multiple rules. Nor could it undercut the motivation for the idea that constraints must be involved somehow in derivations. Indeed, one of the foundational notions of autosegmental phonology (cf. Kisseberth, this volume) was the Obligatory Contour Principle (OCP, the ban on identical adjacent feature values—for example, the disfavoring of High High sequences in many tone languages). Pervasive evidence came to light that the OCP motivated a great variety of phonological actions, all of which involved adjustments to avoid successive identical feature specifications. Even outside of the sphere of autosegmental phonology, constraints and templates (and their satisfaction) played an increasing role. For example, as syllable structure became central to phonological theory, it became clear that syllabification, as well as phenomena that interact with it, could not be reduced to the sort of rules and rule interactions found in classical generative phonology (cf. the particularly significant treatment of vowel epenthesis in Itô 1989).

Rules in generative phonology are a melding together of a structural description and a structural change. The structural description identifies the trigger of a phenomenon; the

structural change is an action that is the response to the trigger. In the 1990s, Optimality Theory took the step of removing the actions from phonological theory. Anything could be an action; the issue is whether any given action is actually the response to the trigger chosen by a language (or phrased differently, the means by which a language satisfies the constraint set). OT took the further step of taking the triggers to be a (universal) set of constraints, but with the recognition that constraints could be in conflict and that these conflicts are resolved in terms of the ranking of constraints (formalizing an intuition found in the earlier work of Stampe 1973).

From the beginning, it was recognized that Optimality Theory represented a solution to the conspiracy problem. A constraint like *CC{C, #}, by being ranked higher than faithfulness constraints, can trigger vowel epenthesis or consonant deletion. Furthermore, this same constraint, by being ranked higher than some constraint motivating vowel loss, could serve to block the loss of a vowel if this loss would lead to a violation of *CC{C, #}.

Even the much maligned notion of 'global rules' turns out to be critical to the OT enterprise. OT replaced language-specific rules, tailored to the particular alternations observed in a given language, with a collection of universal constraints. But a given constraint is not implemented fully or even at all in every language. The task of an OT analysis (and of the child acquiring an OT grammar) is to explain when a given constraint is implemented in a language and when not. Critical to this task is the introduction of a set of constraints called 'Faithfulness' constraints that are satisfied by preferring a representation that maintains an underlying property as opposed to obeying a particular constraint. Ranking a F(aithfulness)-constraint higher than a P(honological)-constraint

preserves the underlying feature specification. Ranking the P-constraint higher than the F-constraint allows the underlying specification to be modified. It should be obvious that Faithfulness is the equivalent of a 'global' condition on P-constraints that restricts them from being applicable to underlying representations (i.e. 'organic' elements in the pre-generative characterization). Faithfulness constraints are, of course, an extremely 'powerful' device, as indeed is the very notion of a 'universal constraint' in Optimality Theory. But the notion of Faithfulness is at the very heart of the complex task of navigating how P-constraints can be universal and yet not universally fully manifested. In generative phonology, particular languages simply lack some rules that are clearly manifested in other languages. In OT, the phonological constraints are universally present in all languages; Faithfulness constraints bear the burden of explaining why underlying elements do not actually adjust their shape to these universal constraints in certain cases. Models of the acquisition of OT grammars have proposed that in the initial state of phonotactic learning, markedness constraints dominate input-output faithfulness (reflecting the predominance of Stampean natural processes) while output-output constraints dominate markedness minimizing paradigmatic alternation (cf. Tesar & Smolensky 2000 for extended discussion of the acquisition of OT grammars).

Recall that Kisseberth (1973) suggested that derivational history (i.e. global rules) potentially might serve to account for opacity. Discussion of the opacity issue largely receded from attention as generative phonology gradually shifted its focus from rules and their ordering to representations. The issue returned to the forefront in Optimality Theory because opacity now became pervasive in every phonological system. Many of the universal set of constraints are violated in any language (in other words, these constraints

are opaque). Although as discussed earlier, Faithfulness constraints can be used to prevent underlying forms from obeying a constraint, they do not actually explicate the full extent of opacity in phonology. And while efforts have been made to develop an OT account of opacity, these attempts have come at considerable cost in empirical coverage and an increase in theoretical machinery. It seems safe to say that phonological opacity represents an issue from the 1970s that is still unresolved (Vaux 2008).

It is not surprising that abstractness (which is closely intertwined with opacity) likewise remains a controversial issue, decades after the 1970s. Neither the exploration in the 1980s of representations (whether within the generative tradition or in radical departures like Government Phonology (Kaye, Lowenstamm, & Vergnaud 1985, 1990; Ritter, this volume) nor the development of Optimality Theory (van Oostendorp, this volume) showed much interest in abstractness issues. Arguments with respect to representations or the structure of OT relied on analyses that often were assumed without much regard for their degree of abstractness or the extent of the evidence supporting them. It is not surprising that OT faced some of the same backlash as *SPE* with respect to abstractness. The abstractness of phonology is reemerging as a central problem in efforts to model the acquisition of OT grammars (see Heinz and Rawski, this volume).

In summation, the derivational concerns of the 1970s are not just concerns of historical interest; and while they do not necessarily remain a focus of attention at every step in the journey to a deeper understanding of a speaker's phonological knowledge, they remain significant challenges to the ultimate attainment of that understanding.

20.8 Summary and conclusion

This chapter reviewed some of the questions that arose concerning the model of phonology inaugurated by *SPE* where an evaluation measure placed a premium on generating morpheme alternants from a single underlying form by an ordered set of rules defined over distinctive feature matrixes. These questions centered around the abstractness of underlying representations, the ways in which the rules apply to derive the surface forms, and the existence of structural constraints separate from the sound changes themselves that shape and constrain the derivation. The recent literature is revisiting many of these basic questions with tools derived from computational modeling and experimental psychology. Time will tell whether the results of the past fifty years can be sustained or must be reconsidered.

References

Anderson, Stephen. 1969. West Scandinavian vowel systems and the ordering of phonological rules. PhD thesis, MIT.

Anderson, Stephen. 1974. *The organization of phonology*. New York: Academic Press.

Baković, Eric. 2007. A revised typology of opaque generalizations. *Phonology* 24: 1–43.

Benua, Laura. 1997. Transderivational identity: Phonological relations between words. PhD thesis, University of Massachusetts, Amherst.

Bethin, Christina. 1978. Phonological rules in the nominative singular and genitive plural of the Slavic substantive declension. PhD thesis, University of Illinois, Urbana-Champaign.

Blevins, Juliette. 2004. A reconsideration of Yokuts vowels. *International Journal of American Linguistics* 70: 33–51.

Bloomfield, Leonard. 1939. Menomini morphophonemics. *Études phonologiques dédiées à la mémoire de M. le prince N.S. Trubetzkoy*, 105–15. Prague: Travaux du Cercle Linguistique de Prague 8.

Brame, Michael. 1972. On the abstractness of phonology: Maltese ʕ. In Michael Brame, (ed.), *Contributions to generative phonology*, 22–61. Austin: University of Texas Press.

Buccola, Brian. 2013. On the expressivity of Optimality Theory versus Ordered Rewrite Rules. Glyn Morrill and Mark-Jan Nederhof, (eds), *Proceedings of Formal Grammar 2012 and 2013*, 142–58. Berlin-Heidelberg: Springer-Verlag.

Calabrese, Andrea. 1985. Metaphony in Salentino. *Rivista di Grammatica Generativa* IX–X: 3–140.

Calabrese, Andrea. 1988. Towards a theory of phonological alphabets. PhD thesis, MIT.

Chafe, Wallace. 1968. The ordering of phonological rules. *International Journal of American Linguistics* 34: 115–36.

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

Clements, George N. 2002. Feature economy in sound systems. *Phonology* 20: 287–333.

Clements, George N. and Samuel Jay Keyser. 1983. *CV phonology: A generative theory of the syllable*. Cambridge, MA: MIT Press.

de Rijk, Rudolf. 1970. Vowel interaction in Bizcayan Basque. *Fontes Linguae Vasconum* 2(5): 149–67.

Donegan, Patricia and David Stampe. 1979. The study of Natural Phonology. In Daniel Dinnsen, (ed.), *Current approaches to phonological theory*, 126–73. Bloomington, IN: Indiana University Press.

Halle, Morris. 1962. Phonology in generative grammar. *Word* 18: 54–72.

Halle, Morris. 2019. Aspects of the morphophonology of the verb in Latin and in German and English. *Linguistic Inquiry* 50:1, 3–12.

Halle, Morris and George N. Clements. 1983. *Problem book in phonology*. Cambridge, MA: MIT Press.

Hansson, Gunnar Ólafur and Ronald Sprouse. 1999. Factors of change: Yowlumne vowel harmony then and now. In M. Caldecott, S. Gessner, and E.-S. Kim (eds.), *Proceedings of WSCLA IV*, 39–57. Vancouver: Department of Linguistics, University of British Columbia.

Henriksen, Nicholas. 2017. Patterns of vowel laxing and harmony in Iberian Spanish; data from production and perception. *Journal of Phonetics* 63: 106–26.

Hockett, Charles. 1967. The Yawelmani basic verb. *Language* 43: 208–22.

Hockett, Charles. 1973. Yokuts as testing ground for linguistic methods. *International Journal of American Linguistics* 39: 63–79.

Hooper, Joan B. 1976. *An introduction to Natural Generative Phonology*. New York: Academic Press.

Householder, Fred. 1965. On some recent claims in phonological theory. *Journal of Linguistics* 1: 13–34.

Hyman, Larry. 1970. How concrete is phonology? *Language* 46: 58–76.

Idsardi, William. 1998. Tiberian Hebrew spirantization and phonological derivations. *Linguistic Inquiry* 29: 137–73.

Idsardi, William. 2000. Clarifying opacity. *The Linguistic Review* 17: 337–50.

Itô, Junko. 1989. A prosodic theory of epenthesis. *Natural Language & Linguistic Theory* 7(2): 217–59.

Johnson, C. Douglas. 1972. *Formal aspects of phonological description*. The Hague: Mouton.

Kaye, Jonathan. 1974. Opacity and recoverability in phonology. *Canadian Journal of Linguistics/ Revue canadienne de linguistique* 19(2): 134–49.

Kaye, Jonathan. 1980. The mystery of the tenth vowel. *Journal of Linguistic Research* 1: 1–14.

Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. 1985. The internal structure of phonological representations: A theory of Charm and Government. *Phonology Yearbook* 2, 305–28.

Kaye, Jonathan, Jean Lowenstamm, and Jean-Roger Vergnaud. 1990. Constituent structure and government in phonology. *Phonology* 7(2): 193–231.

Kenstowicz, Michael. 1976. Rule application in pregenerative American phonology. In Andreas Koutsoudas (ed.), *The application and ordering of grammatical rules*, 259–80. The Hague-Paris: Mouton.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. Cambridge, MA: Blackwell.

Michael Kenstowicz.1997. Base identity and uniform exponence: Alternatives to cyclicity. In Jacques Durand and Bernard Laks (eds), *Current trends in phonology: Models and methods,* 363–94. Salford: University of Salford.

Kenstowicz, Michael. 2005. Paradigmatic uniformity and contrast. In Laura Downing, T. A. Hall, and Renate Raffelsiefen, (eds), *Paradigms in phonological theory*, 145–69. Oxford: Oxford University Press.

Kenstowicz, Michael and Charles Kisseberth. 1971. Unmarked bleeding orders. In Charles Kisseberth, (ed.), *Studies in generative phonology*, 1–12. Edmonton, Alberta: Linguistic Research Inc.

Kenstowicz, Michael and Charles Kisseberth. 1977. *Topics in phonological theory*. New York: Academic Press.

Kenstowicz, Michael and Charles Kisseberth. 1979. *Generative phonology: Description and theory*. New York: Academic Press.

King, Robert. 1980. The history of final devoicing in Yiddish. In Marvin I. Herzog, Barbara Kirshenblatt-Gimblett, Dan Miron, and Ruth Wisse, with the assistance of Alan Huffman (eds), *The Field of Yiddish. Studies in language, folklore, and literature*. *Fourth collection*, 371–430. Philadelphia, PA: Institute for the Study of Human Issues.

Kiparsky, Paul. 1965. Phonological change. PhD thesis, MIT. [Reprinted by the Indiana University Linguistics Club, 1971.]

Kiparsky, Paul. 1968a. Linguistic universals and language change. In Emmon Bach and Robert Harms, (eds), *Universals in linguistic theory*, 191–212. New York: Holt.

Kiparsky, Paul. 1971. Historical linguistics. In William O. Dingwall, (ed.), *Survey of linguistic science,* 576–649. College Park, MD: Linguistic Program, University of Maryland.

Kiparsky, Paul. 1968. How abstract is phonology? Bloomington, IN: The Indiana University Linguistics Club. Also in Osamu Fujimura, (ed.), *Three dimensions of linguistic theory*, 5–56. Tokyo: TEC, 1973. Reprinted in Paul Kiparsky, *Explanation in phonology*, 119–64. Dordrecht: Foris, 1982.

Kiparsky, Paul. 1973. Abstractness, opacity and global rules. In Osamu Fujimura (ed.), *Three dimensions of linguistic theory*, 1–136. Tokyo: TEC.

Kiparsky, Paul. 1982. From cyclic to lexical phonology. In Harry van der Hulst and Norval Smith, (eds), *The structure of phonological representations (part I),* 131–75. Dordrecht: Foris.

Kisseberth, Charles W. 1969. On the abstractness of phonology: The evidence from Yawelmani. *Papers in Linguistics* 1: 248–82.

Kisseberth, Charles W. 1970. On the functional unity of phonological rules. *Linguistic Inquiry* 1: 291–306.

Kisseberth, Charles W. 1973a. On the alternation of vowel length in Klamath: A global rule. Michael Kenstowicz and Charles Kisseberth, (eds), *Issues in phonological theory*, 9–26. The Hague: Mouton.

Kisseberth, Charles W. 1973b. Is rule ordering necessary in phonology? In Braj B. Kachru, Robert B. Lees, Yakov Malkiel, Angelina Pietrangeli, and Sol Saporta (eds), *Issues in linguistics: Papers in honor of Henry and Renée Kahane,* 418–41. Urbana, IL: University of Illinois Press.

Kisseberth, Charles W. and Mohamed Imam Abasheikh. 1974. A case of systematic avoidance of homonyms. *Studies in the Linguistic Sciences* 4: 107–24.

Koutsoudas, Andreas, Gerald Sanders, and Craig Noll. 1971. The application of phonological rules. Bloomington, IN: The Indiana University Linguistics Club. Reprinted in *Language* 50: 1–28 (1974).

Kuroda, S.-Y. 1967. *Yawelmani phonology*. Cambridge, MA: MIT Press.

Lakoff, George. 1973. Some thoughts on trans-derivational constraints. In B.B. Kachru, R.B. Lees, Y.Malkiel, A. Pietrangeli & S. Saporta (eds), *Issues in Linguistics: Papers in Honor of Henry & Renée Kahane,* 442–52. Urbana: University of Illinois Press.

Liberman, Mark. 2017. Towards progress in theories of language sound structure. In Diane Brentari and Jackson J. Lee (eds), *Shaping phonology: Festschrift for John Goldsmith*. Chicago: University of Chicago Press.

Newman, Stanley. 1944. *Yokuts language of California*. New York, NY: Viking Fund.

Odden, David. 2003. *Introducing phonology*. Cambridge: Cambridge University Press.

O'Hara, Charlie. 2017. How abstract is more abstract? Learning abstract underlying representations. *Phonology* 34: 325–45.

Paradis, Carole. 1988. On constraints and repair strategies. *The Linguistic Review* 6: 71–97.

Perlmutter, D. M. 1968. Deep and surface constraints in syntax. PhD thesis, MIT. [Published by Holt, Rinehart and Winston, New York, NY, 1971.]

Postal, Paul. 1968. *Aspects of phonological theory*. New York, NY: Harper and Row.

Prickett, Brandon. 2019. Learning biases in opaque interactions. *Phonology* 36: 627–53.

Prince, Alan and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Cambridge, MA: MIT Press.

Ross, John Robert. 1967. Constraints on variables in syntax. PhD thesis, MIT.

Sapir, Edward. 1915. Notes on Judeo-German phonology. *The Jewish Quarterly Review* 6: 231–66. Reprinted in David G. Mandelbaum (ed.), *Selected writings of Edward Sapir in language, culture, personality*, 252–72. Berkeley, CA: University of California Press, 1949.

Saporta, Sol. 1965. Ordered rules, dialect differences, and historical processes. *Language* 41: 218–24.

Shibatani, Masayoshi. 1973. The role of surface phonetic constraints in generative phonology. *Language* 49: 87–106.

Singh, Rajendra. 1987. Well–formedness conditions and phonological theory. In Wolfgang Dressler, Hans Luschützky, Oskar Pfeiffer, and John Rennison (eds), *Phonologica 1984*, 273–85. Cambridge: Cambridge University Press.

Sommerstein, A.H. 1974. On phonotactically motivated rules. *Journal of Linguistics* 10(1): 71–94.

Stampe, David. 1973. *A dissertation on Natural Phonology*. PhD thesis, University of Chicago. [Reprinted by Garland Publishing Co., New York, NY, 1979.]

Tesar, Bruce and Paul Smolensky. 2000. *Learnability in Optimality Theory*. Cambridge, MA: MIT Press.

Vaux, Bert. 2008. Why the phonological component must be serial and rule-based. In Bert Vaux and Andrew Nevins (eds), *Rules, constraints, and phonological phenomena*, 20–60. Oxford: Oxford University Press.

Vennemann, Theo. 1970. The German velar nasal: A case for abstract phonology. *Phonetica* 22: 65–81.

Wiese, Richard. 1996. *The phonology of German*. Oxford: Oxford University Press.

Wilbur, Ronnie. 1973. *The phonology of reduplication*. PhD thesis, University of Illinois, Urbana-Champaign. [Printed by the Indiana University Linguistics Club.]