CHAPTER ONE

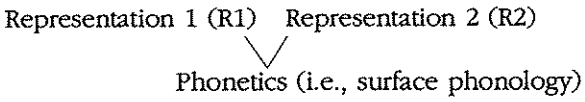
Global Determinacy and Learnability in Phonology

B. Elan Dresher
University of Toronto

Harry van der Hulst
Leiden University/HIL

Most current theories of phonology allow for the possibility of different phonological representations for a particular phonetic unit (e.g., a segment type). Here we take the expression *phonetic unit* to refer to a unit in the phonological representation that forms the input to the phonetic interpretation component, that is, the surface phonological representation. A schematic representation of a two-to-one (in principle, many-to-one) relation between an underlying representation and a surface representation is given in (1):

(1) Many-to-one mapping of phonological representations to phonetics



It need not be the case that both R1 and R2 are lexical or underlying representations. Depending on the theory, multiple sources for a surface representation may arise in the course of a phonological derivation. We therefore take underlying to refer to nonsurface (lexical or intermediate) levels of representation.

Positing multiple underlying sources for a single surface unit may be the result of theories that are as yet too loosely constructed or imprecise, but typically that is not the case. Rather than resulting from theoretical insecurity, this variability is introduced deliberately in order to explain differences in

processes that involve these units. In such cases, the phonological representation is underdetermined by the local phonetic properties of the relevant unit and can be established only by taking into account a wider range of data. In older versions of generative grammar the additional data would typically involve processes that affect this unit. We say that grammars in which such cases arise have the property of *global determinacy*.

In this chapter we examine a number of cases of this type and speculate on the consequences of global determinacy for the theory of language acquisition. We do not attempt to present a systematic overview. Our goal is more modest, in that we merely wish to point out that global determinacy is not a problem associated only with older versions of generative phonology, where it was recognized as a problem of (excessive) abstractness of lexical representations. Perhaps the contrary is true. Recent advances in phonology have led to an enormous increase in representational possibilities for identical phonetic events. Even though specific possibilities are usually pervasively grounded in analyses of a set of facts that may not be analyzable otherwise, the addition of every possibility introduces further ambiguity into the analysis of other sets of data that may be compatible with more than one available account.

The fact that a choice must be made between R1 and R2 raises a learnability question. The logical problem of language acquisition is the following: Construct a theory so that, for all languages, it is possible to establish a grammar on the basis of a relevant set of data, which, as is usually added, underdetermines the grammar to be attained. Given global determinacy, crucial information for certain parts of the grammar is unavailable in the local context to which those parts directly pertain. Hence, with strict reference to the local context, several grammars could be selected. In that case, there must be a reasonable way of bringing nonlocal factors into the decision space, through cues in the data that lead the language learner to the correct grammar, rather than to other grammars that are possible in general but happen to be incorrect in the given case. To arrive at a theory that explains how a grammar can be acquired, we must be able to specify what Lightfoot (1989) called a *learning path*; that is, it must be possible to specify exactly which cues determine the choice of one analysis over the other. In Lightfoot's terms, we must identify the child's *trigger experience*.

In our case, this means that there must be a reasonable way for the language-learning child to determine which representation corresponding to a particular phonetic surface form is called for if this is not locally determined. It is not sufficient to show that a unique analysis can be chosen once all the evidence is known (either because all others are wrong, or because it is the best one, according to some evaluation metric). Though consideration of large data sets and complicated lines of argumentation may in practice lead the investigator to decide which representation is correct,

we must assume that children fix parameters without knowing all the evidence, or even a lot of evidence, unless we admit that our model does not aim at accounting for the fact that children fix parameters on the basis of very limited evidence.

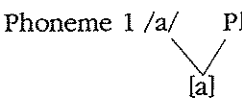
In a broad sense, the issue of global determinacy is central to what phonology is all about. Phonology would be a trivial discipline if there were a strictly local mapping from the continuous speech signal (or surface phonology) onto the discrete mental representations that we have reason to believe form part of linguistic capacity. Much research in phonology suggests that theories limited to local determination (a strict form of biuniqueness) are descriptively inadequate. It is also clear, however, that too much flexibility leads to an unconstrained theory that allows us to construct too many representations, most of which are never called for.

Our goal, then, is to indicate some of the major causes of global determinacy that arise in variants of generative phonology, including current multilinear approaches. Of course, we are not suggesting that all the cases we discuss form part of a single coherent model. Which cases are relevant depends on the specific ideas one has regarding such issues as underspecification, unary features, feature geometry, dependency relations, system-dependent or relative interpretation of features, abstract underlying representations, rule ordering, and deletion. Nevertheless, we hope to show that the problem of global determinacy arises in every nontrivial theory of phonology and has certain recurring characteristics, no matter what guise it comes in. On the moralistic side, we note that current phonological theorizing may suffer from a lack of consideration for learnability issues raised by the explosive increase of representational possibilities, especially those involving underspecification and geometrical feature arrangements.

1. STRUCTURALIST PHONEMICS

In structuralist phonemics, the task implied in (1) would be framed as one of assigning a phonetic speech sound to its proper phoneme. Suppose, for example, that the phonetic segment in question is the vowel [a], as in (2), and that in the language under consideration the sound [a] is an allophone of two different phonemes, the phoneme /a/ and the phoneme /e/ (we are assuming here a liberal structuralist theory in which phonemes may overlap in this way):

(2) Structuralist phonemics: Overlapping phonemes



This variability in phonemic membership could be posited in order to explain differences in processes that involve this segment. For example, suppose that [a] that belongs to phoneme /e/ is associated with palatalization of preceding velars, whereas [a] from /a/ is not. In such cases, the phonological representation is underdetermined by the local phonetic properties of the segment and can be established only by taking into account a wider range of data. These data could be phonological, involving, for instance, palatalization or other relevant processes that distinguish the two representations, or morphological, as when [a] alternates with [e] in some cases but not in others.

Because global determinacy raises learnability problems, there have been recurring attempts to limit the theory of grammar to local determinacy. In structuralist phonemics, this was accomplished by imposing a number of constraints on phonemic representations. One was the constraint against mixing of levels, ruling out recourse to morphological information in identifying phoneme membership. Another was what later became known as the requirement of biuniqueness (Chomsky, 1964), which ruled out overlapping phonemes in identical contexts and required that an allophone in a given context could belong to only one phoneme. These constraints were motivated precisely by a desire to impose local determinacy for purposes of acquisition. Thus, Bloch (1941) argued that schwa must be a phoneme in English, because otherwise a child would not be able to identify which phoneme any given schwa belongs to on the basis of local cues.

This is one way of making the acquisition problem more tractable, but the price is too high. Much research in phonology suggests that theories limited to local determination are descriptively inadequate, essentially because such theories must give up the idea that the relation between allomorphs is expressed in terms of a common unique underlying representation. So it appears that we have to deal with global determinacy not by ruling it out, but by other means.

2. SOURCES OF GLOBAL DETERMINACY IN UNILINEAR GENERATIVE GRAMMAR

The theory of SPE (Chomsky & Halle, 1968), which posited fully specified, two-dimensional (unilinear) phonological representations, did away with the constraints imposed by many structuralist theories, opening the way to more descriptively adequate grammars. In the process, the problem of global determinacy made an impressive comeback. SPE allowed for derivations as in (3):

(3) Abstractness in linear generative phonology (counterbleeding order)

Underlying	/. . . k e . . ./	/. . . k a. . . /
Palatalization	c e	-
Lowering	c a	-
Surface	[. . . c a . . .]	[. . . k a . . .]

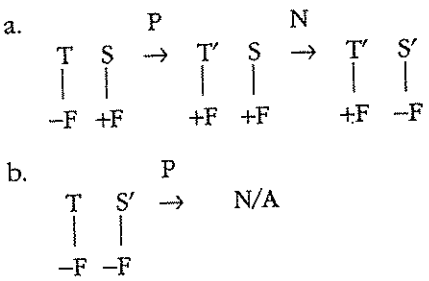
As in the example in (2), suppose we find that a language has a rule of palatalization that changes /k/ to /c/ before the front vowels /i/ and /e/. Suppose further that we also find this change occurring before some instances of [a], even though there is no trace at the surface of the phonetic element that usually causes palatalization. The SPE theory nonetheless allowed for the possibility of specifying these cases of [a] with a palatal feature, to trigger the process of palatalization. Then, before phonetic interpretation takes place, this feature is changed, in this case as a result of lowering, leading to a merger with [a] derived from underlying /a/, which lacks the palatalizing feature in the first place.

The key to this kind of abstractness is that palatalizing [a] behaves as if it has the palatalizing feature. The fact that it has this feature underlyingly cannot be determined by the local phonetic properties of [a]. It is only because [a] behaves as if it has this feature with respect to a process that refers to it that we can learn that it must have the feature underlyingly. Crucial to this type of analysis is the counterbleeding relation between palatalization and lowering, because if lowering were to apply first, palatalization could not apply.

A classical and much debated case of this kind was presented in Hyman's analysis of Nupe (Hyman, 1970). In Nupe, consonants are labialized (C^w) before the round vowels /u/ and /o/, and palatalized (C^y) before the front vowels /i/ and /e/. Before the low vowel /a/ (which is neither round nor front), we find labialized, palatalized, and plain consonants. Hyman argued that there are three underlying sources for /a/, that is, /E/, /O/, and /a/. The first two cause preceding consonants to be palatalized and labialized, respectively, and then merge with /a/. For a child to choose /E/ and /O/ as the underlying source of /a/, it must be determined that there is a rule palatalizing and labializing consonants in completely different contexts. Then the child must decide that the contrast between C, C^w, and C^y is not phonemic, which implies that the latter two must be the result of spreading from the local context, even when occurring before /a/. This is a clear case of what we mean by global determinacy.

More generally, this type of situation can be represented schematically as in (4):

(4) Apparent overapplication of P: Trigger S merges with nontrigger S′



(Here, even though we are talking about the unilinear youth of generative phonology, we adopt a quasi-autosegmental notation. We do not exploit the autosegmental potential until we reach section 3, however.)

In (4a), a segment type S causes a process P that involves the spreading of some feature [+F] to another segment T. S, however, appears on the surface as S′, which is not [+F] (i.e., the phonetic interpretation of this S′ shows no trace of the phonetic contribution that [+F] generally makes). The SPE theory, nonetheless, allowed for the possibility of specifying what appears as surface [S′] with the feature [+F] (i.e., as underlying /S/), so that it can trigger the process P. Then, before phonetic interpretation takes place, [+F] associated with S is changed to [−F] by rule N. Rule N leads to the merger of S with S′; S′ is specified [−F] in the first place and so does not trigger P, as shown in (4b).

Sometimes, S actually surfaces in other occurrences of the same morphemes, so that we can infer the presence of S throughout the whole paradigm, given the principle that allomorphy is reduced to a unique underlying morpheme shape. SPE, however, did not impose alternation as a condition on this kind of abstractness. Hence, the term *abstractness* came to refer mainly to the possibility of positing underlying phonological (i.e., featural) information in a morpheme that does not appear on the surface anywhere in the paradigm. To have this kind of abstractness, SPE had to allow extrinsic rule ordering (in (4), a counterbleeding relation), that is, ordering between rules that cannot be derived from properties, phonological or otherwise, of the rules in question. Abstractness and extrinsic rule ordering make it possible to relate very distantly related surface forms (e.g., *father* and *paternal*). In principle, the underlying form of [kæt] could be any string of phonological material, including the empty string.

The learning problems entailed by this kind of global determinacy led to suggestions to abandon or restrict either abstractness or extrinsic rule ordering, or both. Kiparsky's (1968/1982a) seminal paper argued that abstractness that is not supported by alternation is either ruled out (the strong Alternation Condition) or limited to circumstances in which at least two processes make

reference to S (the weak Alternation Condition). Other phonologists went further in attempting to limit abstractness, returning to theories with local determinacy. Such a position was taken by proponents of Natural Generative Phonology (NGP) (Hooper, 1976; Vennemann, 1973), leading to the result that all rules that were not surface true were excluded from the phonology (with wide-ranging consequences for the treatment of allomorphy). Grammars adhering to the constraint that all rules must be surface true do not allow abstract underlying forms and do not need extrinsic rule ordering. (Recently the NGP surface-true constraint has been revived in Declarative Phonology; see Scobbie, 1992).

It is clear, however, that from the point of view of learnability, the issue of abstractness cannot be resolved by considering only the distance between underlying and surface representations, the need for extrinsic rule ordering, or the relation between rules and surface phonotactics. Indeed, the question cannot even be raised unless this is done against the background of a particular theory of UG. We now show that the learnability problems that are raised by the type of abstractness discussed here need not be too serious, given certain assumptions regarding principles of UG.

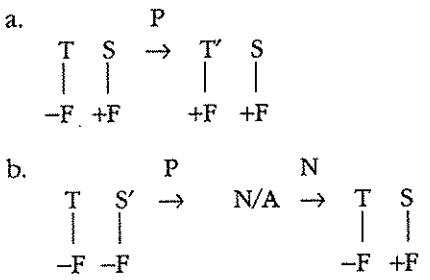
First, we observe that the nonlocal factors that enter into the lexical assignment of nonsurfacing features to segments are rather limited in the sense that the cues are, partly at least, present in the immediate context of the segments in question. If, then, UG imposes strong constraints on the form of phonological rules, abstractness could emerge as a natural consequence, posing no learnability problems. To return to our earlier examples, suppose that it were a principle of UG that a segment could be palatalized only by an adjacent segment characterized by the features [−back, −low]. Then, if it can be established that /k/ is palatalized adjacent to [a], it follows that the learner must assign these features to the representation of [a], automatically converting it to /e/, whether that appears on the surface or not. With reference to (4), if the feature [+F] can be spread by P only from an adjacent segment that is characterized by [+F], then, if it can be established that [+F] has been spread by P to T′ from S′, the learner must assign [+F] to S′, converting it to /S/. Where there is no option, there is no learnability problem.

Our Nupe example has shown, however, that in order to decide that T′ has picked up [+F] from a neighboring segment, a global decision must be made that T′ is not itself underlying. In Nupe, there are no alternations between T and T′, so global consideration of the segment inventory is required. In other cases, where T′ alternates with T, the alternation could provide evidence, as long as the learner knows about the allomorph with T and can decide that T, not T′, is underlying; these decisions also involve global considerations.

In the cases discussed here, a segment spreads a property that it lacks on the surface. The reverse was also allowed in SPE: A segment bearing

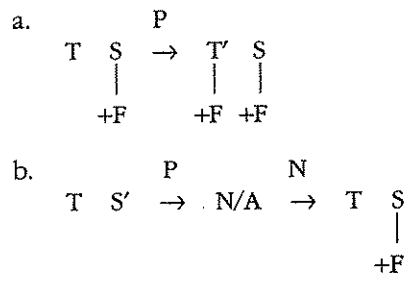
[+F] on the surface fails to participate in a process spreading [+F]. One typical case is shown schematically in (5):

(5) Apparent underapplication of P: Nontrigger S′ merges with trigger S



In SPE, the distinction between changing and adding information could not be made because forms had to be fully specified at all levels of a derivation. In theories allowing underspecification (cf. section 4) or unary features (cf. section 4.4), however, this distinction becomes viable. (We show later that exploitation of the autosegmental machinery blurs this distinction again because overapplication can also be handled without changing information.) The kind of abstractness exemplified in (5) does not crucially depend on specifying segments with information that is changed or deleted before reaching the surface, as was the case in (4), where [+F] must be changed to the opposite value (if F is a binary feature) or deleted (if F is a unary feature). The example in (5) can be recast as in (6), for example, where a segment is provided with additional information after a rule that refers to this kind of information has applied:

(6) Apparent underapplication of P: Feature-adding account



This type of case typically arises with respect to redundant properties. Suppose there is a process of vowel rounding triggered by front rounded vowels, but not by back rounded vowels; suppose also that there are no

back unrounded vowels. One way to explain why back vowels do not trigger rounding would be to propose that rounding, being nondistinctive for back vowels, is not specified on back vowels. Our rule N, then, would be a redundancy rule filling in rounding on back vowels at the end of the derivation. A similar situation exists when voiced obstruents trigger voicing assimilation, whereas sonorants (redundantly voiced in most models) do not. The relation between P and N is counterfeeding, because if N were to apply first, P would apply as well.

Vowel harmony systems provide us with many examples of this kind. In Hungarian, no contrast exists between [i] and [i̮]—we find only the former. This segment, however, does not trigger palatal harmony in a certain set of cases. This fact can be accounted for by assuming that [–back] is not specified on high, nonround vowels. Harmony applies, and only then is [–back] filled in (cf. van der Hulst, 1987).

Cases (5) and (6), like case (4), involve a counterrelation in terms of rule ordering as well as a rule P that is not surface true. Again, as in the case in (4), assumptions regarding UG influence the extent to which this kind of variability (i.e., S′ is underlying /S/ or faithful to the surface form) is held to pose a challenge to learning. Suppose we say that redundant values are always filled in late; then this kind of underapplication becomes a consequence of a principle of UG. But again it requires global consideration of the segmental inventory to arrive at the correct grammar.

It turns out, however, that such a position cannot be maintained. In fact, it has been argued (for example, in Archangeli & Pulleyblank, 1986) that the opposite is closer to the truth; redundant values are inserted at the moment a phonological rule makes reference to them. A clear case of this type would be that of sonorants triggering voicing assimilation on neighboring obstruents. Here, as in the case of overapplication, one could propose a UG constraint on the form of phonological rules that produces underapplication as a natural consequence, thus posing no learnability problems. In this case, the principle of UG would be that a segment not triggering the spreading of some feature F cannot have F if the language shows signs of F spreading in other cases.

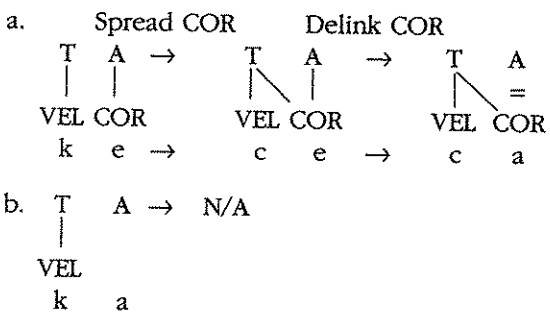
In unilinear generative phonology, cases of over- and underapplication of rules involve underlying representations that cannot be locally determined but require consideration of how processes apply, as well as inventories and/or related allomorphs, even if certain UG constraints on the form of phonological rules are assumed.

3. AUTOSEGMENTAL THEORY

The rise of multilinear theories distracted attention from issues of abstractness, and it is probably fair to say that most phonologists proceeded on the assumption that a reasonable degree of abstractness appears to be required

by the evidence. For the sake of comparison, let us translate the case in (4) in terms of nonlinear phonological representations:

(7) Spreading and delinking in autosegmental phonology



Let capital T be shorthand for a voiceless stop unmarked for place, so T linked to VELAR is a /k/. Let A be a nonhigh vowel; and, for the sake of discussion, let us say that linking a vowel to CORONAL makes it front and nonlow, that is, /e/. Suppose also that the rule of palatalization involves the spreading of the CORONAL place node onto a preceding velar, causing it to become /c/. Following this, CORONAL is delinked from the vowel, which becomes [a], merging with underlying /a/, which does not cause palatalization.

The change from linear to nonlinear representations involves an enrichment of the descriptive apparatus of phonological theory. One possible benefit of these increased representational resources is that we may be able to put constraints on derivations without giving up descriptive adequacy. For example, the derivation in (7) can still be construed as requiring a crucial extrinsic ordering between spreading and delinking, an ordering that would have to be learned. However, we can exploit the geometry of representations to give us further possible solutions not previously available. In this case, we could propose, as in (8), that only CORONAL nodes that are multiply linked may be delinked from a vowel; this, one could argue, could follow from a principle saying that one-to-one association represents the default case and is in fact forced if no loss of information results:

(8) A formulation of delinking that allows for intrinsic ordering in (7): Delink multiply linked COR from a vowel if COR appears elsewhere in the local context (due to spreading).

Given (8), delinking cannot apply until after spreading, so we could allow these rules to apply freely whenever their structural descriptions are met,

an option that was not open to us before. If all cases of overapplication could be analyzed in this way, extrinsic counterbleeding orders would no longer be necessary. Apart from the fact that this principle would not eliminate extrinsic ordering in general (it would still be necessary for cases of underapplication), it is probably not the case that delinking can be derived as an automatic consequence in all cases. In the Nupe case, for example, /E/ and /O/ are arguably no more complex than /e/ and /o/. Both /E/ and /e/ combine frontness with lowness. Both involve spreading of [–back] (or [front]), but only the former must then undergo delinking.

A particular form of underspecification is available in autosegmental phonology that makes it possible to obtain overapplication effects in new ways. It is well known that autosegmental theory allows one to specify segments only with respect to a single tier. For example, to take an extreme case, when vowel-initial words behave as if they begin with a consonant, we can posit an (unsyllabified) empty skeletal position that is identified merely as consonantal (by some feature): There may be no need to posit an abstract fully specified consonant. This, of course, is possible only in cases where the abstract entity surfaces as zero. Assuming that an unsyllabified skeletal point will not be phonetically interpreted, there is no need for a neutralization rule. (In geometrical versions of autosegmental theory, postulating floating material that receives no phonetic interpretation has become even easier, given the possibility of floating root nodes.)

Another type of case where autosegmental phonology has removed the need to posit abstract segments and neutralization rules involves the device of floating features. Consider the following example. Suppose that stems containing back vowels trigger fronting on affixal vowels. Before autosegmental phonology, one could posit abstract front vowels in the stem. Now, it is possible to posit a floating feature [–back] in the stem that will not be linked to the stem vowels but will spread to affixal vowels. When the stem occurs without an affix, one would again assume that unassociated features receive no phonetic interpretation. An analysis of this type has been proposed in van der Hulst (1987) for Hungarian, but this type of approach is common in analyses of tonal downstep that involve floating tones (Clements & Ford, 1979).

The question to be asked is how such strategies can be constrained. In any event, it is clear that floating material replaces (or adds to) old forms of abstractness without resolving issues of global determinacy.

4. UNDERSPECIFICATION THEORY

In a theory such as that of SPE, abstractness introduced by feature-changing derivations is the main source of local indeterminacy. Once we allow for underspecification, a different type of approach becomes possible in cases

of underapplication, as we already saw in the cases of (5) and (6). In this section, we discuss underspecification in more detail.

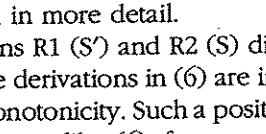
In (6), the underlying representations R1 (S′) and R2 (S) differ in that R2 contains more information than R1; the derivations in (6) are incremental or, put differently, have the property of monotonicity. Such a position is inherent in underspecification approaches: In cases like (6), for example, we might propose (as we suggested in our example) that R1 (S′) is an underspecified version of R2 (S), and that the filling in of omitted feature specifications is part of a phonological derivation.

This type of case involves the underspecification of a potential trigger. Underspecification also plays a role in the representation of targets and material intervening between trigger and target. In many of these cases, a single situation allows for more than one analysis, although, in principle, there is a unique analysis once we have seen all the evidence. Again, it turns out that the decision in favor of the correct analysis usually rests not on local considerations but rather on properties of the overall analysis.

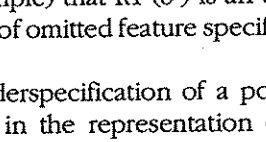
For example, a segment /i/ can be totally unspecified underlyingly in one language, whereas it may be specified as [+high, −back] in another. Or in one language all occurrences of the specification [+ATR] ([+Advanced Tongue Root]) are left out, whereas in another language all specifications [−ATR] are so treated. In both cases the choice is dependent on two factors, typically. On the one hand, underspecification rests on contrast: Noncontrastive features can be left unspecified (contrastive underspecification, Steriade, 1987). On the other hand, one must carefully consider how phonological processes apply. It is crucial to know whether segments intervening between potential triggers and targets interfere with (i.e., block) spreading. Although such factors are partially local with respect to a determination of the underspecification of segments that are directly involved in a process (as trigger, target, or intervener), these factors also have consequences for the representation of segments that are not involved in the relevant process at all. For example, it may be the case that /u/ is unspecified for [high] because /i/ must be unspecified for [high], because the behavior of /i/ indicates that it does not have specifications at all. The assumption here would be that only one value of each feature can be lexically unspecified (radical underspecification, Archangeli & Pulleyblank, 1986). Hence if /i/ is not specified as [+high], /u/ cannot be so specified either.

Another example, not involving radical underspecification, is this. Suppose that in language L_1, /k/ changes to [c] before /i/, whereas it does not do so in the same environment in language L_2. A traditional analysis would assume that L_1 has a rule that L_2 lacks. It is also possible, though, that the /k/ in L_1—call it $/k/_1$—is underspecified for some property F that /i/ can then spread to it. By contrast, $/k/_2$ in L_2 is specified for F and so cannot be influenced in this regard by /i/:

(9) a. L_1: Rule k>c applies



b. L_2: Rule k>c does not apply



Like (6), (9b) is a case of apparent underapplication of a rule, but this time it is the target of the rule that is misleading. In neutral contexts (i.e., where there is no rule that refers to F), a /k/ specified [−F] (a nontarget) looks the same on the surface as a /k/ not specified for F.

Alternatively, it could be proposed that $/k/_1$ has some property G that is a prerequisite for receiving F from /i/ (e.g., a docking site), whereas $/k/_2$ lacks G. Approaches like these are suggested by the strategy of setting up phonological representations in such a way that processes simply follow from general well-formedness conditions. In geometrical theories, versions of this approach involve the presence or absence of nodes/features on which the spreading feature is dependent.

Similar strategies can be followed with respect to intervening material. An intervener, I, may block interaction with respect to F between S1 and S2 in one language, but not in another language. This distinction could be made by specifying I in the former as [−F], and leaving I unspecified for F in the latter, assuming that I's value for F is predictable.

Segment structure in geometrical models can also vary from one language (or context) to another in ways that involve underspecification (perhaps better called nonspecification) that is inherent in systems of feature geometry. If a segment always lacks a node, then of course there is no learning problem. But it has been proposed that the presence or absence of a node specification for a given segment may be variable. For example, much research (e.g., the collection of works in Paradis & Prunet, 1991) suggests that the coronal node is absent in some languages, in that it is transparent to various spreading rules and OCP (Obligatory Contour Principle) effects, but not in other languages, which display a different range of behavior. If this is true, then it is an open question whether the variation is simply parametric, meaning that languages can vary freely in having certain nodes specified or not, or whether there is some principle that determines what nodes must be specified. Avery and Rice (1989) proposed that the presence or absence of a node depends on whether or not there are contrasts under that node.

Another example involves the representation of laryngeals. As extensively documented in McCarthy (1991), laryngeal consonants [h] and [ʔ] classify with pharyngeals and uvulars in a large number of Semitic languages. This calls for an explanation. One solution would be to specify laryngeals with a pharyngeal articulator node in addition to the laryngeal features. Steriade (1987) discussed other cases, however, in which laryngeals must lack all place information in order to explain why laryngeals and no other consonantal segments are transparent to the spreading of the vocalic place node. Further, alternations such as /t/ ~ /ʔ/ and /s/ ~ /h/ have been treated as arising from the delinking of a place node. Taken together, these facts suggest that the representation of laryngeals may differ from one language to another in ways that do not seem to affect their phonetic interpretation too much, if at all. Sandler (1990) argued that laryngeals in Hebrew show both kinds of behavior simultaneously.

Cases in which R1 is simply less specified than R2 need not involve a phonological derivation that changes R1 into R2 before phonetic interpretation takes place. We might leave R1 as it is, which implies that the phonetic interpretations of R1 and R2, [[R1]] and [[R2]] respectively, are different in the sense that [[R2]] is narrower or more precise than [[R1]]. For example, a segment specified as [–back] has a more or less well-defined phonetic description, whereas a segment unspecified for this feature shows the effect of interpolating the two closest neighboring specifications of the feature [back]. Such a difference has been suggested in work by Keating (1988). The question is how a child decides in favor of this kind of phonetic underspecification in such cases.

The issue raised by the possibility of phonetic underspecification is not unrelated to the issue of using unary instead of binary features. Phonological theories making use of unary features reduce indeterminacy significantly, but not entirely. Whether redundant voicing in sonorants is specified or not is arguably a problem for such theories if [voice] is the specified pole of the phonetic opposition voiced/voiceless. A further example involves the property of voicing in the context of Government-based phonology. It is claimed that some voiced obstruents are represented as lacking the element *tense*, whereas others are positively characterized by possessing the element *low tone*. The difference is largely one of phonological behavior, although phonetic effects are claimed to exist (cf. Harris, 1990). Underspecification of unary features may be eliminated in theories that make use of variable dependency relations. We return to this point in the next section.

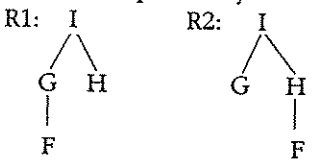
In many other cases, however, unary feature theories do not allow the option, available in binary feature theories, of specifying the default value, that is, the value that is not specified lexically. In this sense unary theories are more restricted, thus raising fewer learnability problems (cf. van der Hulst, 1989).

5. COMPETING STRUCTURES

The richer phonological representations posited by various versions of three-dimensional nonlinear phonology introduce further possibilities of indeterminacy. Thus, R1 and R2 in (1) might simply be distinct competing representations that receive the same phonetic interpretation without the intervention of any derivation at all. Here we distinguish two types of cases, which we refer to as *variable grouping* and *variable dependency*.

The first type, variable grouping, is found in proposals for variable or multiple dependency of nodes or features, in which some phonetic properties of sounds can be represented in different ways. This situation is illustrated schematically in (10), where the feature or node F may depend on either G or H, though the phonetic outcome is the same:

(10) Variable dependency of a node or feature



Piggott (1992), for example, proposed that the feature [nasal] is dominated either by a node he called spontaneous voicing or by the soft palate node. This difference plays a role, Piggott argued, in systems that have nasal harmony. The choice between the two landing sites for the feature [nasal] determines which types of segments can be transparent to nasal spreading. In other words, a decision regarding the location of the feature [nasal] in the feature tree is nonlocal in the sense that it depends not on any phonetic property of the nasal sounds themselves, but rather on the way other segments behave vis-à-vis the harmony process.

Another case of this type was suggested by Yip (1990). In this case it is the feature [lateral] that is dominated either by the spontaneous voicing node or by the articulator [CORONAL]. She referred to this situation as *double dependency*. It is interesting to note in this connection that in Dependency Phonology, properties like nasality and laterality are represented twice, once within the categorial component (in terms of the components |C| and |V|) and once as separate components |N| and |L|. Van der Hulst (in press), however, developed a variant of Dependency Phonology that does not have this kind of variability.

The second type of case of competing structures, variable dependency, does not involve variability in the geometrical locus of features. In various forms of Dependency Phonology (cf. Anderson & Ewen, 1987) or Govern-

ment Phonology (Kaye, Lowenstamm, & Vergnaud, 1985), two components can be combined in two different ways:

(11) Variable dependency/government relations
a. a → i component |a| governs component |i|
b. i → a component |i| governs component |a|

Both (11a) and (11b) are complete and distinct. If used contrastively, these representations receive different interpretations: (11a) is a lower and (11b) is a higher mid vowel. If a language lacks the relevant contrast, however, we potentially have (at least) two ways of representing identical phonetic events. We might then adopt one of two strategies: Either we choose one of the two, depending on phonological behavior, but irrespective of what exactly the phonetic height is (perhaps we will sometimes even be unable to choose on phonetic grounds); or we could leave the government relation unspecified. The former case raises issues of competing structures, as well as the issue of system-dependent or relative interpretation of phonological constructs. The latter case raises issues of phonological or phonetic underspecification, depending on whether we fill in the dependency relation in the phonology at all.

Another example concerns the representation of uvulars. It is typically not the case that all uvulars group with consonants articulated lower in the throat, which raises the possibility that there may be two different representations for this place of articulation: roughly, velarized pharyngeals or pharyngealized velars (often for fricatives and stops, respectively). The distinction could be expressed in terms of a dependency difference, again raising the issues of double representation.

Variable dependency may obviate the need for underspecification of unary features. We can take advantage of the fact that a feature may be present in two ways—as a head or as a dependent—and propose that the mere presence of a feature does not in itself entail that it can be called upon by a spreading process. Rather, a spreading rule can be made to refer only to heads, or only to dependents. Thus, a feature may be phonetically present but fail to spread as a result of its status as either a head or a dependent. Demirdache (1988) argued along these lines that the vowel /i/ in Hungarian cannot spread its frontness if the relevant component is a head rather than a dependent.

6. SUMMARY

We have reviewed six kinds of cases in which different underlying phonological representations eventually receive a nondistinct phonetic interpretation:

(12) Six sources of global determinacy
ia. Old-style abstractness (over- and underapplication):
R1 could receive a phonetic interpretation [[R1]] but is changed into R2 so that we arrive at [[R2]].
ib. Floating material:
R1 and R2 are distinct, but certain information in R1 does not receive a phonetic interpretation.
iia. Phonological underspecification:
R1 is incomplete and nondistinct from R2. [[R1]] is undefined. Hence R1 must be changed into R2, which can be interpreted as [[R2]].
iib. Phonetic underspecification:
R1 is incomplete and nondistinct from R2. [[R1]] is defined and broader than [[R2]].
iiia. Variable grouping:
R1 and R2 are distinct, but [[R1]]=[[R2]].
iiib. Variable dependency:
R1 and R2 are distinct, but [[R1]]=[[R2]].

All these cases raise serious questions regarding learnability, because a lot must be known about the phonology of specific languages in order to determine which representation is adequate. In other words, the phonological representation of many phonetic events cannot be determined by local considerations.

It is of course not the case that all current phonological theories allow all these kinds of global determinacy. Some theories, in fact, allow very little. For example, dependency and government approaches have been described as very concrete in the sense that underlying representations can differ only incrementally from those that become phonetically interpreted, and then only in very limited cases. In section 1 we also noted that recent theories have reinstalled this kind of monotonicity in a form as extreme as in Natural Generative Phonology (NGP). Scobbie's (1992) Declarative Phonology introduces the notion of unification into phonology. Essentially, all expressions (including underlying forms) obey constraints that are surface true. Surface representations are the result of adding up relevant constraints. Abstractness is not allowed in a model of this kind, which, like NGP, appears to entail a purely suppletive approach to feature-changing allomorphy. The child learning in accordance with Declarative Phonology is more interested in getting the facts right than in arriving at generalizations.

The learnability aspects of old-style abstractness (type (i) in (12)) are discussed in Dresher (1981a, 1981b). We do not take a stand here on the degree of abstractness that we would like to allow, nor, in the case of type (ii), the degree and type of underspecification that phonological theory

should incorporate. Learnability issues surrounding type (iii), as we indicated, will be heavily influenced by the position one takes regarding the arity of features. In theories making use of unary features (specifically, theories from the "AIU" family; Anderson & Ewen, 1987; Kaye, Lowenstamm, & Vergnaud, 1985; van der Hulst, 1989, in press), underspecification plays a very modest role. Variability involving empty class nodes would be eliminated if we deny class nodes, or structural information in general, an independent ontological status. Variability due to flexible feature geometries could be removed by developing a more constrained version of feature grouping. Noting that variable dependency requires the language learner to take into account global information, we believe that this concept is so well founded that it seems better to try to solve the learning-path problem than to remove it from the theory.

7. SOME QUESTIONS OF LEARNABILITY

Although phonological theory is changing rapidly, and the various proposals are tentative and in flux, it is nevertheless not too early to inquire more systematically into the learnability of these various approaches. We believe that an attempt to make explicit the learning paths required by whichever proposals one prefers is bound to yield fruitful insights into the degree to which phonological theory is indeed approaching the goal of explanatory adequacy. An inquiry along these lines would ask certain questions.

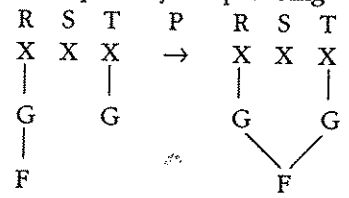
The first question for any particular proposal is (Q1): Does the analysis posit a principle or a parameter? A principle of UG is innate, and independent of experience. A parameter must be fixed on the basis of experience, though the range of possible settings may be given in advance. Not all variation is parametric. Suppose that a principle of UG states that if a language has property I, then it must also have J; otherwise, if it lacks I, it must have K. The choice of I or not I may be parametric, in that a language may freely have I or not. But once that choice has been determined, the choice of J or K comes from UG, and not from experience. Thus, in the case of the laryngeals or coronals, which may or may not have place nodes, the variation is principled if it depends on other aspects of grammar, such as the nature of underlying contrasts in the segment inventory. It is parametric if there is no such determining principle; in such a case, the learner must look for evidence bearing on the setting of the parameter. A similar question can be raised with respect to the other cases discussed earlier. (It should be noted, however, that it is not always easy to distinguish cause and effect in such matters.)

If we are dealing with a parameter, a second question must be addressed (Q2): What is the default (unmarked) parameter setting? And we must also

answer a third question (Q3): What are the cues that trigger a change to the marked setting?

Let us consider the case in which the variability lies in whether or not a segment is specified for a node. If a segment (e.g., S in (13)) lacks a node G, it will be transparent to a process, P, which spreads F on the G tier; if a segment possesses G (e.g., T), it will be affected by P:

(13) Transparency in spreading



Two learning paths are possible in this case. We may take spreading to be the crucial evidence in determining whether or not S or T has G. If the existence of transparency effects is a positive cue, we might assume as default that the node G is present until we see that a segment is transparent to P. Whether G is lacking or not is parametric.

A second route assumes that the geometry of a segment is filled out only as much as is required by contrast, following an algorithm that must be made explicit. In that case, we are really dealing with a principle. Segments are assigned increasingly rich representations as more and more contrasts are learned. At the end of this process, the segments S and T will either have a G node or not; the learning model will never have to look for evidence bearing on this assignment.

The fourth relevant question we see is (Q4): Do cross-parameter dependencies exist? Informal accounts of how a parameter may be set often presuppose that certain parts of the grammar have already been learned. For example, we might say that segment S will be assigned node G if and only if there is an underlying contrast between E and F, dependents of G. But it may not be trivial to determine if the crucial contrast is underlying or derived. Or the spreading process P that supplies the crucial cue as to whether S has G may itself be ambiguous as to which node it is spreading, G or F; the evidence for this may also involve transparency effects, leading to circularity in the learning model.

8. CONCLUSION

We conclude these remarks by considering the implications for learnability of a principle that is currently widely held, to the effect that the main burden of explanation in phonology rests with representations, not rules. We first heard this view expressed some years ago by J. Kaye and J. Lowenstamm, who spoke in terms of projecting the phonology from the segmental inventory. On this

view, the fundamental character of a phonology is determined by its inventory, and the properties of the rules, to the extent there are any, should follow relatively simply from the representations they apply to.

Thinking in terms of acquisition, this view leads naturally, though not inevitably, to a model that is widely assumed, though not often made explicit, and goes something like this: Suppose that a learner begins by learning the segmental inventory, making contrasts only as required. In terms of a feature geometry approach, for example, we might suppose that nodes are added as needed, so as to distinguish contrasting segments. We might call this incremental development. The rules and processes of the phonology apply to the representations so acquired, and facts such as transparency and opacity of segments with respect to processes follow from the representations.

Central to this view are a number of hypotheses that are also often tacit but influential all the same. One is that representations can be established before any rules have been learned, using only the evidence of the inventory (i.e., contrast). A counterexample to this claim would be a case where the representations themselves can be established only on the basis of rules. Such a case can arise, for example, if we have variable representations that must be fixed on the basis of experience. If the relevant experience includes knowing how rules work, then we have to conclude that the representations follow from the rules, and not vice versa.

A related hypothesis is that representations are made increasingly rich in the course of acquisition by adding contrasts, and that one does not go in the other direction, making a contrast that turns out not to be needed and then getting rid of it. Here again the notion of contrast plays a central role. It has not been shown, however, that the notion of contrast that is relevant in any given case can be established in a principled way.

We believe that the preceding remarks demonstrate that modern multilinear phonological theories raise all the same questions regarding learnability as did older SPE-style models (notably, surrounding abstractness) and, in addition, many new ones. Too often, new feature geometries or alternative applications of existing ones are proposed without serious consideration of the learnability issues that are implicitly raised. The main idea behind these comments has been to draw attention to this situation. If the goal of the construction of theories of the language faculty is to explain how it is possible for children to learn grammars on the basis of evidence that largely underdetermines the solution they must come up with, we must be careful about advancing analyses that crucially appeal to global determinacy without specifying a learning path.

ACKNOWLEDGMENTS

These remarks were written as a discussion paper for the GLOW 1991 Workshop on Phonological Acquisition (Leiden, April 1991). The collaboration leading to this paper formed part of the NIAS project on Language Acquisition (Wassenaar, 1990–1991, coordinated by Teun Hoekstra and Harry van der Hulst), with which Elan Dresher was associated as an external participant. The present chapter is a slightly modifed version of the original discussion paper. Elan Dresher would also like to acknowledge the support of grant 410-92-0885 from the Social Sciences and Humanities Research Council of Canada.