Phonology **14** (1997) 83–111. Printed in the United Kingdom
© 1997 Cambridge University Press

*Contrastiveness and faithfulness**

Robert Kirchner
University of Illinois

0 Introduction

A fundamental observation of phonological theory is that, out of the rich sound signal of speech, a small subset of phonetic properties is contrastive in any given language. Standardly, this observation is captured by excluding the non-distinctive phonetic properties from some levels of representation, yielding an abstract categorical representation of the contrasts among the speech sounds. Universally non-contrastive features have generally been assumed to be unspecified throughout the phonological component; thus, for example, Keating (1984) and Lombardi (1991) have argued against inclusion of particular laryngeal features in phonological representation on the grounds that they are never contrastive. Language-particular non-contrastive features have generally been assumed to be unspecified only in underlying representation and early stages of the derivation (e.g. Kiparsky 1982, Archangeli 1984, 1988, Archangeli & Pulleyblank 1986, 1994, Steriade 1987, Clements 1988). Moreover, since the discovery of extensive language-particular gradient alternations,[1] it has standardly been assumed that a subsequent phonetic component is needed, to fill in gradient and universally non-contrastive properties (e.g. Pierrehumbert 1980, Keating 1990; but see Pierrehumbert 1994 for an attack on this view of the phonetics–phonology interface). This model is schematised in Fig. 1.

The crucial property of this model is its representational characterisation of contrastiveness: an abstract representation of pure contrast must be assumed, i.e. underlying representation, with derivations from underlying to surface phonological representation (and perhaps extensive intermediate levels of representation), and from surface phonological representation to the sea of surface phonetic detail. As Steriade (1995a, b) has observed, the assumptions of this model have often been disregarded in practice. For example, the value of the feature [sonorant] is predictable for the classes of [+nasal], [−nasal, −continuant] and [−consonantal] segments; yet the feature defines such an important natural class that phonologists have generally treated it as being present from the earliest

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

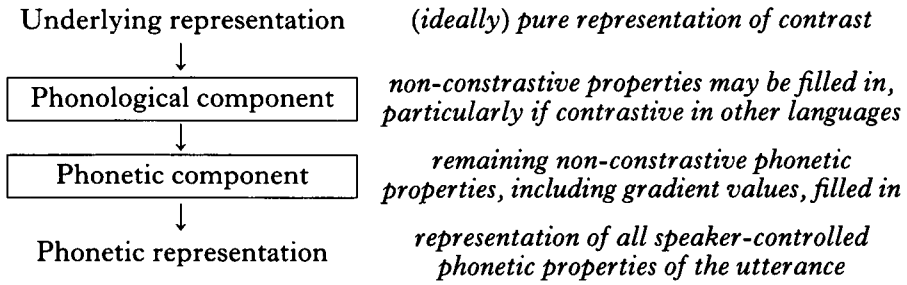


Figure 1
The standard (representational) treatment of contrastiveness

stages of the derivation, even in these classes of segments. Similarly, syllable structure is not contrastive *per se* in any language, and so according to this reasoning ought to be consigned to the phonetic component; yet syllabification is generally considered to be the driving force behind much of phonology. However, I will not dwell on these inconsistencies, and whether they might be reconciled with the model's principal assumptions; rather, I argue that the principal assumptions are superfluous to an adequate treatment of contrastiveness, and therefore should be abandoned.

Note further the all-or-nothing characterisation of contrastiveness in this model. If a phonetic property is admitted to the pantheon of phonological features, it is formally equal to any other feature in its potential for signalling contrasts. Similarly, if a feature is contrastive in a given language, it must be present underlyingly, and there is no representational distinction between features which are contrastive in a broad array of contexts and contrasts which surface only in a particular environment. One may, of course, resort to context-sensitive underspecification (e.g. mid vowels are unspecified for [round] in unstressed syllables), but if there is an independent rule $[-\text{high}] \rightarrow [-\text{round}] / [C_0 __ C_0]_{\sigma(\text{unstressed})}$, or an equivalent constraint, the assumption of underspecification is superfluous (leaving aside the feasibility of having underspecification conditioned by derived properties such as stress). A more accurate observation is that features fall along a continuum of potential contrastiveness. Some distinctions, such as the vowel/consonant distinction, are contrastive in every language; others are contrastive in just a single language, e.g. longitudinal vocal fold tension in Musey voiceless obstruents (Shryock 1995).[2] And of course many phonetic properties, such as the distinction between a 53 ms and 54 ms stop closure, are never contrastive. Similarly, while virtually every feature is subject to some distributional restriction, some features are typically contrastive in extremely narrow contexts. For example, contrastive consonant length is typically restricted to intervocalic position. In the standard approach, constraints or rules may be invoked to neutralise contrasts in particular

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

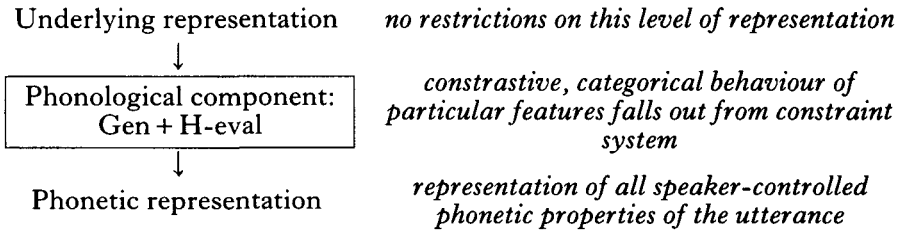


Figure 2
A constraint-ranking treatment of contrastiveness

environments; but if so, we now have two theoretical devices for explaining the absence of contrast: (a) representational restrictions (in UR or throughout the phonology) and (b) rules or constraints requiring the neutralisation of a contrast. One may reasonably inquire whether the latter device is sufficient. Moreover, it is typically the contrasts which are banned outright in many languages which are severely restricted in their distribution in the languages that permit them. For example, Kaun (1994) observes that most languages do not permit a contrast in [round] independent of [back]; but in those languages that do, contrastive rounding is typically subject to vowel harmony, i.e. surface restrictions on the vowels that can occur with it (within some domain). The standard model fails to draw this connection.

I propose an alternative treatment, whereby the notion of contrastiveness emerges from Optimality Theoretic constraint interaction (Prince & Smolensky 1993), specifically with regard to the notion of FAITHFULNESS (minimal divergence of output from input), without resorting to the representational and derivational assumptions of the standard model. This move permits phonological representations to include as much phonetic detail as may be necessary to adequately characterise a speaker's phonetic competence. A significant derivational residuum, namely the post-phonological phonetic component, is thereby eliminated from the theory, as is shown in Fig. 2.

Furthermore, this model captures the observations noted above: a more scalar notion of the potential contrastiveness of a given feature, and the connection between cross-linguistic markedness and restricted distribution of contrasts.

More importantly, this move permits direct reference to phonetic properties which, though never contrastive *per se*, may nevertheless play a role in conditioning phonological processes; we can thereby directly capture phonetic explanations for phonological generalisations which elude formal capture using standard, more phonetically abstract representations. As an illustrative example, consider vowel centralisation in Nawuri, a Kwa language of Ghana (Casali 1995). Short non-back vowels (i ɪ e ɛ) are centralised (i̽ ɨ ə ʌ), except in absolute word-initial or phrase-final position.[3]

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

(1) a. *Non-back vowels centralise, except phrase-finally*

ləmbɨri	'black'
oliŋ	'root'
təkpəri	(type of grass)
ɔ-kɨŋ	'fish'
gɨba: (/gɪ-ba:/)	'hand'
ʧʌmɨnɛ: (ʧɛ-mɪnɛ:/)	'friend'
natɨba (/natɪ ba/)	'walk and come'

b. *Long vowels do not centralise*

bɪ:la	'learn'

c. *Word-initial vowels do not centralise*

ɪ sʌŋ ɪkɨŋ	'The fish (PL) remain'

Similarly conditioned centralisation processes occur in Chamburung, Gichode, Krachi and Gonja (Snider 1989). Rod Casali (personal communication) observes, from both auditory impressions and instrumental measurements, that the 'short' vowels in the blocking contexts are phonetically longer than the vowels in the target contexts, albeit shorter than a truly (i.e. contrastively) 'long' vowel. The extra vowel duration is attributable to cross-linguistically common phenomena of word-initial and phrase-final lengthening (Oller 1973, Klatt 1975). The centralisation process can now be readily understood as articulatory 'undershoot' (Lindblom 1963): the tongue body achieves the peripheral 'front' target only in long *or phonetically lengthened* vowels, when it has enough time to do so without deploying a high velocity fronting gesture.[4] Assume a feature, [partially long], which distinguishes the truly short vowels from the long or phonetically lengthened vowels. By allowing the phonology to refer directly to such a feature, despite its universal non-contrastiveness, the undershoot explanation can be directly captured, e.g. in terms of a rule, [−partially long] → [−front], or a constraint, *[+front, −partially long] (see §5).[5] The alternative, stipulation of a list of contexts where centralisation occurs, is formally inelegant as well as phonetically uninsightful. Note that Nawuri centralisation cannot be relegated to the 'phonetic' component, because, as Casali (1995) demonstates, the centralisation in turn conditions a 'lexical' phonological process of rounding harmony.

This sort of analysis is part of a growing body of research on the role of phonetic factors, including universally non-contrastive and gradient properties, in conditioning phonological processes. Ohala (1983), for example, has shown that the predictable aerodynamic properties of voicing play a large role in explaining the distribution of voiced segments (see also Ohala 1990). Steriade (1993) has shown that the predictable release properties of stops account for their ability to act as contour segments (partially nasalised or affricated). Similarly, Browman & Goldstein (1992) have argued that subphonemic distinctions in degree of overlap among articulatory gestures can explain a variety of assimilation phenomena. Jun (1995) has observed that the relative salience of the

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

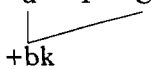
predictable acoustic/auditory consequences of consonantal articulatory gestures can account for certain generalisations concerning the typology of place assimilation (see also Flemming 1995 for an explicit theory of auditory representations in phonology). Hayes & Stivers (1995) have shown that the predictable aerodynamic consequence of velic raising explains the strong cross-linguistic tendency of postnasal stops to be voiced. Moreover, Steriade (1995b) has shown that predictable properties of C–V and V–C formant transitions play a crucial role in explaining the distribution of retroflex consonants. Silverman (1995) shows that a variety of laryngeal and tonal interactions can be explained in terms of predictable consequences of laryngeal gestures; specifically, larngeal specifications are temporally organised so as to achieve salient auditory effects. Similarly, Steriade (1996) shows that a broad variety of laryngeal neutralisation patterns must be understood in terms of a notion of adequate phonetic cues, including non-contrastive properties, such as audible release bursts in the case of glottalised (ejective) obstruents. MacEachern (1996) argues that a variety of laryngeal dissimilation processes refer to a scale of auditory similarity in laryngeal cues. Gordon (1996) shows that the phenomenon of weight- and sonority-sensitive stress (see Kenstowicz 1994) can be understood in terms of a phonetic 'total energy' scale (i.e. the auditory intensity of a given segment times its duration). Finally, Kirchner (in progress) argues that minimisation of articulatory effort is the driving force behind lenition phenomena; therefore, the predictable effort cost of articulatory gestures must be included in the phonological representation.

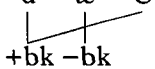
The principal goal of this article is not to present further arguments for the role of predictable phonetic properties in phonology. Rather, it is to show that this research programme can be pursued, with its attendant phonetic enrichment of phonological representations, while actually *improving* the ability of phonological theory to distinguish the contrastive and categorical behaviour of certain properties from the behaviour of the remainder of properties within the sound system.

1 Smolensky's challenge to the standard model

Certain aspects of the standard model, particularly the assumptions surrounding underspecification theory, have been challenged since the development of Optimality Theory (Itô *et al.* 1993, Smolensky 1993, Inkelas 1994, Steriade 1995). In particular, Smolensky has shown that the phonological inactivity of predictable features may be attributed to rankings of a particular class of constraints, rather than to the absence of such features from the representation at some stage of a derivation. For example, consider the standard underspecification-based treatment of the neutral vowels [i e] in Finnish backness harmony (Kiparsky 1981):

(2) a. u I U b. * u æ U





I = [−low, −round, Øback] vowel
U = [−low, +round, Øback] vowel

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

That is, the non-low unrounded vowel in (2a) is unspecified for [back] at the point in the derivation where the spreading rule applies, and therefore is transparent to spreading of [back] onto the following vowel; whereas the [back] specification of non-neutral vowels such as [æ] (2b) makes them opaque. (The failure of [+back] to spread *onto* the neutral vowel is attributed to Structure Preservation, namely the ill-formedness of non-low back unrounded vowels in Finnish.) Assume, however, the Optimality Theoretic notion of faithfulness (minimal departure from identity of input and output), formalised, in part, in terms of feature-specific PARSE constraints (cf. Kirchner 1993, McCarthy & Prince 1995, Orgun 1995). Now the transparency of the Finnish neutral vowels can be analysed in terms of the constraint hierarchy in (3) (modified slightly from Smolensky's presentation):

(3)

	*[−low,−rd, +bk]	ALIGN(+bk-R)	PARSE(bk)	*[−low,+rd, −bk]	*EMBED
a. /u-I-U/					
☞ u-i-u					*
u-i̇-u	*!				
u-i-ü		*!		*	
u-i̇-ü	*!	*		*	
ü-i-ü			*!	**	
b. /u-i-u/					
☞ u-i-u					*
u-i̇-u	*!		*		
u-i-ü		*!	*	*	
u-i̇-ü	*!	*	**	*	
ü-i-ü			*!*	**	
c. /u-i̇-ü/					
☞ u-i-u			**		*
u-i̇-u	*!		*		
u-i-ü		*!	*	*	
u-i̇-ü	*!	*		*	
ü-i-ü			**	*!*	

ALIGN(+bk-R) requires that a [+back] specification be linked to a segment at the right edge of the word (thereby enforcing rightward spreading). *EMBED prohibits embedding of a [−back] domain inside a [+back] domain (i.e. in autosegmental terms, spreading which results in a line-crossing or gapped configuration). The correct outcome is obtained whether we assume an input in which the target vowels are unspecified with respect to [back] (3a), a fully specified input (3b) or even an input whose specifications are *contrary* to the surface values (3c). Since

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

PARSE(back) is ranked below ALIGN(+bk-R), it is better to spread [+back] than to preserve underlying values; and since *[−low, −rd, +bk] is ranked above PARSE(back), back unrounded high or mid vowels are ruled out in all contexts (even if this results in a *EMBED violation). That is, the formal expression of the unmarkedness of front unround vowels is shifted from the representations (i.e. the characterisation of such sounds in terms of fewer underlying features) to the constraints, which directly state the preference for front unround and back round vowels. *The essential observation here is that low ranking of a constraint on faithfulness to a particular feature results in phonologically inert behaviour of that feature within the sound system.* There is no need for restrictions on the presence of particular features at underlying or intermediate levels of representation.

2 A definition of contrastiveness

2.1 Contrastiveness of phonetic representations

At this point, it is necessary to consider afresh precisely what we mean by 'contrastiveness'. Specifically, I wish to develop a characterisation of contrastiveness which is independent of assumptions concerning the systematic presence or absence of particular features or feature values in underlying representation. Pretheoretically, speakers have intuitions that two distinct surface forms (traditionally 'phonetic representations' (PRs)) may or may not count as significantly different, depending on the phonological system, whether or not they are actual forms of the language. For example, [splɪk] and [splɛk] would presumably be considered 'different words' by most English speakers, whereas [splɪk] and [splɪkˀ] (unreleased [k]) would not, though these are all nonce forms. This notion of significant difference, or contrastiveness, appears to require a level of representation distinct from the surface, traditionally 'underlying representation' (UR); and the minimal theoretical assumption required to capture these intuitions of contrastiveness is that two contrastive PRs correspond to distinct URs. However, as illustrated in the previous section, there may be several possible URs for a given surface form under a grammar, and therefore it is necessary to speak of distinct *sets* of URs for some pair of PRs:

(4) *Definition of contrastiveness of PRs*
Two distinct PRs p and p′ are contrastive under grammar G iff the set of URs for p under G is not identical to the set of URs for p′ under G.

Thus, p and p′ are contrastive in (5a) and (b) below, but not in (c) or (d):

(5)

a. u → p u′ → p′	b. u → p u′ → p′	c. u → p p′	d. u → p u′ → p′	e. u → p u′

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

In (5e) there is no contrast, trivially, since there is no distinct pair of PRs: this corresponds to obligatory neutralisation to some surface value (e.g. final laryngeal neutralisation in Korean, where a form such as [kap] could in principle derive from /kap/, /kap^h/ or /kap'/). Note that in the optional neutralisation case (5b), contrastiveness obtains, even though p and p′ have a common UR (u′), since the *sets* of URs for each PR are not identical (e.g. the [w] ~ [ʍ] distinction in certain dialects of English, where e.g. *witch* is invariantly [wɪʧ], but *which* [ʍɪʧ] freely varies with [wɪʧ]).

2.2 Contrastiveness of features

However, phonological grammars refer not to particular forms, but to features and contexts; and therefore we require some notion of contrastiveness as a systematic property of features within sound systems, as well as a property of pairs of forms. 'Feature' is intended here in the broadest possible sense: it encompasses *any* property of a phonological representation, including relations among features, such as prosodic affiliation, and temporal precedence.[6]

Naively, a feature is contrastive just in case it is sufficient to distinguish a pair of contrastive PRs (analogous to the 'minimal pair' criterion of structuralist phonemics).

2.2.1 *Mutually predictable features.* A complicating case, however, is presented when two or more features are mutually predictable. For example, if all sonorants in a sound system are voiced and all voiced sounds are sonorant, there will be no output pair which differs solely with respect to [sonorant], nor solely with respect to [voice]. By the naive definition, *neither* feature is contrastive; whereas we wish to say that one of the features is contrastive, namely the one which corresponds to an underlying minimal distinction ([sonorant]), and the other feature ([voice]) is predictable from it.

(6) u ([−son, −voice]) → p ([−son, −voice])
u′ ([+son, −voice]) → p′ ([+son, +voice])

Note that, under this treatment, the contrastive feature in (6) cannot be determined solely by inspection of the surface forms. Rather, identification of the contrastive feature is relative to the grammar in question: in particular, to the input–output mappings permitted by that grammar. (I leave aside the distinct question of whether we can non-arbitrarily decide between alternative grammars which generate the same surface forms. As a practical matter, it may be necessary in such cases to speak of contrastive feature *sets*: that is, the set {[sonorant], [voice]} is contrastive, without forcing a choice between [sonorant] and [voice].)

2.2.2 *Displaced contrasts.* A further problem is posed by displaced contrasts. For example, in certain dialects of Basque, /e/ raises to [i], and /i/ raises to [i̝][7] when followed by another vowel (e.g. [seme] 'son', [semie] 'the son' *vs.* [eri] 'village', [eri̝e] 'the village'; Hualde 1991). That is, an

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

underlying distinction in [high] is neutralised under hiatus; but there is a surface distinction in [superhigh].

(7) u ([+high, ?superhigh]) → p ([+high, +superhigh])
u ([−high, ?superhigh]) → p ([+high, −superhigh])

We should not, however, treat [superhigh] as a contrastive feature in Basque, although we can find surface minimal pairs, since it is predictable from the underlying value of [high]. On the other hand, we should not treat [high] as being contrastive *in this context*, since there is no surface [high] distinction. However, all cases of displaced contrast involve features which are directly contrastive in some other context in the sound system.[8] In Basque for example, [high] *is* contrastive by virtue of its behaviour in non-hiatus contexts, where it surfaces directly; therefore the URs in (7) are underlyingly distinguished by a contrastive feature. Indeed, it is precisely the contrastiveness of [high] in pairs such as [seme] and [eri] that motivates the identification of [high] as the underlying distinction in the morphologically derived pair [semie] and [erie]. More generally, we need not say that the contrastiveness of a feature F is established by F's behaviour in contexts where an underlying F distinction is displaced: rather, if F is contrastive, this is so because of its behaviour in other contexts. By excluding displaced contrasts from the definition, we are able to provide a formal characterisation of contrastiveness in terms of constraint interaction, while keeping this task distinct from the non-trivial problem of opacity (i.e. the interactions which give rise to displaced contrasts) in non-serial Optimality Theory.

2.2.3 *Free variation.* To handle the complicating cases, then, we must require that a surface F distinction between p and p′ corresponds to a minimal underlying F distinction with respect to u and u′. This condition brings us close to an adequate definition, but fails under a particular circumstance, namely neutralising free variation, shown schematically in (8):

(8) [+F] u → p [+F]
[−F] u′ → p′ [−F]

In this case, if we pick {⟨u,p⟩, ⟨u′,p′⟩} as the crucial pair of input/output mappings, we wrongly claim that F is contrastive, when plainly it is not. To exclude this scenario, we must attach a further condition, namely that one of the URs cannot map to one of the PRs.

2.2.4 *The definition.* We can now define contrastiveness of features as follows:

(9) *Definition of contrastiveness of a feature*
A feature F is contrastive under grammar G iff there is some pair of UR-to-PR mappings {⟨u,p⟩, ⟨u′,p′⟩} under G such that
a. u and u′ differ solely with respect to some specification for F, *and*

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

b. p and p′ differ with respect to a corresponding specification for F, *and*

c. u′ cannot map to p under G.

(Recall that the URs and PRs under discussion are possible, not necessarily actual forms.) I claim, then, that (9) corresponds to an adequate definition of contrastiveness of features. The fact that p and p′ are contrastive PRs follows from condition (c): u′ cannot map to p, therefore p and p′ have distinct sets of URs, and so by definition (4) they are contrastive. To review the complicating cases:

(i) The treatment of mutually predictable features depends on the input/output mappings permitted by a particular grammar, thus the feature [sonorant] in (6) meets the definition (while [voice] does not), since the surface [αsonorant, αvoice] distinction corresponds to an underlying distinction in [sonorant] alone.

(ii) In cases of displaced contrasts, neither the underlying nor the surface distinction is treated as contrastive by virtue of its behaviour in the displaced context, thus neither [superhigh] nor [high] in (7) meets this definition, since the surface distinction corresponds to an underlying distinction in a different feature (though [high] is contrastive by virtue of its behaviour in other contexts).

3 Faithfulness

Before demonstrating the connection between faithfulness to a particular feature and contrastiveness of that feature, we must make explicit our notion of featural faithfulness. McCarthy & Prince (1995) and Orgun (1995) present a generalised notion of faithfulness, defined in terms of correspondence between segments in an input and output, base and reduplicant, etc. That is, a class of faithfulness constraints requires each segment in an input to have a corresponding segment (indicated by coindexation) in the output. A further class of faithfulness constraints requires input and output correspondents to be identical with respect to specific features. I adopt this general conception of faithfulness here. However, the reference to segments is not crucial: we could instead speak of direct correspondence between the features (since segments are ultimately sets of features). Since we are concerned with the general behaviour of features here, this move simplifies the discussion considerably. Assuming a monostratal grammar, 'Input/Output' is equivalent to 'UR/PR', and we can define faithfulness between a UR and PR with respect to some feature F as follows:

(10) PRESERVE(F,INPUT/OUTPUT) (PRES(F,I/O))

For all $\alpha \in \{+,-\}$, for each αF specification in the input there is exactly one corresponding αF specification in the output, and for each αF specification in the output there is exactly one corresponding αF specification in the input.

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

That is, all and only the following mappings with respect to F satisfy PRES(F,I/O):

(11) $+F_i \rightarrow +F_i \qquad -F_i \rightarrow -F_i \qquad \emptyset F \rightarrow \emptyset F$

Privative F can be equated with +F above, assuming that no −F is possible. The class of PRES(F,I/O) constraints therefore subsumes all notions of input/output faithfulness, whether featural, segmental or prosodic.[9] I further assume that no constraint other than PRES(F,I/O) refers to the value of F in UR. As Orgun (1995) notes (extending an observation of Lakoff 1993), a parallel Optimality Theoretic model which permits context-sensitive constraints to refer to underlying properties has all the (presumably excessive) descriptive power of a serial framework, like Chomsky & Halle (1968), with an unlimited number of intermediate levels of representation (but see McCarthy 1995).

A final point concerns the treatment of phonological exchange, i.e. $\alpha F \rightarrow -\alpha F$. Such mappings meet the definition of contrastive features (9), since the definition merely requires a surface distinction between p and p′ corresponding to an underlying minimal distinction between u and u′, even if the surface and underlying values are flipped. However, since we admit of no input/output 'anti-faithfulness' constraints with respect to F, such mappings cannot arise under any grammar. This prediction seems correct (see Janda 1987). Nevertheless, there are a few cases of morphologically conditioned exchanges, e.g. Dinka plural formation (Kenstowicz & Kisseberth 1979): if the base (singular) is [αlong], the derived (plural) form is [$-\alpha$long]. Such alternations can be handled in terms of morphologically conditioned constraints which require some feature in the surface form of the base to map to the opposite value in the morphologically derived form (i.e. morphologically conditioned output/output 'anti-faithfulness'), without referring to the value of the feature in UR. (Henceforth, I use PRES(F) as shorthand for PRES(F,I/O), since we will not discuss output/output faithfulness (or anti-faithfulness) constraints further.)

4 A language-specific predictable property: aspiration in English

To illustrate the connection between PRES(F) and the contrastiveness of F in a sound system, let us consider aspiration of stops in English. As a descriptive approximation, I posit the constraint in (12):[10]

(12) ASPIRATE (ASP): A stop is [+spread glottis] iff it is [−voice], occurring in initial position in a stressed or word-initial syllable.

The English pattern is obtained under the ranking in the tableaux in (13):

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

(13)

	Asp	Pres(spread)
a. pɪl → 'pɪl	*!	
☞ pɪl → 'pʰɪl		*
b. pʰɪl → 'pɪl	*!	*
☞ pʰɪl → 'pʰɪl		
☞ c. spɪl → 'spɪl		
spɪl → 'spʰɪl	*!	*
☞ d. spʰɪl → 'spɪl		*
spʰɪl → 'spʰɪl	*!	

I assume that Pres(voice), the stress assignment constraints, etc. are all ranked above Pres(spread), therefore candidates ['bɪl], [pɪl] (unstressed), etc. are ruled out. Tableaux (13a, b) show that, regardless of underlying specification for [spread], a voiceless stop in initial position within a stressed syllable is aspirated on the surface. Tableaux (13c, d) show that, regardless of underlying specification for [spread], a voiceless stop in a non-syllable-initial position (more generally, a stop in any environment other than initial in stressed syllable) is realised as unaspirated. Moreover, we obtain the same result if the stop is unspecified for [spread]:

(14)

	Asp	Pres(spread)
a. Pɪl → 'pɪl	*!	*
☞ Pɪl → 'pʰɪl		*
☞ b. sPɪl → 'spɪl		*
sPɪl → 'spʰɪl	*!	*

Under this ranking then, for any pair of URs which differ solely with respect to [spread], e.g. /pʰɪl/, /pɪl/ (or /Pɪl/), the PR neutralises to a particular value of [spread]: [+spread] in the aspiration environment and [−spread] elsewhere. Therefore, [spread] is not contrastive under this grammar, by definition (8), at least for the class of stops.

If, however, Asp is ranked below Pres(spread) (and there is no other higher-ranking constraint on the distribution of [spread] in voiceless stops in this context), then [spread] is contrastive, as in Hindi:

(15)

	Pres(spread)	Asp
☞ a. pi → 'pi		*
pi → 'pʰi	*!	
b. pʰi → 'pi	*!	*
☞ pʰi → 'pʰi		
c. Pi → 'pi	*	*!
☞ Pi → 'pʰi	*	

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

URs /pi/ and /p^hi/ are solely distinguished by [spread], and the value of [spread] in their respective PRs, ['pi] and [p^hi], matches the underlying value; therefore, by definition (8), [spread] is contrastive under this grammar.

Finally, consider a grammar where PRES(spread) and ASP are unranked:

(16)

	PRES(spread)	ASP
☞ a. pi → 'pi		*
☞ pi → 'p^hi	*	
b. p^hi → 'pi	*!	*
☞ p^hi → 'p^hi		
c. Pi → 'pi	*	*!
☞ Pi → 'p^hi	*	

UR /pi/ freely varies between ['pi] and ['p^hi], while /p^hi/ is invariantly realised as ['p^hi], a pattern attested in Ao (Gurubasave-Gowda 1975). The PRs for /pi/ are ['pi] and ['p^hi]; and the PR for /p^hi/ is ['p^hi]. That is, a [+spread] UR is invariantly realised as surface [+spread], and a [−spread] specification may be realised as [−spread]; therefore, by definition (8), [spread] is contrastive under this grammar.

Observe that the contrastive or predictable status of aspiration in the foregoing tableaux depends on the satisfaction or violation of PRES(spread) in the winning candidates, which in turn depends on the ranking of PRES(spread) with respect to the constraint on its distribution. The predictable status of stop aspiration in English (13) in no way depends on the feature's absence from any level of phonological representation. (This result is straightforwardly translatable into rule-based terms: a feature is non-contrastive just in case a rule or set of rules neutralises the underlying distinction in all contexts.)

5 Universally non-contrastive properties: partial lengthening

Next, consider a universally non-contrastive property, such as the partial lengthening which occurs in Nawuri, in word-initial and phrase-final vowels. That is, the distribution of the feature [partially long] is subject to the following constraints:

(17) PHRASE-FINAL LENGTHENING: $V]_{Phrase} \rightarrow$ [+partially long]
WORD-INITIAL LENGTHENING: $[_{Word}V \rightarrow$ [+partially long]
*HALF-LONG: [−long] → [−partially long][11]

The phonological relevance of this property in conditioning centralisation

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

is discussed in §0. To complete the analysis, we merely need to rank *[+front, −partially long] above PRES(front).

(18)

	*[+fr, −p.l.]	PRES(front)
a. o-liŋ → oliŋ	*!	
☞ o-liŋ → oli̱ŋ		*
b. natɪ ba → natɪ ba	*!	
☞ natɪ ba → nata̱ ba		*
☞ c. lembiri → ləmbi̱ri]$_{Phrase}$		
lembiri → ləmbi̱ri]$_{Phrase}$		*!
☞ d. bɪ:la → bɪ:la		
bɪ:la → bɛ̱:la		*!

Let us assume that UG contains no corresponding faithfulness constraint, PRES(partially long) (analogous to the standard assumption that [partially long] is unspecified in the phonology). The value of [partially long] is therefore determined by the constraints in (17), regardless of underlying specification for [partially long]. (V· = a [+partially long] vowel, V = a [−partially long] vowel.)

(19)

	PHRASE-FINAL	WORD-INITIAL	*HALF-LONG
☞ a. [...V...]$_{Wd}$ → [...V...]$_{Wd}$			
[...V...]$_{Wd}$ → [...V·...]$_{Wd}$			*!
b. [$_{Wd}$V → [$_{Wd}$V		*!	
☞ [$_{Wd}$V → [$_{Wd}$V·			*
c. V]$_{Phr}$ → V]$_{Phr}$	*!		
☞ V]$_{Phr}$ → V·]$_{Phr}$			*
☞ d. [...V· ...]$_{Wd}$ → [...V...]$_{Wd}$			
[...V· ...]$_{Wd}$ → [...V·...]$_{Wd}$			*!
e. [$_{Wd}$V· → [$_{Wd}$V		*!	
☞ [$_{Wd}$V· → [$_{Wd}$V·			*
f. V·]$_{Phr}$ → V]$_{Phr}$	*!		
☞ V·]$_{Phr}$ → V·]$_{Phr}$			*

As shown in (19), for any pair of URs which differ solely with respect to [partially long], the distinction neutralises in PR; the surface value of [partially long] is conditioned by the position of the vowel within the word or phrase. Therefore by (8), [partially long] is not contrastive. Under the opposite ranking, with *HALF-LONG dominating the lengthening constraints, all non-long vowels would be realised as [−partially long] in all contexts; but in no case can the feature behave contrastively.

Once again, the non-contrastive behaviour of the feature in question emerges from the constraint system; we require no assumption that such

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

properties are excluded from phonological representation, or from any level of representation within the phonological component. The sole distinction between this case and the discussion of aspiration in the previous section is that [partially long] is non-contrastive under any ranking, due to the lack of a PRES(partially long) constraint; and since rankings are all that distinguish phonological systems in Optimality Theory, this amounts to showing that the feature is non-contrastive universally. We are able to include this phonetic property in the phonological representation, thereby capturing insights into centralisation processes; but we do not generate spurious systems in which [partially long] is contrastive *per se*.

This result does not translate elegantly into a rule-based framework. We would have to stipulate that *every* language has a rule or set of rules that neutralise this feature in all contexts, contrary to the general view of rules as language-specific. Nor is it a simple matter to stipulate these universal neutralisation rules. For example, the distinction between released and unreleased stops is universally non-contrastive, though stop release is phonologically relevant in licensing contour segments (partially nasalised or affricated) (Steriade 1993). Nevertheless, this feature neutralises (prepausally) to [+released] in French and [−released] in Korean, and is in free variation in English. In a rule-based framework, we must then posit three distinct neutralising rules for the three languages; and there is no unified formal expression of the feature's non-contrastiveness. In the OT formalism, however, what unifies the three cases is the lack of a PRES (released) constraint in UG; and what distinguishes them is the ranking (or non-ranking) of conflicting constraints on the surface value of [released], such as 'stops must be released' *vs.* 'coda stops must be unreleased'.

6 Generalising the result: the Contrastiveness Theorem

The ability to characterise the predictable or contrastive status of features in terms of the interaction of PRES(F) constraints and the rest of the constraint system (either under particular rankings or universally), rather than in terms of representational restrictions, is not limited to the cases just considered, but rather is fully general.

(20) *The Contrastiveness Theorem*

For all features F, F is contrastive under grammar G iff

a. there is a constraint PRES(F) in UG, *and*
b. there is some PR p such that for any UR u, if the mapping u→p is allowed under G, the mapping satisfies PRES(F) with respect to some F specification in u or p, if any.

(Recall that the URs and PRs under discussion are possible, not necessarily actual, forms.) To prove (20), it is sufficient to show that (i) if both the conditions in (20) hold, F is contrastive, and (ii), if either of the conditions in (20) does not hold, F is not contrastive.

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

(i) Assume that both the conditions in (20) hold: there is a constraint PRES(F), which is unviolated in the mapping between p and all its possible URs with respect to some F specification in p, if any. Therefore, by the definition of PRES(F) (10), for some F specification in p, u must have a corresponding identical F specification, or if p has no F specification, then neither does u. Let u′ be a UR which differs from u solely with respect to F. That is:

(21)

if p contains	*then u contains*	*and u′ contains*
$+F_i$	$+F_i$	$-F_i$ *or* $\emptyset F$
$-F_i$	$-F_i$	$+F_i$ *or* $\emptyset F$
$\emptyset F$	$\emptyset F$	$+F_i$ *or* $-F_i$

Since u′ → p would violate PRES(F), u′ cannot map to p. We next show that the PR for u′ must differ from p with respect to F. For some $\alpha \in \{+, -, \emptyset\}$:

(22)

	Other higher or equally ranked constraints	PRES(F)
☞ a. i. $U\alpha FV$ (= u) → $W\alpha FX$ (= p)	<*a*>	
ii. $U\alpha FV$ (= u) → $W\neg\alpha FX$ (= p'_1)	<*b*>	*
iii. $U\alpha FV$ (= u) → $Y\alpha FZ$	<*c*>	
iv. $U\alpha FV$ (= u) → $Y\neg\alpha FZ$ (= p'_2)	<*d*>	*
☜ b. i. $U\neg\alpha FV$ (= u′) → $W\alpha FX$ (= p)	<*a*>	*
ii. $U\neg\alpha FV$ (= u′) → $W\neg\alpha FX$ (= p'_1)	<*b*>	
☜ iii. $U\neg\alpha FV$ (= u′) → $Y\alpha FZ$	<*c*>	*
iv. $U\neg\alpha FV$ (= u′) → $Y\neg\alpha FZ$ (= p'_2)	<*d*>	

U–Z denote any phonological material. (b.i) and (b.iii) are necessarily losing candidates. ⟨*a*⟩–⟨*d*⟩ show dependencies between the presence of a violation mark in tableaux (a) and (b), depending on assumptions concerning the other constraints. These dependencies follow from the assumption that no constraint other than PRES(F) refers to the underlying value of F: since the URs in (a) and (b) are otherwise identical, all constraints violated by a given PR in (a) must be violated by the same PR in (b), and *vice versa*. ⟨*a*⟩ corresponds to no worse violation marks than are present in ⟨*c*⟩, otherwise candidate (a.i) would lose to (a.iii), contrary to our assumption that p is an output for u. We have already shown that u′ cannot map to p, i.e. candidate (b.i) necessarily loses. Moreover, since the violations in ⟨*a*⟩ are no worse than those in ⟨*c*⟩, (b.iii) must lose as well. Therefore the PR for u′ (either p'_1 or p'_2) differs from p with respect to F. There is therefore a pair of PRs, p and p′, which differ with respect to F; among the inputs to p and p′ are u and u′, respectively, which differ solely with respect to F; and u′ does not map to p. By definition (8), F is contrastive under G. This result is exemplified by the rankings of PRES(spread) and ASP for Hindi (15) and Ao (16).

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

(ii) Assume that either of the conditions in (20) do not hold. There is no PRES(F) constraint in UG, or there is no PR p such that all URs for p satisfy PRES(F).

(23)

	Other constraints	PRES(F)
☞ a. $U\alpha FV\,(=u) \rightarrow W\alpha FX\,(=p)$	<*a*>	
$U\alpha FV\,(=u) \rightarrow W\neg\alpha FX\,(=p')$	<*b*>	*
☞ b. $U\neg\alpha FV\,(=u') \rightarrow W\alpha FX\,(=p)$	<*a*>	*
$U\neg\alpha FV\,(=u') \rightarrow W\neg\alpha FX\,(=p')$	<*b*>	

p′ may also be a winner, in free variation with p, if the violations in $\langle b \rangle$ are equal to $\langle a \rangle$ and there is no PRES(F) constraint. In either case, since PRES(F) can be violated, no other constraint prevents u and u′ from both mapping to p. Since u′ can map to p, by definition (6), F is not contrastive under G. This result is exemplified by the universal status of [partially long] (19), and by the rankings of PRES(spread) and ASP for English (13, 14).

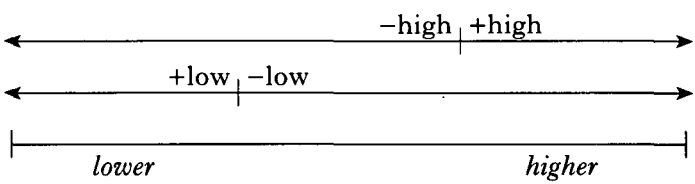
Therefore, the conditions in (20) are both necessary and sufficient to show that a feature is contrastive.

The assumption of strongly parallel, one-step UR-to-PR mapping is not crucial to this result. In a multi-stratal grammar, F is contrastive just in case (20) (substituting 'input/output' for 'UR/PR') holds true for F at each stratum. We have shown that, on a single round of evaluation, an underlying distinction in F maps to an output F distinction just in case (20) holds with respect to F. In a multi-stratal OT grammar, this output is then taken as the input for another round of candidate generation and evaluation. But if (20) characterises the behaviour of F on the second stratum as well, the same result obtains, and so on, regardless of the number of strata.

7 Categorical effects with continuous representations

It is standardly claimed that phonological representations are categorical, whereas phonetic representations are gradient. Consider a phonetic dimension such as vowel height. For purposes of phonological analysis, this phonetic continuum is standardly divided into three regions, low, mid and high, which are formally represented in terms of two binary features: [low] and [high], as shown in (24):

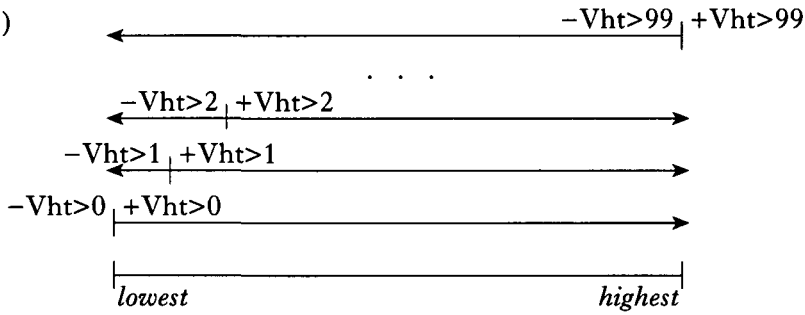
(24) *Vowel height continuum subdivided using two binary features*



https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

Clearly, the claim of phonological categoricalness cannot mean that there is at most a binary distinction for any phonetic dimension; for at least a ternary distinction in vowel height is required.[12] If, however, the claim of phonological categoricalness is that the phonology represents phonetic dimensions in terms of *some* number of discrete, binary features, then the distinction between categorical and gradient representations is vacuous, since this technique can be applied recursively to yield a quasi-continuum. Let the vowel height dimension be subdivided into, say, 100 features of the form [±vowel height (Vht) > *n*], as shown in (25):

(25)



This scale is certainly fine enough for any linguistic phonetic analysis, but if a closer approximation to a true continuum were required, we could simply subdivide the dimension into an even greater number of binary distinctions, such that the increments approach infinitesimality.[13]

If there is any truth to the claim of phonological categoricalness, it lies in the observation that a small number of points along the phonetic dimension will be contrastive in any given language. For example, we do not find languages which have anything approaching 100 contrastive degrees of vowel height. But this is simply a special case of distinguishing between contrastive and non-contrastive properties, which can be adequately handled in terms of the constraint system, without representational restrictions, as shown in the preceding sections.[14] That is, the categorical behaviour of phonological objects emerges from the constraint system, rather than by reifying the categories in an abstract representation.

To make this clear, let us adopt a quasi-continuous representation of vowel height, as in (25), to describe a language with a three-way distinction of vowel height. We simply need to identify the two features which serve as the boundaries between mid and non-mid vowels in the relevant language: for the sake of concreteness, say [Vht > 33] and [Vht > 67]:

(26) [−Vht > 33] = [+high] = high
[+Vht > 33, −Vht > 67] = [−low, −high] = mid
[+Vht > 67] = [+low] = low

To capture the three-way contrast, it suffices to assume that PRES(Vht > 33) and PRES(Vht > 67) are ranked such that the conditions of (20) hold,

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

i.e. for some set of mappings between a PR and all of its URs, the mappings satisfy PRES(Vht > 33) and PRES(Vht > 67); and the conditions of (20) do not hold for the remaining possible distinctions in F1 frequency, e.g. [Vht > 60], either because there is no PRES(Vht > 60) constraint in UG, or because this constraint is inactive due to low ranking.[15]

A class of constraints which might render certain faithfulness constraints inactive in all contexts are the polarisation constraints, which favour maximal dispersion of the values within the relevant phonetic dimension (cf. Liljencrants & Lindblom 1972, Flemming 1995), thereby enforcing a sort of phonetic categoricalness:

(27) POLAR(Vht,binary): Vht = 0 or 100
POLAR(Vht,ternary): Vht = 0, 50 or 100
POLAR(Vht,quaternary): Vht = 0, 33, 67 or 100
etc.

Under the ranking shown in (28), only three surface vowel heights are possible (0, 50 or 100), regardless of underlying values of vowel height, notwithstanding the presence of a constraint PRES(Vht > 60) within the hierarchy.[16]

(28)

	POLAR (tern)	PRES (Vht>33)	PRES (Vht>67)	PRES (Vht>60)	POLAR (bin)
a. Vht = 84					
Vht = 0		*!	*	*	
Vht = 50			*!	*	*
Vht = 60	*!				*
☞ Vht = 100					
b. Vht = 61					
Vht = 0		*!		*	
☞ Vht = 50				*	
Vht = 60	*!			*	*
Vht = 100			*!		*
c. Vht = 29					
☞ Vht = 0					
Vht = 50		*!			*
Vht = 60	*!	*			*
Vht = 100			*!	*	*

Of the three surface values permitted by POLAR(Vht,ternary), the one assigned to a particular UR value is the one which crosses no boundary, i.e. violates no active PRES constraint. For example, in (28a), mapping from an underlying value of 84 to a surface value of 100 does not cross the 33 or 67 boundaries enforced by the active constraints PRES(Vht > 33) and

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

PRES(Vht > 67); if, however, the UR is lower than 67, as in (28b), the surface value which crosses no boundary is 50 Hz (though it violates the inactive PRES(Vht > 60)). Again, the ternary surface distinction is attributable entirely to the constraint system; it does not depend on the exclusion of the non-contrastive vowel height distinctions from phonological representation.

Ironically, this approach now appears to be *too* categorical, i.e. predicting a narrow set of precise and invariant surface values for vowel height, when in fact we observe a great deal of surface variation: indeed, the whole point of introducing quasi-continuous representations is to capture such gradient variation. However, we can reintroduce variation into the picture in a number of ways.

First, to the extent that the variation is completely *free*, the polarisation constraints, rather than identifying precise values, might specify permissible *ranges* of values:

(29) POLAR(Vht,ternary) (*revised*): Vht = 0–12, 45–53 or 89–100.

Second, *context-sensitive* variation can be captured in terms of conflict between the polarisation constraints and other, context-specific, constraints on vowel height, e.g. articulatory constraints favouring reduced vowels in fast speech:

(30) REDUCE: *[−Vht > 20] (i.e. no very low vowels) in fast speech.

(31)

	REDUCE	PRES(Vht>33)	POLAR(tern)
a. Vht = 15 (fast speech)			
Vht = 0	*!		
☞ Vht = 21			*
Vht = 50		*!	
b. Vht = 15 (normal speech)			
☞ Vht = 0			
Vht = 21			*!
Vht = 50		*!	

The result here is a (somewhat simplified) case of context-sensitive variation in the frequency of Vht, depending on speech rate.[17]

Finally, I do not attempt to spell out here the relation between this gradient but still abstract scale of vowel height and actual physical measures thereof: i.e. frequency of the first formant in Hz (acoustically) or Mels (perceptually), or (roughly speaking) tongue body height in cm (articulatorily). Consequently, there is room for *interspeaker variation* in this relation. If we are simply concerned with modelling the production of a single speaker, it is possible (if non-trivial) to substitute e.g. Hz values for the abstract values, as we shall see in the following section. However,

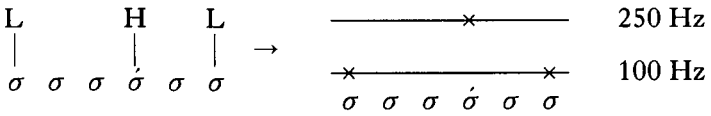
https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

interspeaker variation remains an unsolved problem in this, as in all existing theories of phonology and phonetics.

8 Gradient phonology: interpolation between phonetic targets

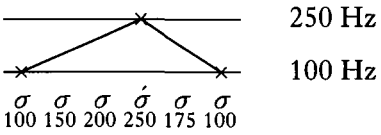
As a further illustration of the extension of Optimality Theoretic phonology to gradient phenomena, consider the device of interpolation between phonetic targets, which has played an important role in models of phonetic implementation. Pierrehumbert (1980) assumes that surface phonological representations may be highly underspecified for particular features, e.g. tone. In the mapping to the phonetic component, the tonal specifications correspond to targets on the F0 frequency dimension.[18]

(32)



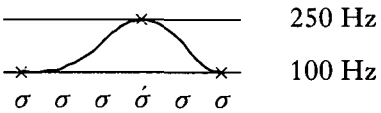
In phonetic realisation, F0 frequency values for the underspecified syllables are filled in by a function which linearly interpolates between the targets:

(33)



Finally, the interpolation lines are 'smoothed', to get rid of abrupt transitions:

(34)



The interpolation device yields results which accord fairly closely with actual measurements over a broad range of phenomena. And though various refinements of this model have since been proposed (e.g. Pierrehumbert & Beckman 1988, Huffman 1989, Cohn 1990, Keating 1990, Choi 1992), the role of interpolation functions in phonetic realisation seems empirically well established, as well as conceptually plausible. That is, any excursion in F0 values presumably corresponds to displacement of some articulators; and smooth, interpolatory articulatory trajectories are presumably less effortful than abrupt, jerky trajectories.

However, this phenomenon may be captured in terms of Optimality Theoretic constraint interaction, without recourse to underspecification, nor to a post-phonological interpretive component:

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

(35) a. L%: Phrase begins and ends with a value around the F0 floor (100 Hz).

b. STRESS-TO-PITCH: The most prominent (i.e. nuclear stressed) syllable of a phrase has a value near the F0 ceiling (250 Hz).

c. LAZY: Minimise articulatory effort.

d. PRES(F0): (stands for the set of PRES constraints for F0 dimension features; since they are all inactive in this case, we need not concern ourselves with the internal ranking of this set, nor with assessing relative violations thereof).

As a first approximation, the articulatory effort involved in producing F0 contours can be equated with peak velocity, measured as the sum of the squares of the difference between each F0 value and the following value in the array:[19]

(36) $E = \sum_{i=1}^{n-1} (F0_i - F0_{i+1})^2$

where n = the number of F0 values in the array.

The differences are squared to get rid of negative values, and to magnify the cost of steep local increases or decreases in F0 frequency. For example, given the array of F0 values in (33), $E = (100-150)^2 + (150-200)^2 + (200-250)^2 + (250-175)^2 + (175-100)^2 = 18750$.

Under the following ranking, we see that the effect of linear interpolation obtains:

(37)

	200 σ	100 σ	100 σ	100 σ́	100 σ	100 σ	L%	STRESS-TO-PITCH	LAZY	PRES(F0)
a.	100 σ	100 σ	100 σ	100 σ́	100 σ	100 σ		*!	0	*
b.	200 σ	100 σ	100 σ	250 σ́	100 σ	100 σ	*!		55000	*
c.	100 σ	100 σ	100 σ	250 σ́	100 σ	100 σ			45000!	*
☞ d.	100 σ	150 σ	200 σ	250 σ́	175 σ	100 σ			18750	*

LAZY is best satisfied by an F0 monotone (candidate a). However, the higher-ranked constraints, L% and STRESS-TO-PITCH, require low initial and final F0 values, and a high F0 value on the nuclear stressed syllable. Among the candidates which satisfy these constraints (c and d), the intonation contour which best satisfies LAZY is (d), in which the F0 values for the intermediate syllables rise or fall linearly towards the next target syllable, just as if these values were filled in by linear interpolation, but without requiring a distinct, post-phonological level of representation.

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

Indeed, this is a standard result of Optimality Theoretic reanalysis: the ranking of violable constraints captures an effect attributed to derivational ordering of operations in rule-based frameworks (see Prince & Smolensky 1993: 23).

Moreover, we could easily adopt a closer approximation to a continuous temporal representation, without changing the substance of the analysis. Assume that the representation is divided into intervals of 10 ms, a fine enough scale for any linguistic phonetic analysis. Rather than having an F0 array with a single value per syllable, we now have a much richer array, with an F0 value for each 10 ms interval. However, LAZY can be evaluated over any array of F0 values. Because each difference is squared, the raw value of E increases with the richness of the array (i.e. the sampling rate); nevertheless, for a candidate set displaying F0 values at a given sampling rate, LAZY correctly identifies the candidate which minimises the steepness of local increases or decreases in F0 frequency.

(38)

syllable [la], sampled every 10 msec.	LAZY
☞ 100-106-112-118-124-130-136-142-148-154-160-166-172-178-184-190-196-202-208-214 lllllllllllllllllllllllllaaa

9 Conclusion

Let us return to the observation that, contrary to the standard approach, contrastiveness is not an all-or-nothing property; that features fall along a continuum of potential contrastiveness. In the approach developed herein, the contrastiveness of a feature follows from the satisfaction of PRES(F), which in the OT framework in turn depends on the ranking of PRES(F) relative to potentially conflicting constraints, namely constraints on the surface distribution of F. The higher the ranking of PRES(F), the more distributional constraints PRES(F) outranks, hence the broader the contexts in which F is contrastive.

Intuitively, the position of PRES(F) in the constraint hierarchy of a grammar for a given speaker corresponds to the degree to which the speaker attends to feature F in the mapping between input and output. Thus, for example, speakers of Hindi presumably attend to stop aspiration distinctions in a way that English speakers do not. Although constraint ranking is generally a language-specific matter, it must be recognised that certain featural distinctions are inherently more salient than others, e.g. [consonantal] (characterised by abrupt, large-scale changes in amplitude) *vs.* [longitudinal vocal fold tension] (principally cued by subtle F0 distinctions in the beginning of a following vowel) mentioned in §0. The notion of inherent salience can be formalised in terms of a set of universal ranking conditions such that for a certain feature F, PRES(F) outranks certain constraints on the distribution of F; or PRES(F) outranks PRES(G) (for some other feature G), reflecting the claim that F is inherently more salient than G (cf. Jun 1995). Features which are inherently highly salient have corresponding faithfulness constraints which are universally highly ranked; while inherently subtler features have lower-ranked faithfulness constraints. And, as discussed in §§5–6, universally non-contrastive features lack a faithfulness constraint altogether.

Moreover, this approach captures the connection between the frequent non-contrastiveness of some feature and its restriction to narrow environments when it is contrastive. The lower the ranking of PRES(F), the more constraints on the distribution of F dominate PRES(F), hence the narrower the contexts in which F is contrastive, and the greater the likelihood that F is not contrastive in any context at all. Central to this treatment of contrastiveness is the Optimality Theoretic notion of faithfulness. The analogue of faithfulness in rule-based frameworks is a mere default state, the absence of neutralisation rules with respect to a particular feature in a given language; thus there is no corresponding formal expression in such a framework of a feature's cross-linguistic tendency to resist neutralisation.

In sum, I have shown that some standard assumptions concerning phonological representation and the phonetics–phonology interface warrant re-evaluation in light of Optimality Theory. Specifically, the motivation for excluding non-contrastive properties from the phonological representation, or any derivational stage therein, evaporates under Opti-

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

mality Theoretic analyses which include feature-specific faithfulness constraints. By the contrastiveness theorem, the contrastiveness of a particular feature depends entirely on whether there is a corresponding faithfulness constraint which is satisfied under some set of mappings, which in turn depends on the position of the constraint within the constraint hierarchy. This result extends even to properties which are not contrastive in any language, if we simply assume that such properties lack a corresponding PRES constraint. Finally, the distinction between categorical and gradient properties, standardly assumed to characterise the difference between phonological and phonetic representation, proves to be a special case of the previous result. Consequently, we may capture the categorical and contrastive *behaviour* of particular phonetic properties (and the predictable or gradient behaviour of the remainder) in terms of constraint interaction, while using *representations* which in principle may contain complete phonetic detail, including gradient properties. Thus, Optimality Theory permits the removal of the representational blinders imposed by the standard treatment of contrastiveness: a move which promises to lead to new insights into the phonetic bases of many phonological phenomena.

NOTES

* I wish to express thanks for helpful comments to Donca Steriade, Bruce Hayes, Rod Casali, Edward Flemming, audiences at Berkeley Linguistics Society and the South-West Optimality Theory workshops, at which earlier versions of this paper were presented, and three anonymous reviewers. All errors are, of course, my sole responsibility.

[1] This includes alternations once thought to be categorical, e.g. final devoicing in German (Port & O'Dell 1985); and place assimilation in casual speech in English (Barry 1985, 1991, Nolan 1992).

[2] See generally Ladefoged & Maddieson (1996) for a surprisingly large number of features which are contrastive in only one or two languages. The list of potentially contrastive features which they document, incidentally, is considerably larger than the feature inventory standardly contemplated by phonologists. Therefore, even if we take potential contrastiveness as a necessary condition for inclusion, a large expansion of the feature inventory beyond the standard 15 or so features is clearly in order.

[3] Casali (1995) characterises the centralisation environment as 'interconsonantal'; however, VV sequences in Nawuri are subject to a set of hiatus-avoiding processes (coalescence and glide formation), which yield long vowels (Casali 1996). Therefore, the interconsonantal condition reduces to the (surface) vowel length condition.

[4] Further supporting the undershoot account, Rod Casali (personal communication) reports that centralisation occurs to a lesser degree when the target vowel follows a coronal consonant, i.e. when the tongue body is already relatively 'front' due to the advancement of the tongue blade for the coronal articulation.

[5] More generally, undershoot can be captured in terms of a family of constraints disfavouring effortful articulations; this idea is deployed in §8 with regard to intonation, and is more fully developed in Kirchner (in progress). Nevertheless, the computation of the effort level feature which conditions the Nawuri centralisation under this approach must refer to phonetic duration, so the phonological relevance of phonetic duration is still established.

[6] That is, if in a given representation feature F temporally precedes feature G, we

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

can treat the relation between F and G as a feature itself: [+precedes(F,G)]. Similarly, if syllable σ dominates F, there is a feature [+dominates(σ,F)]. I am not proposing this as the most perspicuous notation for featural relations: I am merely showing that the notion of feature can encompass all such relations.

[7] Hualde (1991) transcribes this as [iy]; phonetically, the vowel in [iy] is presumably higher than plain [i], due to coarticulation with the glide. I assume that it is the extreme closeness of the vowel which gives rise to the stronger percept of a glide in the transition to the following vowel, rather than *vice versa*.

[8] *Modulo*, of course, analyses using displaced underlying distinctions purely as diacritics marking exceptional forms; Kiparsky (1973) argues persuasively against such use of underlying features.

[9] Prince & Smolensky's (1993) arguments for independent ranking of segmental PARSE and FILL constraints concerned the distinction between languages that break up consonant clusters by inserting a vowel *vs.* deleting a consonant, and between languages that break up vowel sequences by deleting a vowel *vs.* inserting a consonant. These distinctions can be reanalysed in terms of independent ranking of PRES(consonantal) and PRES(vocalic) constraints, with [consonantal] and [vocalic] treated as distinct, privative features. Indeed, Prince & Smolensky (1993: ch. 7) are forced to posit such a distinction in any case (in their terms, PARSE(C) *vs.* PARSE(V) and FILL(Ons) *vs.* FILL(Nuc)).

[10] Phonetic studies (e.g. Pierrehumbert & Talkin 1992) indicate that the degree of aspiration in English actually varies gradiently, depending on the stress level and phrasal position of the relevant syllable. To handle this gradient variation, a more sophisticated, gradient version of the ASP constraint is required (see §7 on gradient properties in the phonological representation); nevertheless, such elaboration does not change the essence of the present analysis, namely conflict between faithfulness and some constraint on the surface distribution of aspiration.

[11] [+long] entails [+partially long] by definition.

[12] Nor is non-binarity unique to vowel height: ternary (or greater) distinctions are also required for the F0 frequency dimension (i.e. tone), and for the 'sonority' dimension (e.g. Clements 1990).

[13] Similarly, the categorical treatments of aspiration and vowel duration in §§4–5 can be recast in gradient terms by adopting more continuous representations of voice onset time and vowel duration, respectively.

[14] This point has also been made by Steriade (1995b) and Flemming (1995), though Flemming's approach relies on a family of constraints which refer explicitly to the maintenance of contrasts (by globally comparing inventories of possible distinct outputs), rather than relying on faithfulness constraints.

[15] A constraint is 'active' on an input if it is satisfied by some candidates and violated by others, and no higher-ranking constraint has already ruled out all of the satisfiers or all of the violators (Prince & Smolensky 1993: 82).

[16] Flemming (1995) claims that polarisation of values only comes into play with respect to contrastive features. This generalisation, if correct, might be captured in terms of a ranking condition such that no constraint may be crucially ranked between the PRES constraints and any active POLAR constraints and any active POLAR constraint for a given phonetic dimension.

[17] More gradient context-sensitive variation can be obtained simply by adopting more sophisticated context-specific constraints: for example, a REDUCE constraint which makes the amount of reduction/raising inversely proportional to duration, e.g. 'vowel height must be no lower than 0 (the 'ideal' minimum vowel height) + k/T (a constant divided by the duration of the vowel)'. Alternatively, reduction (like Nawuri centralisation) can be understood as articulatory undershoot, an effort minimisation effect (see §8 and Kirchner in progress). Further note that register-sensitive variation may also be regarded as a special case of context-sensitive variation: the relevant 'context' includes speech register (adequately formalised), in addition to phonetic properties. For an approach to OT modelling of register-sensitive variation, see Kirchner (1995, in progress).

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

[18] The mapping from tonal specifications to phonetic targets is itself non-trivial, and I will not attempt to analyse this here; however, it seems plausible that adjustments of target values, e.g. due to downdrift and downstep, can likewise be handled by constraint interaction. For example, downdrift can be attributed to a constraint which requires an F0 peak to be lower (by some amount) than all preceding F0 peaks within some prosodic domain.

[19] An anonymous reviewer notes: 'the tacit assumption here is that an F0 monotone would constitute the least effortful thing for a speaker to do. I think that this is probably inaccurate and suggest that sustaining an F0 monotone requires a degree of precision antithetical to the notion of LAZY.' The reviewer is, strictly speaking, correct; the least effortful intonation contour probably involves some declination. However (36) and (39) are merely intended as a rough analogue of articulatory effort; a more sophisticated theory would quantify effort in terms of articulatory gestures directly, using all available techniques of articulatory experimentation. In any case, by drawing a connection between interpolatory movement and effort minimisation, I am not diverging from the assumptions of the standard target-and-interpolation framework: for Pierrehumbert (1980) likewise assumes that effort minimisation lies behind interpolatory movement.

[20] The lower the value of k (for $k \geqslant 0$), the lower the penalty for abrupt transitions, hence the closer the contour to a straight line interpolation.

REFERENCES

Archangeli, D. (1984). *Underspecification in Yawelmani phonology and morphology*. PhD dissertation, MIT.

Archangeli, D. (1988). Extrametricality in Yawelmani. *The Linguistic Review* **4**. 101–120.

Archangeli, D. & D. Pulleyblank (1986). *The content and structure of phonological representations*. Ms, University of Arizona & University of Southern California.

Archangeli, D. & D. Pulleyblank (1994). *Grounded phonology*. Cambridge, Mass.: MIT Press.

Barry, M. (1985). A palatographic study of connected speech process. *Cambridge Papers in Phonetics and Experimental Linguistics* **4**. 1–16.

Barry, M. (1991). Temporal modelling of gestures in articulatory assimilation. *Proceedings of the 12th International Congress of Phonetic Sciences* Vol. 4. Aix-en-Provence: Université de Provence. 14–17.

Browman, C. & L. Goldstein (1992). Articulatory phonology: an overview. *Phonetica* **49**. 155–180.

Casali, R. (1995). Labial opacity and roundness harmony in Nawuri. *NLLT* **13**. 949–663.

Casali, R. (1996). *Resolving hiatus*. PhD dissertation, UCLA.

Choi, J. (1992). *Phonetic underspecification and target interpolation : an acoustic study of Marshallese vowel allophony*. PhD dissertation, UCLA.

Chomsky, N. & M. Halle (1968). *The sound pattern of English*. New York: Harper & Row.

Clements, G. N. (1988). Toward a substantive theory of feature specification. *NELS* **18**. 79–93.

Clements, G. N. (1990). The role of sonority cycle in core syllabification. In Kingston & Beckman (1990). 282–333.

Cohn, A. (1990). *Phonetic and phonological rules of nasalization*. PhD dissertation, UCLA.

Docherty, G. & D. R. Ladd (eds.) (1992). *Papers in laboratory phonology II : gesture, segment, prosody*. Cambridge: Cambridge University Press.

Flemming, E. (1995). *Auditory representations in phonology*. PhD dissertation, UCLA.

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

Gordon, M. (1996). Three- and four-way weight hierarchies: a phonetically driven account of weight. Paper presented at GLOW Workshop on Weight Effects, Athens.
Gurubasave-Gowda, K. (1975). *Ao grammar*. Mysore: Central Institute of Indian Languages.
Hayes, B. & T. Stivers (1995). *The phonetics of post-nasal voicing*. Ms, UCLA.
Hualde, J. (1991). *Basque phonology*. London: Routledge.
Huffman, M. (1989). *Implementation of nasal : timing and articulatory landmarks*. PhD dissertation, UCLA.
Inkelas, S. (1994). *The consequences of optimization for underspecification*. Ms, UCB.
Itô, J., A. Mester & J. Padgett (1993). Licensing and redundancy: underspecification in Optimality Theory. Ms, UCSC.
Janda, R. (1987). *On the motivation for an evolutionary typology of sound-structural rules*. PhD dissertation, UCLA.
Jun, J. (1995). *A constraint-based analysis of place assimilation typology*. PhD dissertation, UCLA.
Kaun, A. (1994). *The typology of rounding harmony*. PhD dissertation, UCLA.
Keating, P. (1984). Phonetic and phonological representation of stop consonant voicing. *Lg* **60**. 286–319.
Keating, P. (1990). The window model of coarticulation: articulatory evidence. In Kingston & Beckman (1990). 451–470.
Kenstowicz, M. (1994). Sonority-driven stress. Ms, MIT.
Kenstowicz, M. & C. Kisseberth (1979). *Generative phonology*. New York: Academic Press.
Kingston, J. & M. Beckman (eds.) (1990). *Papers in laboratory phonology I : between the grammar and physics of speech*. Cambridge: Cambridge University Press.
Kiparsky, P. (1973). Phonological representations. In O. Fujimura (ed.) *Three dimensions of linguistic theory*. Tokyo: TEC. 3–136.
Kiparsky, P. (1981). Vowel harmony. Ms, MIT.
Kiparsky, P. (1982). Lexical morphology and phonology. In I.-S. Yang (ed.) *Linguistics in the morning calm*. Seoul: Hanshin. 3–91.
Kirchner, R. (1993). Turkish vowel harmony and disharmony: an Optimality Theoretic account. Paper presented at the Rutgers Optimality Workshop 1.
Kirchner, R. (1995). Contrastiveness is an epiphenomenon of constraint ranking. *BLS* **21**. 198–208.
Kirchner, R. (in progress). *Lenition typology in phonetically based Optimality Theory*. PhD dissertation, UCLA.
Klatt, D. (1975). Vowel lengthening is syntactically determined in a connected discourse. *JPh* **3**. 129–140.
Ladefoged, P. & I. Maddieson (1996). *The sounds of the world's languages*. Oxford: Blackwell.
Lakoff, G. (1993). Cognitive phonology. In J. Goldsmith (ed.) *The last phonological rule*. Chicago: University of Chicago Press. 117–145.
Liljencrants, J. & Lindblom, B. (1990). Numerical simulation of vowel quality systems. *Lg* **48**. 839–862.
Lindblom, B. (1963). Spectrographic study of vowel reduction. *JASA* **35**. 1773–1781.
Lombardi, L. (1991). *Laryngeal features and laryngeal neutralization*. PhD dissertation, University of Massachusetts, Amherst.
McCarthy, J. (1995). Remarks on phonological opacity in Optimality Theory. Ms, University of Massachusetts, Amherst.
McCarthy, J. & A. Prince (1995). Faithfulness and reduplicative identity. In J. Beckman, L. Dickey & S. Urbanczyk (eds.) *Papers in Optimality Theory*. Amherst: GLSA. 249–384.
MacEachern, M. (1996). Laryngeal similarity effects in Quechua and Aymara. *WCCFL* **15**.

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press

Nolan, F. (1992). The descriptive role of segments: evidence from assimilation. In Docherty & Ladd (1992). 261–289.
Ohala, J. (1983). The origin of sound patterns in vocal tract constraints. In P. MacNeilage (ed.) *The production of speech*. New York: Springer. 189–216.
Ohala, J. (1990). There is no interface between phonetics and phonology: a personal view. *JPh* **18**. 153–171.
Oller, D. (1973). The duration of speech segments: the effect of position in utterance and word length. *JASA* **54**. 1235–1273.
Orgun, O. (1995). *Correspondence and identity constraints in two-level optimality theory*. Ms, UCB (available as ROA-62 from the Rutgers Optimality Archive).
Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. PhD dissertation, MIT.
Pierrehumbert, J. (1994). Knowledge of variation. *CLS* **30:2**. 232–256.
Pierrehumbert, J. & M. Beckman (1988). *Japanese tone structure*. Cambridge, Mass.: MIT Press.
Pierrehumbert, J. & D. Talkin (1992). Lenition of /h/ and glottal stop. In Docherty & Ladd (1992). 90–117.
Port, R. & M. O'Dell (1985). Neutralization of syllable-final voicing in German. *JPh* **13**. 455–472.
Prince, A. & P. Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder.
Shryock, A. (1995). *Investigating laryngeal contrasts: an acoustic study of the consonants of Musey*. PhD dissertation, UCLA.
Silverman, D. (1995). *Phasing and recoverability*. PhD dissertation, UCLA.
Smolensky, P. (1993). *Harmony, markedness, and phonological activity*. Paper presented at the Rutgers Optimality Workshop 1.
Snider, K. (1989). *North Guang comparative wordlist*. Legon: Institute of African Studies.
Steriade, D. (1987). Redundant values. *CLS* **23:2**. 339–362.
Steriade, D. (1993). Closure, release, and nasal contours. In M. Huffman & R. Krakow (eds.) *Nasals, nasalization, and the velum*. Orlando: Academic Press. 401–470.
Steriade, D. (1995a). Underspecification and markedness. In John Goldsmith (ed.) *The handbook of phonological theory*. Cambridge, Mass. & Oxford: Blackwell. 114–174.
Steriade, D. (1995b). Neutralization and the expression of contrast. Ms, UCLA.
Steriade, D. (1996). *Licensing laryngeal features. UCLA Working Papers in Phonology* **1**.

https://doi.org/10.1017/S0952675797003291 Published online by Cambridge University Press