Cambridge Studies in Linguistics 121

The Contrastive Hierarchy in Phonology

B. Elan Dresher

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521889735

This page intentionally left blank

THE CONTRASTIVE HIERARCHY IN PHONOLOGY

'Contrast' – the opposition between distinctive sounds in a language – is one of the most central concepts in linguistics. This book presents an original account of the logic and history of contrast in phonology. It provides empirical evidence from diverse phonological domains that only contrastive features are computed by the phonological component of grammar. It argues that the contrastive specifications of phonemes are governed by language-particular feature hierarchies. This approach assigns a key role to abstract cognitive structures, challenging contemporary approaches that favour phonetic explanations of phonological phenomena. Tracing the evolution of the hypothesis that contrastive features play a special role in phonology, it shows how this insight has been obscured by misunderstandings of the role of the contrastive feature hierarchy. Questioning the widely held notion that contrast should be based on minimal pairs, Elan Dresher argues that the contrastive hierarchy is indispensable to illuminating accounts of phonological patterning.

B. ELAN DRESHER is Professor in the Department of Linguistics at the University of Toronto. His recent publications include *Formal Approaches to Poetry: Recent Developments in Generative Metrics* (ed., with Nila Friedberg, 2006) and *Contrast in Phonology: Theory, Perception, Acquisition* (ed., with Peter Avery and Keren Rice, 2008).

In this series

80 ANDREA MORO: *The raising of predicates: predicative noun phrases and the theory of clause structure*
81 ROGER LASS: *Historical linguistics and language change*
82 JOHN M. ANDERSON: *A notional theory of syntactic categories*
83 BERND HEINE: *Possession: cognitive sources, forces and grammaticalization*
84 NOMI ERTESCHIK-SHIR: *The dynamics of focus structure*
85 JOHN COLEMAN: *Phonological representations: their names, forms and powers*
86 CHRISTINA Y. BETHIN: *Slavic prosody: language change and phonological theory*
87 BARBARA DANCYGIER: *Conditionals and prediction*
88 CLAIRE LEFEBVRE: *Creole genesis and the acquisition of grammar: the case of Haitian creole*
89 HEINZ GIEGERICH: *Lexical strata in English*
90 KEREN RICE: *Morpheme order and semantic scope*
91 APRIL McMAHON: *Lexical phonology and the history of English*
92 MATTHEW Y. CHEN: *Tone Sandhi: patterns across Chinese dialects*
93 GREGORY T. STUMP: *Inflectional morphology: a theory of paradigm structure*
94 JOAN BYBEE: *Phonology and language use*
95 LAURIE BAUER: *Morphological productivity*
96 THOMAS ERNST: *The syntax of adjuncts*
97 ELIZABETH CLOSS TRAUGOTT and RICHARD B. DASHER: *Regularity in semantic change*
98 MAYA HICKMANN: *Children's discourse: person, space and time across languages*
99 DIANE BLAKEMORE: *Relevance and linguistic meaning: the semantics and pragmatics of discourse markers*
100 IAN ROBERTS and ANNA ROUSSOU: *Syntactic change: a minimalist approach to grammaticalization*
101 DONKA MINKOVA: *Alliteration and sound change in early English*
102 MARK C. BAKER: *Lexical categories: verbs, nouns and adjectives*
103 CARLOTA S. SMITH: *Modes of discourse: the local structure of texts*
104 ROCHELLE LIEBER: *Morphology and lexical semantics*
105 HOLGER DIESSEL: *The acquisition of complex sentences*
106 SHARON INKELAS and CHERYL ZOLL: *Reduplication: doubling in morphology*
107 SUSAN EDWARDS: *Fluent aphasia*
108 BARBARA DANCYGIER and EVE SWEETSER: *Mental spaces in grammar: conditional constructions*
109 HEW BAERMAN, DUNSTAN BROWN and GREVILLE G. CORBETT: *The syntax–morphology interface: a study of syncretism*
110 MARCUS TOMALIN: *Linguistics and the formal sciences: the origins of generative grammar*
111 SAMUEL D. EPSTEIN and T. DANIEL SEELY: *Derivations in minimalism*
112 PAUL DE LACY: *Markedness: reduction and preservation in phonology*
113 YEHUDA N. FALK: *Subjects and their properties*
114 P. H. MATTHEWS: *Syntactic relations: a critical survey*
115 MARK C. BAKER: *The syntax of agreement and concord*
116 GILLIAN CATRIONA RAMCHAND: *Verb meaning and the lexicon: a first phase syntax*
117 PIETER MUYSKEN: *Functional categories*
118 JUAN URIAGEREKA: *Syntactic anchors: on semantic structuring*
119 D. ROBERT LADD: *Intonational phonology second edition*
120 LEONARD H. BABBY: *The syntax of argument structure*
121 B. ELAN DRESHER: *The contrastive hierarchy in phonology*

Earlier issues not listed are also available

CAMBRIDGE STUDIES IN LINGUISTICS

The Contrastive Hierarchy in Phonology

THE CONTRASTIVE HIERARCHY IN PHONOLOGY

B. ELAN DRESHER

University of Toronto

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521889735

First published in print format 2009

ISBN-13 978-0-511-64136-7 eBook (NetLibrary)

ISBN-13 978-0-521-88973-5 Hardback

Contents

Acknowledgements

I would like to thank the members of the project on Markedness and the Contrastive Hierarchy in Phonology in the Department of Linguistics at the University of Toronto for creating the context for this study. Keren Rice has been my co-Principal Investigator in this enterprise for more than fifteen years, and her work on contrast-driven complexity and markedness with Peter Avery, our collaborator from the beginning, sparked my interest in investigating the mechanics of contrast and its history in phonological thought.

Many students have participated in the project over the years and it is not possible to name them all here, but some merit special thanks in connection with this book. Zhang Xi first suggested to me, on a plane over the Atlantic some years ago, that I should try to write up a monograph with roughly the contours of the present work. His own research on Manchu-Tungus languages has an important place in this book, and it has been a pleasure collaborating with him. During the writing of this book I have also enjoyed working particularly closely with Daniel Currie Hall and Sara Mackenzie, and I thank them for the many ideas and cases they have contributed to this study.

For comments on earlier drafts I am grateful to Peter Avery, Vincent DeCaen, Chiara Frigeni, Daniel Currie Hall, Manami Hirayama, Yoonjung Kang, Sara Mackenzie, Christine Pittman, Keren Rice, Vanessa Shokeir, Anne St-Amand and three anonymous readers who reviewed the manuscript for Cambridge University Press.

I am grateful to have had the opportunity to teach material related to this study in the Advanced Phonology seminar at the University of Toronto in Fall 2001 and 2004, at the 2003 LOT Summer School in Tilburg, and at a three-day joint mini-course with Keren Rice at the Center for Advanced Study in Theoretical Linguistics (CASTL), University of Tromsø, in September 2007.

I thank the audiences at the following annual meetings: the MOT Phonology Atelier (1994, 1998, 2002, 2003); the CLA (1998, 2002, 2006, 2008); BLS 29 (2003); GALA, Utrecht (2003); ConSOLE XIII, Tromsø (2004); Mayfest, University of Maryland (2006); Lacus Forum (2006); PaPI'07, Braga (2007);

and the LAGB (2007). Some of this material was also presented at the Second International Conference On Contrast in Phonology, Toronto (2002), and at the following workshops: Phonology workshop held at the Meertens Institute Amsterdam (2000); Workshop on Features, Schloß Freudental, Germany (2000); the COLING-2004 Workshop, Geneva; the CUNY Workshop On Phonological Features (2005); and the Inuktitut Linguistics Workshop, Toronto (2008). I also benefited from presenting talks on this material to audiences at CUNY, MIT, McGill, the University of Ottawa, the University of Konstanz, the University of Maryland, Harvard and Cornell.

I would also like to thank colleagues with whom I have discussed issues related to this study, including Nick Clements, Paula Fikkert, Morris Halle, Harry van der Hulst, Larry Hyman, Bill Idsardi, Aditi Lahiri, Andrew Nevins, Glyne Piggott and Jean-Roger Vergnaud.

I am grateful to Andrew Winnard, my editor at Cambridge University Press, for his encouragement and help in bringing this project to completion, and to Leigh Mueller for copy-editing the manuscript. Thanks also to Anne St-Amand for assisting with the indexes and proofreading.

This research was supported in part by grants 410-99-1309, 410-2003-0913 and 410-08-2645 from the Social Sciences and Humanities Research Council of Canada.

I am grateful to my colleagues and students in the Department of Linguistics at the University of Toronto for creating an excellent working environment, and to our former Chair, Diane Massam, for making it possible for me to take a sabbatical leave in 2007–8 to complete this book. I would also like to thank the Principal and staff of University College for providing me with a wonderful place to work during my leave.

I would like to thank Norbert Hornstein for general inspiration, and finally, for her long-suffering support of this project over the years, I am most grateful to Alana Johns.

1 *Introduction*

1.1 Approaches to contrast

The notion of contrast has been central to linguistics since Saussure's famous dictum that 'dans la langue il n'y a que des différences . . . sans termes positifs' ['In a language there are only differences, and no positive terms'] (Saussure 1972 [1916]: 166). 'The sound of a word', according to Saussure, 'is not in itself important, but the phonetic contrasts which allow us to distinguish that word from any other'.[1] That is, a phoneme is identified not only by its positive characteristics – for example, the fact that it sounds like [i] – but also by what it is not – that is, by the sounds it contrasts with.

The notion of contrast can be understood at several different levels. At the most basic level, it can refer simply to whether two sounds contrast in a language or not. In English, for example, [i] is different from [ɪ], and these vowel sounds alone are able to differentiate words in the language: *sheep* [ʃip], for instance, is different from *ship* [ʃɪp]. This contrast recurs in many other word pairs, such as *cheap* ∼ *chip, seat* ∼ *sit, seen* ∼ *sin, meal* ∼ *mill, reed* ∼ *rid*, and so on. Compare this situation with that obtaining in Israeli Hebrew, which has a single phoneme in the part of the vowel space where English has two. This phoneme, which can be represented as /i/, is pronounced somewhere between English /i/ and /ɪ/ (Chen 1972). In final open syllables it may be pronounced more like [i], in closed syllables it may tend more to [ɪ], but these sounds do not serve to distinguish words; that is, they play no contrastive role in the language. Chen (1972: 216) observes that the vowels in the English loanwords *jeep* and *chips* are pronounced the same by many Hebrew speakers; because they fall in between the two English vowel phonemes, English speakers tend to hear them as [ɪ] in *jeep* (as in the first syllable of *gypsum*) and [i] in *chips* (as in *cheap*).

1 'Ce qui importe dans le mot, ce n'est pas le son lui-même, mais les différences phoniques qui permettent de distinguer ce mot de tous les autres' (Saussure 1972 [1916]: 163). The English translation is by Harris (Saussure 1986: 116).

Acquiring the phonological contrasts of a language is one of the more challenging tasks for a language learner, and determining what the contrasts are is a basic aspect of phonological description, and a prerequisite to further analysis. In all the examples that follow I will assume that we know what the contrasts are at this most basic level.[2]

But there is much more to contrast between sounds, and phonologists have traditionally been concerned with further aspects of contrast. One can study the phonetics of contrast to see, for example, how perceptually salient the difference between sounds is. For example, the contrast between [i] and [u] is more perceptible than the contrast between [i] and [ɨ]. It is reasonable to suppose that good contrasts will be favoured in inventories over poor ones (Liljencrants and Lindblom 1972, Flemming 2004), and this fact could have synchronic and diachronic consequences. This is an interesting topic, which I will refer to as 'phonetic contrast', because it is concerned with the surface phonetics of contrasts between sounds.

However, the study of phonetic contrast has not been the central preoccupation of phonologists or phonological theory since Saussure. On the contrary, an influential current of phonological theory – the main stream, for much of the twentieth century – has held that the phonetics do not determine the way sounds pattern in the phonology of a language. In the first issue of *Language*, in the seminal paper that popularized the term 'sound pattern', Edward Sapir wrote: 'And yet it is most important to emphasize the fact, strange but indubitable, that a pattern alignment does not need to correspond exactly to the more obvious phonetic one' (1925).

By 'pattern alignment' Sapir meant the arrangement of the phonemes of a language, their place in the phonological system. I will argue that the pattern alignment of a phoneme is a function of its contrastive properties; hence, according to Sapir and many other phonologists, the contrastive status of a phoneme is not determined by its phonetics. What does determine it? This is the topic of this book: 'phonological contrast' in my terms. Phonological contrast refers to those properties of phonemes that are distinctive in a given phonological system. In most theories of phonology, this means determining which *features* are contrastive and which are redundant.[3]

2 This is not to say that it is a trivial matter to determine what the basic phonemes of a language are, or whether certain contrasts are predictable or must be encoded in the lexicon.

3 Some phonological theories do not posit features in the classical sense, but some other set of primitives. Such primitives are also liable to enter into contrastive relations of the type discussed in this book.

For example, given a language in which there is a contrast between /i/ and /u/, we want to determine, out of the various ways that these sounds differ, which particular dimension is the one most relevant to the phonology of the language. In a theory that posits features like [back] and [round], for example, the question arises whether /i/ and /u/ contrast with respect to one, or the other, or both, of these features.

Jakobson (1962b [1931]) discusses this question with respect to a number of Slavic vowel systems. He cites the observation of B. Hála that, except for a short front vowel ä that occurs in dialects of Central Slovak, the simple vowels of Slovak 'correspond completely both in their production and in the auditive impression they produce to the vowels of Standard Czech'. Jakobson notes (1962c: 224) that the presence of ä in Slovak, though 'a mere detail from a phonetic point of view . . . determines the phonemic make-up of all the short vowels'. The 'phonemic make-up' of a vowel phoneme, like Sapir's pattern alignment, can be equated with its contrastive properties. Jakobson diagrams the Czech and Slovak short vowels as in (1).[4]

(1) Czech and Slovak vowel systems (Jakobson 1962c: 224)

a. Standard Czech

i				u
	e		o	
		a		

b. Standard Slovak

i	u
e	o
ä	a

In Slovak there is a distinction between the low vowels /ä/ and /a/: the former is more front (acute, in terms of Jakobson's features), and the latter is more back (grave). In Czech the low vowel /a/ is not opposed to another low vowel. Therefore, even though the /a/ of Slovak and the /a/ of Czech are phonetically almost identical, Jakobson considers it to pattern as a back vowel in Slovak, whereas in Czech it is neutral with respect to tonality, having no contrastive value except for its height.

The Slovak contrast between the low vowels has consequences also for the status of the non-low vowels, according to Jakobson. Since this contrast does not involve lip rounding (flatness, in Jakobson's features), but only the front/back (acute/grave) dimension, then, presumably by symmetry or feature economy, this distinction may be assumed to hold also of the non-low vowels. That is, Jakobson proposes that the crucial contrast between /i/ ∼ /u/, /e/ ∼ /o/, as well as /æ/ ∼ /ɑ/, is frontness/backness; lip rounding is not a distinctive feature in the Slovak vowel system. In support of this analysis, Jakobson cites the fact that

4 I have inverted and reflected Jakobson's diagrams to the more familiar modern orientation of the vowel space, with high vowels at the top and front vowels to the left.

Central Slovak speakers have little trouble learning to pronounce the French and German front rounded vowels *ü* and *ö*. That is, even though Slovak does not have front rounded vowels, the existence of a front/back contrast independently of lip rounding allows Slovak speakers to combine this dimension with rounding in a new way.

In Czech, tonality is relevant only to the non-low vowels, and Jakobson suggests that the two dimensions of acuteness/gravity and flatness/non-flatness form an 'indissoluble synthesis'. This analysis, he proposes, accounts for the difficulty Czech speakers have in reproducing the German or French front round vowels.

One might think, from these examples, that it is the shape of the inventories alone that determines the nature of the contrastive features in Jakobson's approach. However, Jakobson's further remarks on Russian show that this is not correct. For he observes that Standard Russian has five contrasting stressed vowels, phonetically similar to the five short vowels of Czech. But Russian vowels have front and back allophones determined by neighbouring consonants. Therefore, he proposes that for the Russian non-low vowels flatness (lip rounding) alone is contrastive. In the independence of the tonality features, Russian is more like Slovak than like Czech. The evidence is that Russian speakers 'easily' reproduce the foreign front round vowels *ü* and *ö*.[5]

Jakobson's analysis of these vowel systems requires making a number of decisions: whether the non-low vowels contrast in backness, or roundness, or whether both features are inseparable; and whether the low vowels participate in these contrasts or not. Such decisions are not self-evident. More surprising is that they have seldom been discussed explicitly. But this kind of contrast has been central to phonological theory for a century, because of an abiding intuition that contrastive features are particularly important to the patterning of sound systems. If contrastive features play a special role in phonology, then we need to be clear about what they are and how to identify them.

Before continuing, it may be worth returning to the issue of phonetics and sound patterns. Sapir's view that a pattern alignment may deviate from the phonetics was novel in 1925. Fifty years later it had become linguistic orthodoxy. In recent years the tide has shifted again. Much current work in phonology adopts the hypothesis that phonologies of languages are determined by phonetic principles (see, for example, Pierrehumbert, Beckman and Ladd 2000, and Hayes, Kirchner and Steriade 2004). I will argue that this hypothesis is wrong.

5 I am grateful to Wayles Browne for bringing Jakobson's paper to my attention. For a different view of the contrastive features of Russian vowels, see further section 8.3.

Without denying the contributions that phonetics can make to our understanding of sound systems, I will argue that the influence of phonetics, viewed apart from phonological contrast, has been over-stated.

Therefore, to understand the functioning of phonological systems we need to go beyond phonetics. In particular, I will argue that we need the approach to phonological contrast advocated here.

1.2 Two poles: contrast (negative) versus substance (positive)

Linguistic theory has never actually adopted Saussure's position, as expressed in the dictum at the head of this chapter, in its pure form. If a phoneme is indeed to be defined purely in negative terms, as a unit in opposition to the other phonemes in the inventory, then the phonemes of different phonological systems would become incommensurable. For example, a phoneme /i/ in a three-vowel system /i, a, u/ would be an entirely different object from an /i/ that is part of a four-vowel system /i, e, a, u/. Even two different three-vowel systems of the form /i, a, u/ could not be compared, since the contrasts in these systems would presumably not be identical. Thus, comparative and historical linguistics would become impossible, as any change in the nature of the contrasts in a system from one dialect to its neighbour, or from one historical stage to the next, would result in incomparable systems.[6]

Considerations of cross-dialect comparisons aside, it is simply not the case that linguistic units are characterized in purely negative terms. Though granting that what we call the phoneme /i/ in one language may differ in various substantive respects from the /i/ of another language, the symbol /i/ is not a purely abstract symbol devoid of any phonetic correlates. Designating a phoneme as /i/ suggests that it is realized as some sort of front high vowel. Similar observations hold for other linguistic units.

Alongside the view that phonemes are defined in purely negative terms, phonological theory has also contained the opposite tendency, and this, too, from the very beginning of modern synchronic linguistics. Bloomfield (1933: 79) assumed that 'phoneme-features will be present in the sound-waves', and thus launched the search for phonetic 'invariants', the notion that a phoneme is consistently characterized by certain acoustic cues. As was quickly observed, for example by Twaddell (1935), laboratory investigation had not revealed such cues up to then, and there was no reason to suppose they would ever

6 Such considerations have been raised as a critique of a purely structuralist dialectology; see Kiparsky (1965) for discussion.

be found. Twaddell then argued for a contrastive approach to defining the phoneme.

Throughout the history of phonological theory, then, there has been a recurring tension between these two views of how members of phonological inventories should be defined. At one pole, phonemes are defined negatively, in terms of how they contrast with other phonemes in an inventory. On this view, the types of oppositions a phoneme enters into are the most important determinants of its phonological behaviour. The other pole of this duality defines phonemes positively, as encoding substantive properties. These properties, on this account, are mainly what govern phonological behaviour, and the make-up of other phonemes in the system is of lesser importance.

The balance between the negative and positive approaches has been set differently at different times and in different schools of linguistics. If the formulations of Saussure and Bloomfield can be taken to represent each pole in a relatively pure form, then the theories of the Prague School represent a position in which they maintain a balance. Jakobson (1941) emphasized the oppositional nature of phonemes; but these oppositions are made in terms of distinctive features that have substantive content. For example, in his theory of how the system of distinctive features develops in the course of acquisition, Jakobson proposes that learners begin with an undifferentiated representation which first splits into a consonant (C) and a vowel (V). This formal opposition has phonetic content: V represents a sound of greater sonority, and C one of lesser sonority. The first split among the vowels is then likewise one between a vowel of greater sonority (say, /a/) and one of lesser sonority (such as /i/).[7] When these oppositions are maximized, the optimal syllable turns out to be /pa/.

Similarly, Trubetzkoy (1939, 1969) appeals to both contrastive and substantive properties in characterizing segmental systems. An example is his discussion of the phoneme /r/ in a variety of languages. German has two liquids, /r/ and /l/, which form, in his terms, an isolated bilateral opposition; that is, they are set apart from all other consonants by being liquids, and the distinction between them is unique to this pair. Trubetzkoy observes (1969: 73) that the 'phonemic content' of German /r/ is 'very poor, actually purely negative: it is not a vowel, not a specific obstruent, not a nasal, nor an *l*. Consequently, it

7 The theory of van der Hulst (1995, 1996, 2005) is similar in spirit, extending the idea of a basic C∼V contrast to the entire phonological system. Thus, consonants (C) split into C (obstruents) and V (sonorants), and each of these classes may be further divided into C and V groups. Similar divisions hold in the vowels. At every level, V represents a more sonorous or vowel-like sound and C represents its contrasting term. The precise meaning of each C and V contrast depends on its context.

also varies greatly with respect to its realization.' He notes that some speakers pronounce /r/ as a dental vibrant, some as a uvular vibrant, some as a noiseless guttural spirant, and it varies a great deal in different contexts as well. By contrast,

> Czech *r* has a much richer phonemic content, as it stands in a relation . . . not only to *l* but also to a special Czech phoneme *ř*: *r* and *l* are the only liquids, *r* and *ř* are the only two vibrants of Czech. *r* is distinguished from *ř* in that it is not an obstruent but a liquid; from *l* in that it is a vibrant. For this reason, Czech *r* is always, and in all positions, pronounced as a clear and energetically trilled sonorant.

In Gilyak (also called Nivkh, a language isolate spoken in Russia along the Amur River and on Sakhalin Island) (2c), /r/ is opposed to a voiceless spirant, and the two fall into place as the dental members of a series of oppositions between voiced and voiceless spirants, from which it follows that Gilyak /r/ is always dental.[8]

(2) /r/ in different languages (Trubetzkoy 1939)

a. German	b. Czech	c. Gilyak				
r - l	ř	v̭	r	z	ɣ	ʁ
	|	|	|	|	|	|
	r - l	f	r̥	s	x	χ

Trubetzkoy concludes that 'the phonetic realization of *r*, the number of its variants, etc., can be deduced from its phonemic content'.

Trubetzkoy's 'phonemic content', like Sapir's pattern alignment and Jakobson's phonemic make-up, can be understood as the sum of a phoneme's contrastive features. For Trubetzkoy, too, a major concern was to establish what these contrastive features are.

Though generative phonology was influenced by the Prague School and by American structuralism in various ways, the emphasis on contrast did not carry over into the developing theory exemplified by Chomsky and Halle's *Sound pattern of English* (*SPE*, Chomsky and Halle 1968). Generative phonology has tended to emphasize the substantive aspect of phonological entities, and has downplayed the importance of contrast. More than that, the classic argument of Halle (1959) against the structuralist taxonomic phonemic level can be understood as an argument against the relevance of contrastive features as well.

Halle (1959) considers the voicing and devoicing of Russian obstruents when preceding another obstruent. Most Russian obstruents come in voiced

8 The Gilyak phonemes are listed as in Maddieson (1984: 416).

and voiceless pairs, such as /t/ ~ /d/, /k/ ~ /g/, /s/ ~ /z/, and so on. A few obstruents, including /ts/, /tʃ/ and /x/, have no voiced counterpart. So whereas voicing and devoicing are neutralizing in most cases (in structuralist terms, they are morphophonemic processes, changing one phoneme into another existing phoneme), in the case of the unpaired obstruents they are allophonic rules, creating voiced sounds [dz], [dʒ] and [ɣ], that belong to no other phoneme. Despite this difference, Halle argues that the same rules of voicing and devoicing apply equally to the unpaired obstruents as well as to the paired ones, and that one would lose a significant generalization if one were to separate these rules into two different components of the grammar, as would be the case in neo-Bloomfieldian structuralist phonology.[9]

One could understand this argument – wrongly, I will contend – as bearing on the relevance of contrast to phonology. If we suppose that paired phonemes have a contrastive value for the feature [voiced], whereas the unpaired phonemes have no such contrastive feature, then it follows that the same voicing and devoicing rules apply to and are triggered by contrastive as well as by non-contrastive feature values. But the assumption that the unpaired phonemes do not have a contrastive voicing feature is not necessarily correct. I will argue, in fact, that the evidence suggests that it is not correct in the case of Russian.

The general antipathy to contrast in generative phonology is exemplified also by Anderson's *Phonology in the twentieth century* (S. R. Anderson 1985), still the standard history of the field. The theme of this work is summed up in the subtitle, *Theories of rules and theories of representations*. The history of phonology, in this influential view, is about the tension between rules and representations. In Anderson's analysis, the early emphasis on contrast had, if anything, negative consequences for phonological theory.

Nevertheless, I will argue throughout that contrast is too central to be kept out of phonological theory for long, and I will show that it gradually leaked back into generative phonology in various forms. It is one of the aims of this work to reconnect phonology with its roots in this respect and to establish phonological contrast as a central principle of phonological theory.

1.3 Plan of the book

In chapter 2, I will look at contrast from a logical point of view. I will show that phonologists have followed two different and incompatible approaches in arriving at what the contrastive features in phonology are in any given case.

9 See chapter 4 and Dresher (2005) for further discussion.

One method finds contrastive features based on minimal differences between fully specified phonemes; the second assigns contrastive features based on an ordering of features into a hierarchy. I will argue on logical grounds that the former approach, pairwise comparison, is wrong, and that the latter, the contrastive hierarchy, is right.

The ultimate criterion of adequacy relevant to phonological theory is not logic, however, but empirical adequacy. In the rest of this book I argue that the contrastive hierarchy is not only logically correct, but also empirically supported by the evidence of phonological systems. I will try to show also why a proper understanding of the contrastive hierarchy and its role in phonology has been difficult to attain.

In chapter 3, I look at how contrast is treated in the pioneering works of modern phonological theory, focusing on the work of Sapir and Trubetzkoy from the 1920s and 1930s, and looking briefly also at some later work by Martinet, Jakobson and Lotz, and Hockett. I argue that early phonological theory largely adhered to what D. C. Hall (2007) calls the Contrastivist Hypothesis, which holds that phonological computation operates only on contrastive features. The method for assigning such features, however, remained unclear. I show that the two approaches identified in chapter 2 are found in the work of these theorists – sometimes coexisting – though often implicitly. A review of the relevant cases leads to the conclusion that pairwise comparison tends to predominate where an analysis is based on abstract theorizing with no real empirical consequences; but when a contrastive analysis is advanced to capture an empirical generalization, it tends to employ the contrastive hierarchy. This is as we might expect, if indeed pairwise comparison is a faulty method and the contrastive hierarchy is the correct way to determine contrasts.

Chapter 4 is devoted to the work of Roman Jakobson and his collaborators in the 1950s. In this work the contrastive hierarchy was proposed to be the only method to assign contrastive features to phonemes (a principle not always followed in practice). However, as the decade progressed, the contrastive hierarchy became disconnected from the Contrastivist Hypothesis, as other rationales came to the fore, and the emphasis changed from the contrastive function of features to the information theoretic roles of underspecification. Deprived of a connection to the empirical workings of the phonology, the contrastive hierarchy became vulnerable to arguments against underspecification in phonology and was soon abandoned.

Chapter 5 reviews the role of contrast in generative phonology. Having rejected underspecification and the Contrastivist Hypothesis, generative phonology required some other mechanisms for capturing the sorts of generalizations captured in earlier theories by the Contrastivist Hypothesis. I review three subtheories that arose to meet this need: markedness theory, underspecification theory and feature geometry. Each of these subtheories does some of the work of the Contrastivist Hypothesis and contrastive hierarchy, but none serves as an adequate replacement for them.

Chapter 6 looks at contrast within Optimality Theory (OT). I argue, contrary to some claims, that OT is not in itself a theory of contrast; while OT is at home with feature hierarchies, I show that only certain types of hierarchies do the work of the contrastive hierarchy. I show how the contrastive hierarchy can be incorporated into a serial version of OT.

Chapter 7 presents a series of case studies in support of the contrastive hierarchy and the Contrastivist Hypothesis within the framework of Modified Contrastive Specification (MCS), drawing on research done mostly at the University of Toronto since the early 1990s. These studies also show the insufficiency of competing accounts, discussed in chapter 8. These include purely phonetic approaches, and accounts based on other ways of incorporating contrast into phonological theory. I will show that the minimal pairs approach to assigning contrastive features remains influential in contemporary phonological theory, but continues to suffer from the shortcomings identified in the earlier chapters. I argue that the contrastive hierarchy continues to provide a more adequate approach to contrast. Chapter 9 is a brief conclusion.

2 *The logic of contrast*

2.1 Contrastive specification: an elusive problem

It is far from obvious how to decide, for a given phoneme in a given language, which of its features are contrastive and which are not. The problem is made even more elusive by the fact that it does not appear to be difficult. In particular situations we may have intuitions about what the answer must be. But our common-sense intuitions may lead us astray, in this area as in others. Or we may find that we can follow more than one logical chain of reasoning, each of which may appear to be sound, but which lead to different and incompatible conclusions.

To give something of the flavour of this problem, both its seeming obviousness and real difficulty, I would like to begin with a quote from Stephen Anderson (1985: 96–7). Anderson is illustrating Trubetzkoy's (1939) notion of *phonemic content*, intended to be the sum of the contrastive properties of a phoneme: 'If we consider [English] /t/, for example, we can see that this segment is phonologically *voiceless* (because it is opposed to /d/), *non-nasal* (because opposed to /n/), *dental* (because opposed to /p/ and /k/), and a *stop* (because opposed to /s/ and to /θ/).'

Anderson is not proposing a detailed analysis of English; he is simply illustrating what some of the contrastive features of English /t/ would be in a Trubetzkoyan analysis, and presumably in any analysis of contrast that used these features. And yet, none of the features listed above are uncontroversially contrastive. Assuming that English /t/ and /d/ differ only in their laryngeal specifications, it is possible that /t/ is contrastively voiceless, though other laryngeal features are also possible: thus, /t/ and /d/ differ also in *aspiration* (/t/ is aspirated, /d/ is not), and in *tension* (/t/ has a tenser articulation than /d/). It is not obvious which of these laryngeal features is contrastive in English. Similarly, we could agree that /t/ differs from /p/ and /k/ with respect to its place of articulation, but it is not obvious that this feature should be designated dental as opposed to the more general *coronal*.

The other contrasts are even more problematic. The segment /t/ is opposed to /n/ not only in nasality but also in *voicing* (/t/ is voiceless, /n/ is voiced) and *sonority* (/t/ is an obstruent, /n/ is a sonorant): how do we know that nasality is the contrastive feature and not one of the others? And while /t/ differs from the continuants /s/ and /θ/ in being a stop, it also differs from these phonemes in various other ways. For example, /t/ is *non-strident*, in contrast to /s/, and *apical* in contrast to /θ/. How do we know, then, that the contrastive features of /t/ are those designated by Anderson and not any of the alternatives?

This example is not intended to show that Anderson (1985) was being particularly imprecise; on the contrary, Anderson is more careful than most, and his discussion of the contrastive features of /t/ is entirely typical of what one finds throughout the literature. Anderson's choices are not obviously wrong, but it is not clear that they are right, either. More fundamentally, he provides no procedure for making such distinctions, nor does he discuss how such decisions were made in the history of phonology. Given the centrality of the issue in many phonological theories, this is a striking omission, in my view, and yet again entirely typical of most treatments of the subject. There has been much discussion of the *status* of contrastive representations in phonology; Anderson (1985), for example, is particularly concerned with the question of whether only contrastive features should be included in lexical representations, or all features. This has been a central issue in phonological theory, but it presupposes the answer to a more humble question: how do we decide which features are contrastive in any given segment?

We will see that this more basic question has been answered in different ways. One way proceeds from making pairwise comparisons between the segments of an inventory; the other involves successively dividing up the inventory by an ordered list of features. These approaches are not equivalent, and typically yield different results. Both have a certain common-sense appeal; but I will argue that one of them cannot be correct.

To illustrate each approach, we will look at a very simple problem: how to specify the features that distinguish the three bilabial stops /p, b, m/, such as occur, for example, in Standard French. This problem has been treated by numerous authors over the years. To illustrate the two approaches, we will consider the analyses of Martinet (1960) and Jakobson and Lotz (1949). Both of these analyses emerge from the Prague School and share certain general background assumptions. But their approaches to contrastive specification are quite different.

2.2 Contrastive specification by pairwise comparisons

Martinet (1964: 62–4) considers how to isolate the relevant (i.e., contrastive) features of the Standard French consonants. To simplify the discussion we will focus here only on the bilabial stops /p, b, m/. Martinet proposes that /p/ is contrastively 'unvoiced'; /b/ is 'voiced' and 'non-nasal'; and /m/ is 'nasal'. We can convert these specifications into two binary features, [voiced] and [nasal]: [+voiced] is equal to 'voiced', [−voiced] is equal to 'unvoiced', [+nasal] is equal to 'nasal', and [−nasal] is equal to 'non-nasal'. In these terms, the specifications proposed by Martinet amount to those in (1).

(1) Contrastive specifications for French bilabial stops (Martinet 1964)

	p	b	m
[voiced]	−	+	
[nasal]		−	+

Martinet arrived at these specifications by isolating those features that serve to distinguish phonemes that are minimally different in terms of their full feature specifications. To follow his reasoning, let us start with the full (not just the contrastive) specifications of the phonemes /p, b, m/ for the features [voiced] and [nasal], shown in (2).

(2) Full specifications for French bilabial stops

	p	b	m
[voiced]	−	+	+
[nasal]	−	−	+

We observe that /p/ and /b/ differ only with respect to the feature [voiced]. Therefore, by any definition, this feature must be contrastive in these segments; if it were absent, we could not distinguish /p/ from /b/. By the same token, /b/ and /m/ are distinguished only by the feature [nasal], which must, too, be designated as contrastive. Let us circle these two undisputedly contrastive features:

(3) Circled features certainly contrastive

	p	b	m
[voiced]	⊖	⊕	+
[nasal]	−	⊖	⊕

What about the features that have not been circled? These are the features Martinet leaves out of his contrastive specifications, so evidently in his view they are not contrastive. He explains (1964: 65) why /m/ is not to be considered as contrastively [+voiced]:

> It is likewise to be noted that . . . the segments /m n ñ/ are not only nasal but also voiced. However, here voice cannot be dissociated from nasality since in this position there are no voiceless nasals. This is why /m n ñ/ do not figure in the class of the 'voiced' elements, which are defined as such solely in virtue of their opposition to 'voiceless' partners.

By similar reasoning we can see why /p/ is not classed as 'non-nasal', even though it is phonetically non-nasal, just like /b/. It is because /p/, unlike /b/, has no nasal 'partner'; such a partner would have to be otherwise identical to /p/, that is, a voiceless nasal stop /m̥/. Since there is no such phoneme in French, /p/ is not contrastively non-nasal.

This method proceeds in terms of pairwise comparisons. It designates as contrastive all and only features that serve to distinguish between pairs of phonemes. An explicit algorithm for extracting contrastive features by this method was proposed by Archangeli (1988).[1] I will call this the Pairwise Algorithm, given in (4).

(4) Pairwise Algorithm (Archangeli 1988)
- a. Fully specify all segments.
- b. Isolate all pairs of segments.
- c. Determine which segment pairs differ by a single feature specification.
- d. Designate such feature specifications as 'contrastive' on the members of that pair.
- e. Once all pairs have been examined and appropriate feature specifications have been marked 'contrastive', delete all unmarked feature specifications on each segment.

Pairwise comparison seems to make sense, and it has been widely used in phonology (not always explicitly) as a way to isolate contrastive features.

2.3 Contrastive specification by feature ordering

The above analysis of the contrastive features of the Standard French bilabial consonants is not the only one in the literature. An entirely different analysis is given by Jakobson and Lotz (1949). As with the Martinet example above, I will focus only on their analysis of the bilabial consonants, extracting it from

1 Archangeli (1988) presents this algorithm as part of an argument *against* the sort of contrastive specification proposed by Steriade (1987). Her argument is that the algorithm is faulty, and hence so is contrastive specification. I will show that while the algorithm, and the general approach it instantiates, are indeed faulty, contrastive specification does not necessarily depend on this approach. A more elaborate algorithm was formulated by van den Broecke (1976); see section 2.5.5 below for discussion.

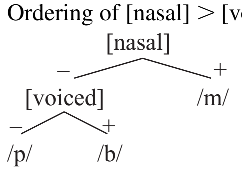
their larger analysis of the contrasts in the French consonant system. I will also modify their features to conform with the example we have been using; specifically, I will continue to use [voiced] in place of their *tense* (tense stops are voiceless, non-tense stops are voiced). With these adjustments, their analysis of the contrastive features for the bilabial consonants is as in (5).

(5) Contrastive specifications for French bilabial stops (Jakobson and Lotz 1949)

	p	b	m
[voiced]	−	+	
[nasal]	−	−	+

Notice that the contrastive specifications in (5) differ from Martinet's in (1) in that /p/ in (5) is specified as [−nasal], a specification omitted in (1). Jakobson and Lotz arrived at a different contrastive specification from that of Martinet because they used a different method. They themselves do not make their method explicit, but we can reconstruct it from later work by Jakobson and his collaborators. Rather than make pairwise comparisons of fully specified segments, they put all the potentially distinctive features into an ordered list, and divide the inventory successively on the basis of this list until every segment has received a distinct representation. In this case they order [nasal] ahead of [voiced]. We will represent the ordering of feature [F] ahead of feature [G] by the notation '[F] > [G]'. We can represent the result of feature ordering by a tree, as in (6).

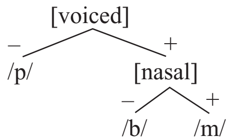
(6) Ordering of [nasal] > [voiced] applied to /p, b, m/



First we divide the inventory into two sets on the basis of the feature [nasal]: one set contains those phonemes that are nasal and the other contains those that are non-nasal. In the small inventory of bilabial consonants we are concerned with, /m/ is the only nasal consonant, and so is already distinct from the others. There are two non-nasal consonants, however, and they need to be distinguished by the feature [voiced], which is contrastive only in the [−nasal] set. Thus, we obtain the specifications in (5).

When we derive contrastive specifications from ordered features, the ordering makes a difference. To see this, consider what we would obtain if we reversed the order of the two features [voiced] and [nasal]. The results are shown in (7) in the form of a tree, and in (8) as a table of contrastive specifications.

(7) Ordering [voiced] > [nasal] applied to /p, b, m/



(8) Contrastive specifications with the ordering [voiced] > [nasal]

	p	b	m
[voiced]	–	+	+
[nasal]		–	+

This time we first divide the inventory on the basis of the feature [voiced]: /p/ is the only voiceless consonant, and this feature suffices to set it apart from the other consonants. There are two voiced consonants, however, and they need to be distinguished by the feature [nasal], which is now contrastive only in the [+voiced] set.

On this approach, contrastive specifications are determined by splitting the inventory by means of successive divisions, governed by an ordering of features (Jakobson, Fant and Halle 1952; Cherry, Halle and Jakobson 1953; Jakobson and Halle 1956; Halle 1959). An algorithm corresponding to this idea, the Successive Division Algorithm (SDA; Dresher 1998b, 2003, 2008, based on the work of Jakobson and his collaborators cited above), is given in (9):

(9) The Successive Division Algorithm

a. Begin with *no* feature specifications: assume all sounds are allophones of a single undifferentiated phoneme.
b. If the set is found to consist of more than one contrasting member, select a feature and divide the set into as many subsets as the feature allows for.[2]
c. Repeat step (b) in each subset: keep dividing up the inventory into sets, applying successive features in turn, until every set has only one member.

2 This algorithm does not require any particular set of features. I assume that the set of relevant distinctive features is given by the theory of features, whatever that may turn out to be.

The algorithm in (9) is a very general formulation for defining contrast and redundancy for members of an inventory.[3] It designates feature values as being contrastive or redundant in terms of an ordering of features, which I will call a *contrastive hierarchy*.[4] In this approach, contrast is a matter of *relative scope* or ordering of contrastive features.

2.4 Contrastive specification as a logical problem

We have now seen two different approaches to deriving contrastive feature specifications. These are not simply two ways of arriving at the same answer; the fact that they yield different answers shows us that the methods are fundamentally different, and inconsistent with each other. In the example of the bilabial stop consonants characterized in terms of the features [voiced] and [nasal], pairwise comparison yields the contrastive representations in (1), and feature ordering gives the representations in (5). To be more precise, feature ordering can give us more than one answer, depending on how the features are ordered. In the above example, we obtain (5) in the ordering [nasal] > [voiced], and (8) with the ordering [voiced] > [nasal]. This is another important way in which the two approaches differ: given a phonemic inventory and a fixed set of features, pairwise comparison always gives the same answer (if it gives an answer at all); feature ordering can give different answers.

The flip side of contrast is *redundancy*, which is often equated with *predictability*: if, after we remove a feature, we can predict what it is, based on our knowledge of the inventory and the other features, then it stands to reason that it is redundant; and if it is redundant, it cannot be contrastive, or so it would

3 A more procedurally explicit version of the SDA is as follows:

a. In the initial state, all tokens in inventory I are assumed to be variants of a single member. Set I = S, the set of all members.
b. i) If S is found to have more than one member, proceed to (c).
   ii) Otherwise, stop. If a member, M, has not been designated contrastive with respect to a feature, G, then G is *redundant* for M.
c. Select a new *n*-ary feature, F, from the set of distinctive features. F splits members of the input set, S, into *n* sets, $F_1 - F_n$, depending on what value of F is true of each member of S.
d. i) If all but one of $F_1 - F_n$ is empty, then loop back to (c). (That is, if all members of S have the same value of F, then F is not contrastive in this set.)
   ii) Otherwise, F is *contrastive* for all members of S.
e. For each set F_i, loop back to (b), replacing S by F_i.

4 As far as I know, the earliest appearance of this term in print in this sense is in Walker (1993).

appear. This kind of reasoning, which I will argue is flawed, would appear to support the pairwise method. For both features omitted from the specifications in (1) are predictable from the other specifications, and can be filled in by redundancy rules as shown in (10).

(10) Contrast by pairwise comparison

a. Contrastive specifications

	p	b	m
[voiced]	−	+	
[nasal]		−	+

b. Redundancy rules

i. [−voiced] → [−nasal]   ii. [+nasal] → [+voiced]

Let us define *logical redundancy* as in (11).

(11) Logical redundancy

If Φ is the set of feature specifications of a member, M, of an inventory, then the feature specification [F] is *logically redundant* iff it is predictable from the other specifications in Φ.

The omitted features in (1), repeated as (10a), are logically redundant in the sense of (11) because they are predictable from the other features, given this inventory. Thus, because /p/ is the only [−voiced] member of the inventory, its feature value [−nasal] is predictable by rule (10bi); similarly, the value [+voiced] for /m/ is predictable by rule (10bii) because /m/ is the only [+nasal] phoneme.

The specifications derived from feature ordering do not omit all logically redundant features. In the ordering [nasal] > [voiced] (5), /p/ is contrastively specified as [−nasal], even though this specification is logically redundant, as we have seen. And in the order [voiced] > [nasal] (8), the value [+voiced] for /m/ is not omitted, though it, too, is logically redundant.

The concept of contrast that emerges from feature ordering, then, is not based on logical redundancy as defined in (11). Nevertheless, *in any particular feature ordering* some features are defined as redundant (equivalent to those features not designated as contrastive). To avoid confusion, let us call this type of redundancy *system redundancy* and define it as in (12).

(12) System redundancy

The feature specification [F] is *system redundant* iff it is not contrastive in terms of the method used for determining which features are contrastive in an inventory.

System redundancy is relative to a particular method for designating features as contrastive, whereas logical redundancy is fixed for a given inventory and set of features. Since a specification that is not logically redundant is not predictable under any procedure, it follows that the specifications designated as system redundant in any system of contrastive specification will also be logically redundant. The converse does not necessarily hold: a specification may be logically redundant but not system redundant, as we have seen. Many discussions of redundancy in phonology fail to distinguish the two types of redundancy, and this conflation of two different concepts can lead to considerable confusion.

Which approach to contrastive specification is correct? From the point of view of phonology, the question is ultimately an empirical one: which of these approaches, if any, yields representations that are significant in the phonology? The answer to this question could conceivably be 'neither', if in fact contrastive specifications play no special role in the phonology. I will continue to assume, however, that phonology is sensitive to contrastive specifications and that empirical evidence can be adduced to show that the feature ordering approach is correct. Evidence to this effect will be presented in subsequent chapters.

Putting this sort of empirical evidence aside for now, it was already recognized by Trubetzkoy (2001[1936]: 15) that the question of contrast (what he called 'the concept of the opposition') 'is not exclusively a phonological concept, it is a logical one, and the role it plays in phonology is strongly reminiscent of its role in psychology. It is impossible to study phonological oppositions (of which phonemes are only the terms) without analyzing the concept of the opposition from the point of view of psychology and logic.'

In the remainder of this chapter I will consider the logic of contrast. We will see that the pairwise approach suffers from severe logical problems. Feature ordering appears to be impeccable from a logical point of view, though it challenges us to order the features correctly for every language.

2.5 Arguments against the pairwise approach to contrastive specification

2.5.1 Distinctness

Let us consider the contrastive specifications in (1) a bit more closely. They are repeated here as (13) for convenience.

(13) Contrastive specifications by the pairwise method

	p	b	m
[voiced]	−	+	
[nasal]		−	+

Recall that these representations were derived by making a pairwise comparison between /p/ and /b/ on one side, and between /b/ and /m/ on the other. Each of these involves a minimal difference in one feature, which must therefore be contrastive. Let us call these *minimal pairs*, defined as in (14):

(14) Definition of a minimal pair
Two members of an inventory that are distinguished by a single feature are a minimal pair.[5]

Minimal pairs play a crucial role in the pairwise approach. But let us now observe that there is in fact a third pairwise comparison we can make in (13), between /p/ and /m/, and it is not obvious that they are properly distinguished in (13). The segment /p/ is characterized as being [−voiced] and /m/ is characterized as [+nasal]. Thus, they are not in contrast with each other along some common dimension. Where /p/ has a specification, /m/ has none, and vice versa. Their specifications look different, but they are not necessarily distinct. Without applying the redundancy rules, we would not know if /p/ and /m/ are distinct from each other or not. But then we have failed in our attempt to represent all the relevant contrasts in the chart.

The representations in (13) would be ruled out by a criterion in the linguistic literature known as the Distinctness Condition, proposed by Halle. He formulates it as in (15), and gives the examples in (16).

(15) Distinctness of phonemes (Halle 1959: 32)
Segment-type {A} will be said to be *different from* segment-type {B}, if and only if at least one feature which is phonemic in both, has a different value in {A} than in {B}; i.e., plus in the former and minus in the latter, or vice versa.

(16) Examples of distinctness and non-distinctness (Halle 1959: 32)
a. {A} is not 'different from' {C}

	{A}	{B}	{C}
Feature 1	+	−	+
Feature 2	0	+	−

5 This kind of *featural* minimal pair differs from the usual sense of 'minimal pair' in linguistics, which is a pair of *words* that differ by a single phoneme: for example, *bit* and *pit*, or *cat* and *cap*. Determination of word minimal pairs does not require us to identify in what way one phoneme is crucially distinguished from another.

b. All three segment-types are 'different'.

	{A}	{B}	{C}
Feature 1	+	−	−
Feature 2	0	+	−

By the terms of the Distinctness Condition, /p/ and /m/ in (13) are not different from each other. Therefore, the pairwise approach fails to contrast these elements of the inventory, and hence fails to provide an adequate set of contrastive specifications, according to the Distinctness Condition.

The Distinctness Condition has not been uncontroversial in linguistic theory, and some readers may question whether it is really necessary. Why *can't* the absence of a specification count as a value distinct from the presence of a value? After all, the system in (13) will result in three distinctly specified members once we apply the redundancy rules, so what is the problem?

The problem is that we are abusing the notion of contrast. Consider a language that has bilabial /p/ and /m/, but lacks /b/ (a fairly common situation, as many languages lack phonemic voiced obstruents). If asked to provide a contrastive specification of such an inventory, would anybody choose (17)? The relation between /p/ and /m/ in (17) is the same as that between /p/ and /m/ in (13); but without the middle member /b/ that forms minimal pairs with both of them, the specifications in (17) appear bizarre. It does not make sense to assert that one member in a two-member set is contrastively voiceless and the other is contrastively nasal. In contrast with what? If something is contrastively voiceless, it can only mean in contrast to something that is voiced, and the same holds for [nasal]: what is contrastively not nasal must be non-nasal (oral).[6]

(17) Contrastive specifications of /p/ and /m/?

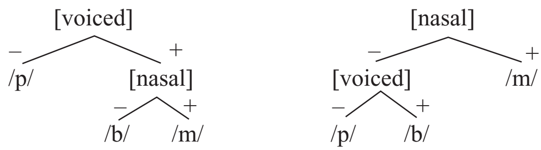
	p	m
[voiced]	−	
[nasal]		+

On further reflection, it appears that the chart in (13) results from a misconstrual of our original observations about the inventory. When we observed above that '/p/ is the only member that is [−voiced]', what we had in mind was that, once we made a contrast between [−voiced] /p/ on one side and [+voiced] /b/ and /m/ on the other, there was no need to further specify /p/ for [nasal]. The relevant contrasts can be pictured as in (18a). And when we observed that '/m/ is the only member that is [+nasal]', we had in mind a picture such as

6 Recall that we are using binary features. In a privative feature system, to be discussed below, the absence of a value acts like a value, and the conclusions drawn above do not hold.

(18b), where, once /m/ is specified [+nasal] and /p/ and /b/ [−nasal], there is no need to further specify /m/.

(18) Two ways of viewing the contrasts
a. [voiced] > [nasal] b. [nasal] > [voiced]



Thus, the observations that /p/ is the only voiceless member and /m/ is the only nasal member are correct, but in terms of contrastive force they derive from two different ways of cutting up the inventory, corresponding to the feature ordering approach. The ultra-minimal specification in (13) results from trying to put together two observations that derive from incompatible ways of dividing up the inventory. For this reason, it fails to adequately contrast /p/ and /m/.

2.5.2 The problem of too many features

Whatever one thinks of the Distinctness Condition and the above logical argument, the inventory in (17) shows us a fundamental problem with the pairwise approach: in many cases, there are too many logically redundant features. In (17), *every* feature specification is redundant given the others: in /p/, [−voiced] predicts [−nasal] and [−nasal] predicts [−voiced], and in /m/, [+voiced] predicts [+nasal] and [+nasal] predicts [+voiced]. Thus, all four feature specifications are logically redundant, but they can't all be omitted! But this is what the Pairwise Algorithm (4) does in such situations: the inventory in (17) contains no minimal pairs, as defined above, because the two members differ from each other by two features, not by one. Therefore, no features are designated as contrastive, and all are removed. Clearly, removing all logically redundant features in this inventory does not work, and the Pairwise Algorithm fails in such cases.

The problem of too many features not only arises in atypical inventories, but is ubiquitous and affects almost every phonological inventory in some way. For example, it arises in the most common vowel inventories.

Consider first the most commonly attested vowel system, the five-vowel inventory /i, e, a, o, u/.[7] If we include only the features [high], [low], [back]

7 According to the UCLA Phonological Segment Inventory Database (UPSID) (Maddieson 1984), sixty-eight of the listed languages (21 per cent of the sample) have five-vowel systems, by far the highest number of any vowel system. Of these, thirty-one have the indicated inventory.

and [round], we already have too many features for the pairwise method to function, as shown in (19).

(19) Five-vowel system, features [high], [low], [back], [round]

a. Full specifications

	i	e	a	o	u
high	+	–	–	–	+
low	–	–	+	–	–
back	–	–	+	+	+
round	–	–	–	+	+

b. Contrastive specifications according to the pairwise method

	i	e	a	o	u	*Minimal pairs*
high	+	–		–	+	{i, e}; {o, u}
low						none
back						none
round						none

The only minimal pairs are {i, e} and {o, u}; the features [back] and [round] double each other for every vowel except /a/, making each other logically redundant, but not leaving behind enough features to make a contrast between /i, u/ and /e, o/, and leaving /a/ without any contrastive feature.

The same point applies a fortiori with the simple three-vowel system /i, a, u/. If we confine ourselves to two features, say [high] and [round], the vowels fall into minimal pairs and the pairwise method can assign them distinct representations.

(20) Three-vowel system, features [high], [round]

a. Full specifications

	i	a	u
high	+	–	+
round	–	–	+

b. Specifications according to the pairwise method

	i	a	u	*Minimal pairs*
high	+	–		{i, a}
round	–		+	{i, u}

Adding one more feature, say [back], wipes out the minimal pairs (21), causing the pairwise method to fail to distinguish them.

(21) Three-vowel system, features [high], [round], [back]

a. Full specifications

	i	a	u
high	+	–	+
round	–	–	+
back	–	+	+

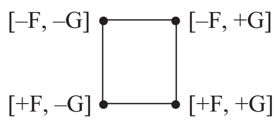
b. Specifications according to the pairwise approach

	i	a	u	*Minimal pairs*
high				none
round				none
back				none

2.5.3 Minimal pairs and feature space

We can approach the problem of algorithms that depend on minimal pairs by considering more generally how inventories fill out the available space of feature specifications. Two binary features, F and G, define a two-dimensional feature space with nodes at four possible values: [−F, −G], [−F, +G], [+F, −G], and [+F, +G]. This space can be diagrammed as in (22).[8]

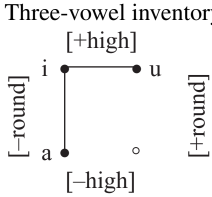
(22) Space defined by two features



The lines connect nodes that are separated by one feature value. Such nodes, which we will call *neighbours*, are minimal pairs. In (22), each node has potentially two neighbours, and hence can form minimal pairs with two other members of the inventory, if they are present.

If an inventory completely fills the feature space, then it is guaranteed that the pairwise method will find sufficient minimal pairs to arrive at a contrastive specification. The pairwise method can tolerate some gaps in the feature space, as long as there are sufficient neighbours. For example, the inventory in (20) can be diagrammed as in (23), where ◦ indicates an unfilled position.

(23) Three-vowel inventory, features [high], [round]

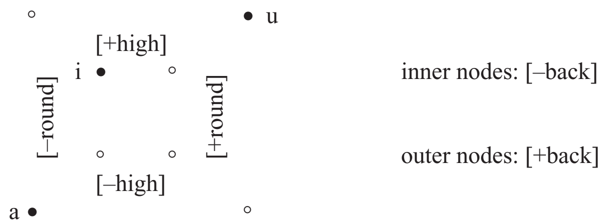


We can observe graphically how adding the feature [back] isolates the members of the inventory in the larger feature space (24). The feature space in (24)

8 Compare the information-theoretic adjacency graphs employed by Shannon (1993 [1956]); see D. C. Hall (2004, 2007) for discussion.

is an expansion of the one in (23): the features [high] and [round] remain as before, but now the inside nodes are [−back] and the outside nodes are [+back]. The addition of the feature [back] exiles the [+back] segments /a/ and /u/ to the outer nodes, away from /i/ that had formerly connected them.

(24) Three-vowel inventory, features [high], [round], [back]



It follows, then, that the pairwise approach to contrastive specification fails in the simplest vowel systems, when all features are taken into account. The reason that this simple fact has not disqualified it long ago as a theory of contrast is that many analysts tacitly reduce the feature set to a minimal set. That is, if an inventory is classifiable using a proper subset of the full set of features, then the 'extra' features are quietly discarded until the set is minimal, but still able to distinguish every member of the inventory.

In such cases, the analyst chooses which logically redundant features to delete and which to retain. Such a choice implies some notion of a hierarchy, and is in fact a tacit use of feature ordering. Therefore, even an algorithm formulated to remove redundancies from fully specified specifications must be supplemented by some device that orders the redundant specifications so that some take priority over others. That is, some notion of a feature hierarchy is required even in a pairwise approach to contrastive specification. But if a feature hierarchy is independently needed, there is no further rationale for the pairwise method, since the hierarchy can do all the work by itself.

2.5.4 The problem of too few minimal pairs

In the type of case discussed above, the pairwise approach can be salvaged (ignoring violations of the Distinctness Condition) by removing features (that is, appealing to a feature hierarchy for this limited purpose) until a minimal set of features remains. But this approach fails in more spectacular ways when faced with inventories that use a minimal set of features whose members do not fill the space of feature values in the right way. In such cases no feature may be removed from the set of relevant features specifying the inventory, but there are still not a sufficient number of minimal pairs to fuel the Pairwise Algorithm.

Consider again the common five-vowel system in (19), this time without the feature [round]. According to the pairwise method, this five-vowel system, fully specified for the features [high], [low] and [back] in (25a), would be underspecified as in (25b).

(25) Five-vowel system, features [high], [low], [back]

a. Full specifications

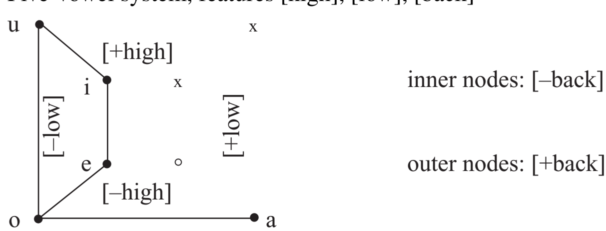
	i	e	a	o	u
high	+	−	−	−	+
low	−	−	+	−	−
back	−	−	+	+	+

b. Specifications according to the pairwise method

	i	e	a	o	u	*Minimal pairs*
high	+	−		−	+	{i, e}; {o, u}
low			+	−		{a, o}
back	−	−		+	+	{i, u}; {e, o}

That the pairwise method gives a contrastive specification at all, whether correct or not, is due to the connectedness of the paths through the space of the three features being considered here. As before, we can model the space and the minimal pair paths through it with a diagram as in (26). An x represents an impossible combination of [+high, +low].

(26) Five-vowel system, features [high], [low], [back]



Archangeli (1988) points out that not every five-vowel system can be assigned a contrastive set of specifications by the Pairwise Algorithm. An example of such an inventory is the vowel system of Maranungku (Tryon 1970), given in (27).

(27) Maranungku, features [high], [low], [back]

a. Full specifications

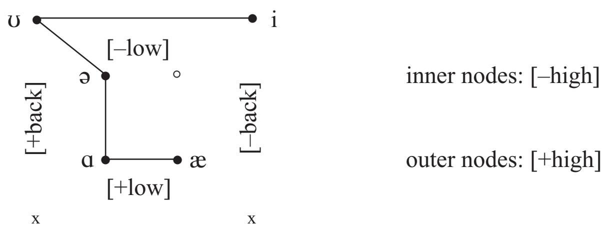
	i	æ	ɑ	ə	ʊ
high	+	–	–	–	+
low	–	+	+	–	–
back	–	–	+	+	+

b. Specifications according to the pairwise method

	i	æ	ɑ	ə	ʊ	*Minimal pairs*
high				–	+	{ə, ʊ}
low			+	–		{ɑ, ə}
back	–	–	+		+	{i, ʊ}; {æ, ɑ}

In this vowel system, /i/ and /æ/ have the same contrastive specification because they occupy parallel positions in a [back] contrast, but have no other neighbours that could further differentiate them in terms of the pairwise method. This situation is represented graphically in the diagram in (28).

(28) Maranungku, features [high], [low], [back]



Now, if it were the case that Maranungku represented a relatively rare situation, one could argue that such examples are not serious problems for the pairwise method. One might reason the other way around: if the Pairwise Algorithm is correct, we expect that actual phonological inventories ought to have sufficient minimal pairs. However, D. C. Hall (2004, 2007) argues that this expectation is seriously misguided. For actual inventories do not aim for minimal phonetic contrasts, free of all redundant differences. On the contrary, there is much evidence that minimal phonological contrasts are *enhanced* by additional phonetic distinctions, or that members of an inventory are *dispersed* so as to maximize the perceptual salience of contrasts (see chapters 7 and 8 for further discussion and references).

As D. C. Hall (2007: 165) demonstrates, the three-vowel inventory /i, a, u/ diagrammed in (24) is problematic for the Pairwise Algorithm because its members differ with respect to too many features. He shows that the algorithm would have no difficulty finding contrastive features for an inventory like /ɨ, ə, ʉ/, whose members are closer together; but this type of inventory is non-existent, whereas /i, a, u/ is the most common three-vowel inventory. It appears that the pairwise approach rests on exactly the wrong assumption about real phonological inventories, which are designed to thwart an algorithm that relies on phonetic minimal pairs.

2.5.5 Extending the Pairwise Algorithm

The pairwise approach as instantiated in the Pairwise Algorithm (4) identifies as contrastive only features that distinguish minimal pairs, and finds no contrastive features for members of an inventory that are distinguished by more than a single feature. There is no reason in principle why pairwise comparison must be limited in this way. To deal with situations where members of an inventory are distinguished by more than one feature, there must be a way of selecting one of them as being contrastive. The simplest way to do this is to order the features, selecting the feature that is highest in the ordering. But this is to adopt feature ordering, and makes pairwise comparison superfluous.

A sophisticated version of pairwise comparison was devised by van den Broecke (1976: 33–4). He wrote a computer program that takes as input a phonological inventory with fully specified feature matrices, with the aim of arriving at a set of contrastive specifications. The first step of this algorithm is equivalent to the Pairwise Algorithm, except that for each pair of phonemes the program records every feature that distinguishes the two. As in the Pairwise Algorithm, features that uniquely distinguish a pair of phonemes are designated as contrastive for that pair.

But whereas the Pairwise Algorithm stops at this point, van den Broecke's program is just getting started. If a pair of phonemes is distinguished by more than one feature, but one of those features has already been marked as contrastive for another pair, then that feature is selected. If none of the distinguishing features has been marked as contrastive elsewhere, then the program creates several columns and in each column marks one of the features as contrastive and the others as redundant. These columns multiply as a function of the number of such choices. For example, for an English inventory of 48 phonemes characterized by 14 distinctive features, van den Broecke reports that the program generated up to 52 columns, with an average number of 10.5 columns per segment.

The next step in van den Broecke's procedure is to assign a relative weighting factor to each candidate contrastive feature based on the number of columns in which the feature is listed as being obligatory. The feature with the highest weighting is selected as contrastive.

As van den Broecke (1976: 35) points out, the specifications arrived at in this fashion are based only on considerations of economy, and do not take phonological patterning into account. As a result, the specifications tend to be highly counter-intuitive and not the specifications that any phonologist would propose.[9] For example, the major class features [vocalic], [consonantal] and [sonorant] are rarely marked as contrastive by this method, because they are predictable from more specific features, like [strident]. Thus, [sonorant] is marked as contrastive in only a single segment, /ð/, a typically strange result of this method.

Van den Broecke does not advocate this method for arriving at contrastive specifications; on the contrary, he presents it to show that attempts to remove redundant features based only on a notion of feature economy or minimality (a criterion allied to logical redundancy, as it aims to reduce the set of specifications to a minimum) result in unnatural contrastive specifications that no phonologist would posit. I am unaware of any other attempt to apply an algorithm along these lines.

2.5.6 Summary

The preceding sections have argued that the pairwise approach, despite its common-sense appeal, faces serious logical problems. In the cases where it yields a set of contrastive specifications that make all the segments look different, it is not at all clear that the specifications are properly contrastive. More usually, the analyst must remove certain logically redundant features *before* making the pairwise comparisons, thus tacitly putting the features into a partial order. Finally, there are cases where pairwise comparisons simply fail to distinguish some members of an inventory even when the features are reduced to a minimal set. In short, the pairwise approach to contrastive specification is simply too problematic and too sensitive to the vagaries of the distribution of members of an inventory to serve as a principle for assigning contrastive specifications. While one can imagine ways of trying to extend the pairwise approach, the one extension I am aware of (van den Broecke 1976) results in bizarre specifications and has never actually been used in a phonological analysis.

2.6 Feature ordering

The feature ordering approach, instantiated by the SDA, is not subject to these difficulties. As long as the members of an inventory can be distinguished by the full set of relevant distinctive features, the SDA is guaranteed to arrive at

9 But see Jakobson's specifications for Serbo-Croatian discussed in section 4.5.1.

properly contrastive specifications. To see this, we will consider the various difficulties faced by pairwise comparison in turn.

First, all members of an inventory assigned contrastive features by the SDA are guaranteed to pass the Distinctness Condition. This was demonstrated by Halle (1959). Since the SDA works by successively splitting an inventory and does not stop until each segment has been assigned a unique set of features, it is guaranteed that every phoneme will be distinct from every other phoneme in the sense of the Distinctness Condition. In a sense, every time the SDA splits an inventory on the basis of a feature, F, it applies the Distinctness Condition with respect to F: assuming a binary feature, every member of the relevant inventory is assigned either [+F] or [−F]. Therefore, every phoneme in one set is distinct from every phoneme in the other set. Phonemes in the same set are not distinct with respect to F (or with respect to any feature ordered higher than F); but since the procedure iterates, it is guaranteed that eventually every set will have just one member.

Second, the SDA does not depend on any particular distribution of the members of an inventory in feature space. No matter how sparse the inventory or how long the list of features, the SDA functions in the same way. It starts by selecting the first feature in the list, F_1. If this feature is contrastive within the inventory, then the SDA splits the inventory into two sets (assuming binary features), one contrastively specified [$+F_1$], the other [$-F_1$]. It is possible that this feature is not contrastive: it could be that the members of the inventory are all [$+F_1$], or all [$-F_1$]; or F_1 may not be relevant to the members of the inventory (in the case of features defined to apply in limited circumstances). In any of these cases, the SDA will move on to the next feature. As long as the full set of features is capable of characterizing each member of the inventory in a unique way, the SDA is guaranteed to find a unique set of contrastive features. Given a small inventory, the SDA will stop sooner; a larger inventory will require more splits. In every case, the SDA assigns a minimal set of specifications that meet the Distinctness Condition.

Therefore, I conclude that of the two methods for arriving at contrastive specifications, the pairwise method is inadequate on purely logical grounds, whereas feature ordering is logically sound.

The hierarchical approach to contrastive specification imposes a task on language learners and analysts that pairwise comparison does not: it requires that the features be ordered. Although the feature order, or contrastive hierarchy, is crucial to the functioning of the SDA, it is not itself discovered by the SDA. Where does the ordering come from? I will return to this important issue in subsequent chapters.

2.7 Other issues in a theory of contrastive features

The above sections have been concerned with how one decides which features are contrastive in a set of phonemes: whether one does so in terms of minimal pairs and logical redundancy or by feature ordering. This is the most fundamental issue in an investigation of the logic of contrast. However, there are other issues that arise in a theory of contrastive features that bear some discussion here, because they interact with the logical questions considered above, and in some cases may produce results that look different from the ones we arrived at above. In the rest of this chapter I consider three such issues: whether contrast is to be assessed in an inventory as a whole or is limited by position (section 2.7.1); whether features (binary features, in this case) have two values or one value (section 2.7.2); and the relationship between contrastive specification and theories of underspecification (section 2.7.3).

2.7.1 Contrast limited by position

Up to here we have viewed contrasts as being defined over the entire inventory: we have considered, for example, the contrastive specifications of phonemes /p, b, m/ as if these were fixed once and for all for the whole language. But the sounds of a language are arranged *syntagmatically* as well as *paradigmatically*, and phonotactic restrictions can alter the set of contrasts at particular positions in a language. English, for example, observes the restriction that the first in an initial sequence of three consonants must be *s*: hence, *splash*, *stretch*, *squat*, are well-formed words, but no English word results from substituting another consonant for *s* in these examples. If we evaluate the contrastive status of *s* in this position only, it would suffice to specify it as [+consonantal]. Thus, contrastive evaluation limited by position can yield different results from evaluating contrasts globally.

Should we evaluate contrasts globally or by position? And if by position, how do we define the positions? In the above example we singled out a position before two consonants and following a word boundary; clearly, there are many other possible positions of varying degrees of specificity. We could, for example, focus on initial single consonants in monosyllables (*pick*, *tail*, *comb*, *fuse*, etc.), or limit the vowel to [ɪ] (*pick*, *rig*, *win*, *fill*, etc.), or further limit the final consonant to [k] (*pick*, *tick*, *wick*, *sick*, etc.). A variety of prosodic and morphological considerations may also play a role, allowing us to distinguish between stressed and unstressed syllables, stems and affixes, and so on.[10]

10 See Beckman (1997) and Dresher and van der Hulst (1998) for two approaches to picking out positions that have a special status in the phonology.

As in other matters concerning contrast, phonologists have not been consistent in this regard. Unless the sets of relevant positions can be somehow limited in reasonable ways, however, positionally limited contrastive evaluation can get out of hand. Therefore, for the most part I will continue to assume that contrasts are defined globally for phonemes in inventories. However, we will return to this topic in section 7.5, where we will see that there is a place for positionally defined contrasts in phonology.

2.7.2 Types of features: equipollent and privative

The SDA works on all types of features. For purposes of exposition I have been assuming binary features, but this assumption is not crucial to the validity of the arguments against the pairwise method. However, the type of feature adopted can affect the results produced by the SDA, and for this reason it is worth considering this issue here.

There are different kinds of binary features.[11] The kind discussed above have two values, one positive and one negative. As long as we do not attribute special status to + or −, the two values of a binary feature have equal status. To borrow the Prague School term (Trubetzkoy 1969), such features are *equipollent*.

The [+F] ~ [−F] notation introduces an inherent asymmetry, however: [+voiced] feels psychologically different from [−voiceless], because each names the feature after a different one of its values. It is a small step to suppose that the two values are not equal in status. One could be the *default*, or *unmarked*, value, and the other could be the *marked* value. The terms 'marked' and 'unmarked' are also borrowed from the Prague School, who took it somewhat literally, as meaning that an unmarked feature value is simply not indicated, whereas a marked value is indicated by a mark. This kind of contrast can be represented as Ø ~ [F], where Ø represents the *absence* of a mark and [F] is the marked value. In Prague School terminology, this kind of binary contrast is called *privative*.

Privative contrasts impose more structure on representations than equipollent ones, and hence require more information. To make an equipollent contrast between nasal and oral, it is enough to write [+nasal] ~ [−nasal], or, equivalently, [−oral] ~ [+oral] (which name we choose has no significance). To make a privative contrast, we have to decide which is the marked feature.

11 It is also possible to have *multi-valued* features, either *discrete* or *continuous*. There have been proposals that multi-valued, even continuous, features play a role in linguistic theory (Broe and Pierrehumbert 2000; Pierrehumbert, Beckman and Ladd 2000). Nevertheless, I will continue to assume that phonological features are discrete and mainly binary.

Privative features act differently from equipollent ones with respect to contrast. For the sake of discussion, let us suppose that the marked values for the features [voiced] and [nasal] are the positive values in both cases. The *full* (not contrastive) specifications of the simple inventory in (2) will now look like (29).

(29) Full specifications: privative features

	p	b	m
[voiced]		✓	✓
[nasal]			✓

In the fullest possible set of specifications, /p/ is completely unmarked and has no specifications. Looking at the inventory as a whole, the 'full specifications' of (29) look like very minimal contrastive specifications. However, these are not contrastive specifications in the sense of the previous section.

Let us consider what effects contrastive feature ordering has on the specifications in (29). If [voiced] is the first feature, we mark /b/ and /m/ for this feature and leave /p/ unmarked. The next feature, [nasal], distinguishes between /b/ and /m/, and marks /m/ as [nasal]. Now we have the specifications in (30a), which are the same as in (29). That is, the contrastive specifications are the same as the full specifications in this ordering of the features. Proceeding in the other order, we first mark /m/ [nasal] in contrast to /p/ and /b/, which are unmarked; we then draw a contrast between /p/ and /b/ by marking /b/ [voiced], deriving the contrastive values in (30b). In this order, one of the full specifications is omitted.

(30) Contrastive specifications: privative features

a. [voiced] > [nasal]				b. [nasal] > [voiced]		
p	b	m		p	b	m
	✓	✓	[voiced]		✓	
		✓	[nasal]			✓

We observe that the effect of feature ordering is greatly reduced with privative features as opposed to equipollent features. This is because privative features conflate two situations that are distinct in equipollent features; the two are compared in (31).

(31) Contrastive specifications with equipollent and privative features

a. Equipollent features
Member M is contrastively specified for a feature F iff M contrasts with at least one other member with respect to F.

b. Privative features
Member M is contrastively specified for a feature F iff M contrasts with at least one other member with respect to feature F, *and* M is marked for F.

Looking at it from the other side, M will remain unspecified for F in a privative system if *either* (i) F is not contrastive in M, *or* (ii) M is unmarked for F; whereas in an equipollent system, M will remain unspecified for F in case (i), but will receive a value for F in case (ii). This means that in a privative system we cannot tell from the representations which unmarked segments are contrastive; nor can we reconstruct what the scope of a contrast is, because only the marked members of a contrast receive a feature value, leaving it unclear which of the phonemes that are unmarked for a feature are in the scope of the contrast and which fall outside it. It follows that if it is important to know the scope of a contrast and which segments it affects in a privative feature system, we will have to keep track of this information with some machinery in addition to the representations themselves.

2.7.3 Contrast and underspecification

There is a natural, but by no means necessary, connection between contrast and underspecification. In a theory where contrastive feature specifications are assigned hierarchically by the SDA, it is natural to suppose that contrastive specifications are specified and redundant specifications are unspecified. Consider again our example of bilabial stops /p, b, m/, assuming an ordering [nasal] > [voiced]; the contrastive specifications are as in (5). It is natural to assume that the contrastive feature values in (5) are specified whereas the redundant values (in this case, the feature [+voiced] for /m/) are unspecified. This is not necessarily the case, however. We have seen that it is not necessary for all contrastive values to be specified. In a privative feature system, only *marked* contrastive values are specified, as in (30b).

In (30), representations are underspecified beyond the requirements of contrast, by omitting also unmarked contrastive specifications. The converse is also theoretically possible: representations may be specified over and above the requirements of contrast. Thus, it is possible to interpret the SDA not as an algorithm that *assigns* feature values in contrastive fashion, but rather as an algorithm that *designates* which values are contrastive. In such a theory, all possible feature values are always present, but some of them are designated as being contrastive. In this kind of theory, the specifications in (5) can be viewed as shorthand for the more complete listing in (32); specifications designated $_{C}$ are contrastive.

(32) All specifications, [nasal] > [voiced], $_C$ = contrastive

	p	b	m
[voiced]	$-_C$	$+_C$	$+$
[nasal]	$-_C$	$-_C$	$+_C$

We can take a similar approach to markedness; rather than assume that only marked values are specified, as is the case with privative features, we can designate which values of each feature are marked, as in (33). In such a theory, the phonology has the option of targeting all features, or contrastive features, or marked features; this is the approach of Calabrese (2005) and Nevins (2004), discussed further in section 8.6.

(33) All specifications, [nasal] > [voiced],], $_C$ = contrastive, $_M$ = marked

	p	b	m
[voiced]	$-_C$	$+_{C,M}$	$+_M$
[nasal]	$-_C$	$-_C$	$+_{C,M}$

2.8 Conclusions: one approach to contrast left standing

In this chapter we have looked at how we might determine which features are contrastive in a given phoneme. I have identified two approaches to this question: the pairwise approach, based on making comparisons of fully specified phonemes with special attention to minimal pairs, and contrastive specification by feature ordering. The pairwise approach identifies as contrastive only specifications that are not logically redundant. While this may seem to be a point in its favour, I have argued that it is actually the source of a number of insurmountable problems.

The feature ordering approach, on the other hand, poses no logical difficulties. This approach is based on the notion that the scope of a contrast depends on where a feature is ordered in the hierarchy: features ordered higher up take wider scope than features ordered lower down. The feature ordering approach is not dependent on any particular distribution of minimal pairs, and separates the notions of logical redundancy and system redundancy.

Another difference between the two approaches is that pairwise comparison always produces the same results, given an inventory and a set of features, whereas feature ordering can give different results depending on the ordering. Again, this property might at first seem to favour pairwise comparison, for it is automatic and imposes less of a burden on the analyst as well as on the learner, who must determine the correct ordering of features in the latter approach. I will argue, however, that the advantage here again is on the side of feature ordering, for there is empirical evidence that similar-looking inventories can indeed have different contrastive specifications.

I conclude, then, that the SDA applied to features ordered into a contrastive hierarchy must be the basis of any theory of phonological contrast.

Another issue in the determination of contrast involves the syntagmatic dimension, the extent to which contrast is evaluated globally over an inventory or is tied to particular positions. From a logical point of view there is no way to decide this question. I will argue in section 7.5 that there are practical constraints that limit the degree to which contrasts can be tied to particular syntagmatic contexts; but where these constraints do not hold, there is evidence for contrasts limited to certain positions.

An issue orthogonal to the evaluation of contrast is the question of whether features are equipollent or privative. This issue has important implications for the identification of contrasts and the scope of a contrast. Also orthogonal to contrastive specification is whether redundant features are entirely absent from the phonology.

A study of the logical problem of contrast can take us only so far. The important question, from the point of view of phonology, is what role, if any, contrastive specifications actually play in phonological theory, and the extent to which a theory of contrastive specifications helps to illuminate phonological phenomena. In the next chapter we will find some preliminary answers to these questions in the work of some major figures in the formative years of phonological theory.

3 *Contrast in structuralist phonology*

3.1 Introduction

In the preceding chapter, I identified two basic approaches to contrastive specification in phonology, and discussed their status from a logical point of view. I argued that pairwise comparison based on full specifications has severe logical problems, whereas contrastive specification by feature ordering is logically sound. I showed that Martinet's (1960) analysis of Standard French bilabial stops is an example of the former approach, and that Jakobson and Lotz's (1949) analysis of the same phonemes exemplifies the latter approach.

In this chapter I will review some work in structuralist phonology that bears on this issue. This chapter has three main aims. First, I wish to show the extent to which contrastive specification was central to the project of phonological theory in its formative years, roughly the period 1925–50. Though the authors I will survey are central figures in the field whose work has been widely read, I believe that this work has been misconstrued in various ways. There are several reasons for this. First, the authors sometimes use terminology that is unfamiliar to contemporary readers. Second, they are not always very explicit about how their theory is supposed to work, and in some cases, their writings do not add up to a consistent theory. Third, we tend to read these works through the prism of our own preoccupations.

The second aim of this chapter is to investigate in detail the narrower question of how these phonologists determined which features of a phoneme are contrastive in a given language; that is, how they arrived at contrastive specifications. One might come away from the previous chapter with the impression that Martinet and Jakobson represented two different theoretical positions (pairwise comparison versus feature ordering), and one might expect to see arguments for and against these positions. However, this impression would be incorrect. I will show on the contrary that explicit procedures for determining contrasts were not formulated, and the authors I survey did not adopt clear-cut or even consistent positions on this issue. Nevertheless, it is possible to see the two basic approaches to contrast throughout this period.

This review of early work in phonology proceeds not just from a historical interest. The third and most important aim is to recover fundamental insights into the nature of contrast and its role in phonology that should form a part of contemporary phonological theory. To this end, I will show the existence of what appears to be a recurring pattern: when these theorists were thinking abstractly about contrast, they tended to assume something like pairwise comparison; but when they had empirical reasons for proposing contrastive specifications, they tended to apply something like feature ordering. This result supports the finding in the previous chapter that feature ordering is the only viable approach to contrastive specification. Moreover, we will see that it is indispensable in accounting for phonological patterning.

3.2 Sapir: phonetics versus phonological patterning

A central theme in the work of Edward Sapir is the distinction between a phonetic description of the sounds of a language and the way these sounds 'pattern' in the phonology (Sapir 1925, 1933). Sapir emphasizes that the phonological patterning, or 'pattern alignment', of a speech sound can be different from what we might expect from the phonetics. But what does Sapir mean by 'sound pattern'? Through the prism of generative phonology, this term is usually understood, in a wide sense, to be the collective set of rules and representations that make up the phonology of a language, and, in a narrower sense, to refer to the underlying representation of a sound. It can be argued, however, that the 'pattern alignment' of a phoneme refers most specifically to the set of its contrastive properties.

Sapir (1925) presents interesting examples of how systems that appear to be phonetically similar pattern very differently (his languages A and B), and, conversely, how languages that may appear to be phonetically quite different can be very similar in their phonemic patterning (languages C and D). The languages are constructs of Sapir's, though they are based on cases with which he was familiar. Sapir's languages C and D are shown in (1).[1]

1 I have changed Sapir's phonetic symbols to accord with current usage. Next to *q*, Sapir writes '(velar k)', and similarly for all the obstruents in the fourth column. I assume he means uvular rather than velar. Similarly, I represent his 'h (laryngeal ḥ)' as ħ. In this book, I transcribe italicized *a* in my sources as a, except where something like IPA *a* is intended.

(1) Different phonetics, similar patterning (Sapir 1925)

a. Pattern of C

a			ɛ		i			u	
aː			ɛː						
	h			w	j		l	m	n
p		t		k		q			
b		d		g		ɢ			
f		s		x		χ			

b. Pattern of D

æ			e		i			y	
æː			eː						
	h			v	ʒ		r	m	ŋ
p^h		t^h		k^h		q^h			
β		ð		ɣ		ʁ			
f		ʃ		ç		ħ			

Sapir points out that, if one were to be guided only by phonetics, one might suppose that [ʒ] in Language D should be listed under [ʃ] as its voiced counterpart, just as [b] is placed under [p] in system C. Similarly, we might expect that [v] in D should be placed under [f] as its voiced counterpart. Sapir allows that the 'natural phonetic arrangement' of sounds is a useful guide to how they pattern, but he goes on: 'And yet it is most important to emphasize the fact, strange but indubitable, that a pattern alignment does not need to correspond exactly to the more obvious phonetic one.'

It is worth inquiring a bit more closely into the significance of Sapir's inventory charts. We observe first that he specifies no features or other substantive specifications of the listed sounds (except for the parenthetical comments about the obstruents in the fourth column). Therefore, the basis of the similarities in the patterns of C and D is, in the first place, in the typographical arrangement of the sounds. C and D have isomorphic patterns because each segment in one corresponds to a unique segment in the other, and the relative position of each segment in the pattern is the same as that of its correspondent.

But what is the meaning of these typographical layouts? They surely imply some phonological feature specifications, though these are not listed explicitly. Moreover, by arranging the systems of C and D isomorphically, Sapir is suggesting that the specifications for the two languages are parallel, if not identical. He writes that the reasons for placing [ʒ] (which he represents as *j*) in D where it is may have to do with the alternations it enters into (it might alternate with /i/

but never with /ʃ/), or the phonotactic combinations it enters into (for instance, that *v*- and *ʒ*- are possible initials, like /r, m, ŋ/, but that /β, ð, ɣ, ʁ/ are not). He continues, 'In other words, it "feels" like the y [= IPA *j*] of many other languages, and as y itself is absent in D, we can go so far as to say that j [=[ʒ]] occupies a "place in the pattern" that belongs to y elsewhere.'

The notion of 'place in the pattern' can be interpreted in a number of ways. One interpretation in terms of generative theory is that the corresponding segments in C and D have the same underlying specifications, which, in classical generative phonology, are full, not contrastive, feature specifications. Thus, we could understand Sapir to be suggesting that the lexical (underlying) phonological specifications of the phonemes /v/ and /ʒ/, that appear to be in the wrong place in the pattern of D, are to some extent at odds with their phonetics. In derivational generative terms, we can justify their positions in the pattern of D by assuming that they are specified as sonorants rather than obstruents, just like /w/ and /j/ in C, and assume their phonetic forms by late rules that alter some of their specifications. On this interpretation, /v/ and /ʒ/ in D correspond to /w/ and /j/ in C because they *are* those sounds at an abstract level of analysis.

While such an analysis is tenable in these cases, it does not extend to the rest of the phonemes whose phonetics do not deviate so spectacularly from their positioning in the phonological pattern. According to Sapir, /b/ occupies a place in the pattern of C that corresponds to the place of /β/ in D. In this case we have no reason to suppose that one of these sounds derives from the other. Looking at their fully specified lexical representations would not show us in what way they can be said to occupy the same position in the pattern of their respective languages. Rather, in this case, we must understand the notion of 'place in the pattern' to refer to the contrastive role of speech sounds. What /b/ in C and /β/ in D have in common is that they are both the only voiced labial obstruents in their respective languages; no further specifications are required to distinguish them from every other phoneme. That /b/ is also a stop whereas /β/ is a fricative is not relevant to their contrastive positioning, in this analysis.

Similarly, /l/ in C corresponds to /r/ in D because each is the only liquid in the language, and /n/ in C corresponds to /ŋ/ in D by being a non-labial nasal consonant. There is no suggestion that the underlying forms of the corresponding phonemes must be identical.

More generally, both languages C and D can be said to have a series of contrastively voiceless stops (with redundant aspiration in D), a series of contrastively voiced obstruents (which are redundantly stops in C and spirants in D), and a series of contrastively voiceless spirants. Both languages contrast four places of articulation for each obstruent series: labial, coronal, palato-dorsal and post-dorsal; further specification within these broad place categories is not important to the phonological patterning of the system.

Similar considerations hold for the vowels. Both languages have two high vowels, one unrounded and the other rounded; whether the latter is front or back is not significant. Each language also has one mid vowel and one low vowel, the exact tonality of which is not important phonologically. Each language also has long vowels corresponding to the non-high short vowels.

The contrastive patterning common to these languages is shown in (2). In each cell, the first phoneme is from C, and the second is from D. Phonemes that share a cell have the same contrastive specifications.

(2) Contrastive patterning of languages C and D

a. Short vowels

unrounded	rounded	
i/i	u/y	high
ɛ/e		mid
a/æ		low

b. Long vowels

	high
ɛː/eː	mid
aː/æː	low

c. Consonants

			labial	coronal	palato-dorsal	post-dorsal
obstruent	voiceless	stop	p/pʰ	t/tʰ	k/kʰ	q/qʰ
		spirant	f/f	s/ʃ	x/ç	χ/ħ
	voiced		b/β	d/ð	g/ɣ	ɢ/ʁ
sonorant	nasal		m/m	n/ŋ		
	liquid		l/r			
	glide		w/v	j/ʒ		h/h

Sapir's discussion lacks formal rigour and a system of features, but we can recognize in it some seminal ideas that I would like to build on later. First is the notion that not all properties of a sound are equally important, but that certain ones – the contrastive ones – are particularly relevant to the phonology. Second, the contrastive status of a phoneme may differ from what its phonetics might lead us to think; that is, the phonetics of a segment is a guide to its distinctive properties, but is not sufficient to indicate what these are, and may sometimes even be at odds with its phonological status. Finally, we determine what the contrastive properties of a phoneme are by the phonetics in combination with its phonological *behaviour*. This behaviour could consist of its phonotactic restrictions, or the way it alternates, or the effects it has on other phonemes.

3.3 Trubetzkoy: a theory of oppositions

The phonologist who did the most to develop the principle of contrast as an organizing principle of phonology was N. S. Trubetzkoy. His *Grundzüge der Phonologie*, written in the 1930s, is a major statement of the Prague School approach to phonology. In this work, Trubetzkoy's aim is to present an exhaustive account of the various types of contrastive oppositions that can exist in a phonological system. The influence of his approach can be found in many subsequent schools of linguistics. Therefore, Trubetzkoy's approach to contrast is central to our investigation.

Trubetzkoy made major contributions to our understanding of how contrast functions in phonological systems. One of his key insights is that the determination of contrastive features in an inventory is not self-evident but must be established by the analyst on the basis of the patterning of the phonological system. I will also show that the notion of feature hierarchy as a way to depict contrastive relations makes an appearance – perhaps its earliest appearance – in the *Grundzüge*. However, his account of contrastive relations is crucially incomplete, because it is not explicit with respect to how contrasts are assigned. More than that, I will show that Trubetzkoy (1939) does not follow a consistent approach to contrast, but varies between the pairwise approach and feature ordering. The former predominates in the earlier parts of the book, where the discussion tends to be abstract and mainly theoretical, with no direct empirical consequences. In later sections, where Trubetzkoy has empirical evidence for proposing certain contrastive relations, he uses an approach consistent with feature ordering.[2]

3.3.1 Phonologically relevant features and phonemic content

Every phoneme of a language enters into an *opposition* with every other phoneme. It is important to bear in mind that an opposition is a relation between

2 I am grateful to William Sullivan for reminding me that Trubetzkoy died before he was able to complete his book, a fact that may account for some of the inconsistencies noted here.

a *pair* of phonemes. It is not just the number of oppositions, but their particular characters that give structure to a phonological system.

From the outset, Trubetzkoy distinguishes between *phonologically relevant* and *phonologically irrelevant* features of speech sounds: ‘every sound contains several acoustic-articulatory properties and is differentiated from every other sound not by all but only by a few of these properties’ (p. 35).[3] As an example he adduces the German phonemes *k* (/k/) and *ch* (/x/). The latter has two allophones, [x] (what Trubetzkoy calls the ‘*ach* sound’) and [ç] (the ‘*ich* sound’). The sounds that belong to the /k/ phoneme are distinguished in manner of articulation from /x/ by forming a complete closure, whereas /x/ sounds form a stricture; that is, /k/ is distinctively a stop, whereas /x/ is a fricative. As to place of articulation, the fact that the opposition between [x] and [ç] is nondistinctive ‘presents evidence that for *ch* the occurrence of a stricture between dorsum and palate is *phonologically relevant*, while the position of stricture in the back or central dorsal-palatal region is *phonologically irrelevant*’ (p. 36).

From this example we can also see Trubetzkoy’s approach to determining what the distinctive marks of a phoneme are. These are the marks that abstract away from the variation of the surface allophones. Thus, in the case above, variation in the place of articulation of the allophones of German /x/ provides evidence that its place must be specified only as generally dorsal. Similarly, the more radical variation in the place of articulation of German /r/, which is sometimes alveolar and sometimes uvular, reveals that place is a completely irrelevant property of this phoneme.

This diagnostic does not apply in cases of neutralization of underlying distinctive contrasts.[4] For example, in certain positions German voiced obstruents become voiceless, causing /g/ to become [k], which is identical to [k] that derives from phoneme /k/. Since the contrast between /g/ and /k/ is suspended in these positions, it is not possible to define the contrastive features of /g/ in such a way that they abstract away from the variation between [g] and [k]. Allophonic variation is not a necessary condition for determining the phonologically relevant features of a phoneme. We will see below that Trubetzkoy also appeals to other aspects of a sound’s phonological behaviour in discovering its relevant distinctive features.

3 All quotes drawn from the English translation by Baltaxe, cited as Trubetzkoy (1969).

4 As we shall see, neutralization itself is Trubetzkoy’s primary diagnostic for bilateral oppositions, and hence for the structuring of contrasts in an inventory.

The phonologically relevant marks of a phoneme make up its *phonemic content*. 'By phonemic content we understand all phonologically distinctive properties of a phoneme, that is, those properties which are common to all variants of a phoneme and which distinguish it from all other phonemes of the same language, especially from those that are most closely related' (p. 66). Again, this must exclude neutralized variants of phonemes. More important, the phonemic content of a phoneme is a function of the contrastive oppositions it enters into: 'The definition of the content of a phoneme depends on what position this phoneme takes in the given phonemic system, that is, in final analysis, with which other phonemes it is in opposition . . . Each phoneme has a definable phonemic content only because the system of distinctive oppositions shows a definite order or structure' (Trubetzkoy 1969: 67–8).

The above remarks suggest that the phonemic content of a phoneme, that is, the set of its distinctive (contrastive) properties, ought to *derive* from its position in the system of distinctive oppositions. Therefore, we need a way to determine a phoneme's position in the system of oppositions *before* we have determined its distinctive properties. But Trubetzkoy does not explicitly show us how to do this.

Consider again his comments on the German phoneme /r/, cited above in section 1.2. Trubetzkoy (1969: 73) observes that the phonemic content of German *r* is 'very poor, actually purely negative: it is not a vowel, not a specific obstruent, not a nasal, nor an *l*.' How did Trubetzkoy arrive at this conclusion? First, he is assuming a theory of *markedness* wherein one value of a feature is marked (positive) and the other is unmarked (negative). His discussion assumes the markedness values in (3); the feature names are not Trubetzkoy's, but I have chosen them so that the marked value is the positive (+) one.

(3) Markedness of features of German /r/

Feature	*Marked*	*Unmarked*
[obstruent]	obstruents	sonorants
[nasal]	nasals	liquids
[lateral]	laterals	rhotics

Markedness is one ingredient we require to reconstruct Trubetzkoy's analysis, but we need to answer a further question: how did he pick these particular features, and only these, to distinctively characterize German *r*? One way we can arrive at this result is by successively dividing up the German consonantal inventory, by features, following the order given in (3). The phonemes shown in (4) are those listed by Trubetzkoy, and the layout of the chart is based on his remarks.

(4) German consonantal phonemes

p	pf	t		ts			k	
b		d					g	
	f			s	ʃ		x	h
	v			z				
m		n					ŋ	
			l			r		

We can distinguish /r/ from every other phoneme by this procedure:

1. First, divide the inventory by the feature [obstruent], which distinguishes between obstruents (above the line in (4)) and sonorants (below the line). Since *r* is a sonorant, this feature distinguishes it from all obstruents, which no longer need be considered with respect to uniquely characterizing /r/.
2. Among the sonorants the feature [nasal] distinguishes the nasal consonants (in the box) from the non-nasal sonorants, leaving *r* in contrast only with *l*.
3. The final feature, [lateral], distinguishes between *l* (circled) and *r*, and leaves them both with a unique set of features.

The above procedure meets the requirement that the phonemic content of a phoneme, that is, the set of its distinctive (contrastive) properties, follows from its position in the system of distinctive oppositions. Moreover, in this procedure, 'the system of distinctive oppositions shows a definite order or structure'. The order in question is the order of the features, which gives structure to the inventory.

It follows that feature ordering gives us a way to reconstruct what Trubetzkoy may have meant by the statements cited above. However, it is not possible to state definitively that this is indeed what he intended, for he does not give any explicit procedure for how he arrived at his analysis of German *r*. Moreover, some of the other examples he discusses do not appear to work the same way, and other statements are inconsistent with feature ordering.

In the following sections I will first give some examples consistent with the assumption that Trubetzkoy determined contrasts by means of pairwise comparisons. I will then review other examples that are more consistent with contrast through feature ordering.

3.3.2 Some examples that imply the pairwise method

In chapter 3, Trubetzkoy presents what he calls a 'logical classification' of distinctive oppositions. He makes clear that he intends the principles presented in this chapter to be applicable to any systems with contrastive elements made up of features, though his examples are almost all drawn from phonology.

Oppositions are characterized by the properties that distinguish the opposition members, as well as by the properties the members have in common. In a *bilateral* opposition, the sum of the properties common to both opposition members is common to them alone. In a *multilateral* opposition, the basis of comparison is not limited exclusively to the two opposition members.

A question immediately arises: in comparing the opposition members, do we consider *all* their properties, or only their *distinctive* properties? Trubetzkoy's initial answer is decisive (1969: 68): 'Of course, only the phonologically distinctive properties are to be considered.' But he goes on: 'However, some nondistinctive properties may be taken into consideration as well if, on the basis of these properties, the members of the opposition in question are placed in opposition with other phonemes of the same system.' This latter qualification muddies the waters considerably. In this case, Trubetzkoy (1969: 69) presents an example to illustrate why he wants to allow nondistinctive features to play a role: '[T]he opposition *d*-*n* (as in French) is to be considered bilateral because its members are the only voiced dental occlusives. Yet neither voicing nor occlusion is distinctive for *n*, as neither voiceless nor spirantal *n* occur as independent phonemes.'

The only way to make sense of Trubetzkoy's remarks is to assume that he looked for minimal pairs in order to determine which features are contrastive. In particular, we can make the pairwise comparisons in (5).

(5) French consonants: pairwise comparisons

a. *n* ∼ *m* are distinguished by [dental] (or another place feature).
b. *n* ∼ *d* are distinguished by [nasal].
c. These two features suffice to distinguish *n* from every other phoneme as well, regardless of any other distinctions that may exist.
d. The only way *n* could be contrastively [voiced] in a pairwise procedure is for there to exist a voiceless segment with the same full feature specifications as *n* except for [voiced] – that is, a voiceless *n*.
e. Similarly, *n* could be contrastively [non-continuant] if there existed a fricative *n*, identical to *n* in all other features.

To confirm that Trubetzkoy must be (tacitly) following the pairwise method here, it suffices to observe that feature ordering does not give this result. If we used that procedure here, we could, for example, order the feature [voiced] first, distinguishing between voiced (boxed in (6)) and voiceless consonants.

(6) French consonantal phonemes

p	t					k	
b	d					g	
f			s	ʃ			
v			z	ʒ			
m	n				ɲ		
		l					r
					j		

Since both *d* and *n* are voiced, following this division we would only have to subsequently distinguish among the voiced consonants. Continuing to use Trubetzkoy's terms, the next feature could be [occlusive] (horizontal boxes in (7)), which we apply to the set of voiced consonants (voiceless consonants are no longer relevant to distinguishing *n*). Finally, the feature [dental] (vertical box in (7)) narrows the set to just *d* and *n*.

(7) French voiced consonantal phonemes

b	d					g	
v			z	ʒ			
m	n				ɲ		
		l					r
					j		

Note that under this procedure the features for voicing and occlusion *are* distinctive for *n*; in fact, together with [dental], they are the *only* distinctive features so far assigned to *n*. Since Trubetzkoy does not consider the possibility that these features are contrastive in *n*, it must be that he was not thinking of the feature ordering method in this discussion of French.[5]

Further examples of what appears to be extraction of contrasts based on pairwise comparisons can be shown to occur in Trubetzkoy (1939), particularly from the front of the book. To mention just one more here, his analysis of bilateral oppositions in German vowels (Trubetzkoy 1969: 69–70) can be shown to be based on pairwise comparisons applied to fully specified

5 We will see below that Trubetzkoy argues that occlusion does not play any contrastive role in French; rather, he argues that stops and fricatives are distinguished by place, not manner.

representations, and are inconsistent with any feature ordering approach to contrastive specification.

It should be noted that Trubetzkoy adduces no empirical evidence in support of these analyses; thus, it is unclear what consequences, if any, flow from them. This is in striking contrast to the examples to be discussed in the following section, which are all accompanied by empirical justifications.

3.3.3 Examples of hierarchy in Trubetzkoy

One might conclude from the above examples that our earlier interpretation of Trubetzkoy's theory of contrast is simply wrong. Nevertheless, I do not believe that this position can be consistently maintained. For one thing, the notion of 'relevant contrast' is central to Trubetzkoy's entire exposition. If, as he writes in connection with the *d* ∼ *n* opposition, oppositions must normally be established using only contrastive features, the implication is that the contrastive status of a specification is established *before* oppositions are classified as bilateral or multilateral. But such a notion is incompatible with deriving bilateral pairs from full (noncontrastive) specifications.

Second, as we have seen, pairwise comparisons do not yield a result for oppositions that differ by *more* than a single feature: for such cases, which make up the majority of oppositions, the procedure does not decide which feature is the contrastive one and which are redundant.

Consider again Trubetzkoy's analysis of German *r* (see (4)). Recall that Trubetzkoy assigns German *r* the distinctive (negative) features [non-lateral], [non-nasal] and [non-obstruent]. We have seen that *l* ∼ *r* are distinguished by the feature [lateral]. But, except for this pair, the pairwise method yields unclear results.[6] What feature distinguishes *r* from *ŋ*, for example? It could be [nasal], as in Trubetzkoy's analysis; but it could also be an occlusion feature (nasals are stops, in contrast to *r*), or a place feature. Similarly, *r* may be distinguished from *z* by [obstruent], as in Trubetzkoy's analysis, but this is not the only feature that distinguishes these phonemes; other candidates are [strident], or place of articulation. Such choices arise with respect to almost every opposition.

Third, there are places where Trubetzkoy makes crucial use of the notion that bilateral oppositions are a function of the system of contrasts, and that the same inventory can be viewed in different ways, depending on what contrasts have been established. These examples can all be easily reconstructed in terms of a contrastive hierarchy, but not in terms of pairwise comparisons. It is to these cases that we now turn.

6 Even for this pair there are other possible contrasts, such as [continuant].

3.3.3.1 Consonant systems: place versus manner and voicing

Many examples come from Trubetzkoy's discussion of consonant systems. We frequently find two consonants that differ slightly in place of articulation (e.g. bilabial vs labiodental, or dorsal vs laryngeal) as well as in another dimension such as manner (stop vs fricative) or voicing. It is a recurring question whether the contrast is primarily one of place of articulation, from which the other differences are derived, or whether the other dimension is the determining difference. Trubetzkoy resolves these oppositions differently, depending on how the particular opposition fits into the overall system, in a way that implies some hierarchical organization of the relevant features.

Trubetzkoy recognizes a basic series of place contrasts that includes the gutturals (or dorsals), the apicals (dentals) and the labials. To these he adds the sibilants. Some languages have other basic series. He lists the lateral, labiovelar, palatal and laryngeal series.[7] The key point is that 'the phonological concept of series of localization must not be confused with the phonetic one of position of articulation' (p. 124).

German and Czech h Consider, for example, Trubetzkoy's treatment of German and Czech *h*. According to Trubetzkoy (1969: 69), German *h* does not take part in any bilateral oppositions. In particular, it is not in a bilateral opposition with *x*: *h* is laryngeal and *x* is dorsal, so there is no set of features that the two share exclusively. Looking at the Czech consonant inventory in (8),[8] one might suppose that Czech h (more properly, *ɦ*) is similarly isolated.

(8) Czech consonant phonemes: *h* in a separate laryngeal series

p	t	c	k	
b	d	ɟ	g	
	ts	tʃ		
f	s	ʃ	x	
v	z	ʒ		ɦ
m	n	ɲ		
	r	r̝		
	l			
		j		

However, Trubetzkoy (1969: 124) proposes that *ɦ* forms a bilateral opposition with *x*. His reason is that the distinction between these phonemes can be

7 In contemporary feature theory, the features [sibilant] and [lateral] are not typically considered to be places of articulation.

8 Trubetzkoy does not provide an explicit account of the Czech phoneme inventory; the charts in (8) and (9) are based on D. C. Hall (2007: 38).

neutralized, for they behave phonologically like a voiced–voiceless pair, like the other such pairs in Czech: 'The *h* in Czech thus does not belong to a special laryngeal series, which does not even exist in that language. It belongs to the guttural series, for which, from the standpoint of the Czech phonological system, only the fact that lips and tip of tongue do not participate is relevant.' That is, we should diagram the Czech consonants as in (9) rather than as in (8).

(9) Czech consonantal phonemes: *h* part of the guttural series

p	t	c	k
b	d	ɟ	g
	ts	tʃ	
f	s	ʃ	x
v	z	ʒ	ɦ
m	n	ɲ	
	r	r̝	
	l		
		j	

The difference in the contrastive status of German *h* and Czech *ɦ* does not emerge from pairwise comparisons of the phonetic properties of these phonemes with other phonemes in the system. Rather, it is the *phonological behaviour* of these phonemes that is the key to the analysis of their phonological content. Whereas pairwise comparison tells us nothing about the difference between the German *h* ~ *x* opposition and the Czech *ɦ* ~ *x* opposition, we can use feature ordering to implement Trubetzkoy's analysis and capture this distinction. In German, if the feature [laryngeal] is ordered relatively high in the list, it will distinguish *h* from every other consonant, including *x*; therefore, *h* participates in no bilateral oppositions. In Czech, [laryngeal] would be lower in the order; instead, a feature [guttural] (perhaps characterized negatively as [non-coronal] and [non-labial]) and the voicing feature are ordered higher. As there are no distinctive place differences between *ɦ* and *x*, their opposition is bilateral.

As Trubetzkoy (2001 [1936]: 20) remarked in his 1936 article addressed to psychologists and philosophers, the correct classification of an opposition 'depends on one's point of view'; but 'it is neither subjective nor arbitrary, for the point of view is implied by the system'. Feature ordering is a way to incorporate 'point of view' into the procedure of determining contrastive properties. Different orders result in different contrastive features, as is the case with German *h* and Czech *ɦ*.

Greek and French labials In German and Czech, different feature orderings resolve whether the laryngeal consonant makes up a distinct place of articulation, or whether it is to be regarded as part of a more general guttural series. There are other differences between German and Czech that illustrate the same point. For example, Trubetzkoy (1969: 125) proposes that the German bilabials *p*, *b* and *m* form a series distinct from the labiodentals *v*, *f* and *pf*; he makes no such claim for Czech, where the labial consonants presumably make up a single series.

This issue arises explicitly in Trubetzkoy's discussion of Greek and French labial consonants. In Greek, labial and apical stops and fricatives differ in place as well as in occlusion: the fricatives /f, v/ are labiodental in contrast to the bilabial stop /p/, and fricatives /θ, ð/ are interdental in contrast to the stop /t/. The major contrast between these stops and fricatives could thus be based either on place or on occlusion. Trubetzkoy appeals to 'parallel' relations between stops and fricatives at different places of articulation. In the sibilant and dorsal series, /ts, s, z/ and /k, x, ɣ/, respectively, the contrast is unambiguously one of stop versus fricative, since stops and fricatives occur at exactly the same place of articulation. By parallelism, Trubetzkoy proposes that the same contrast should apply to the ambiguous cases, which leads to the conclusion that the minor place splits are phonologically irrelevant. The Greek consonant contrasts can thus be represented as in (10).[9]

(10) Greek: major place, voicing, occlusion > minor place

	Labial	Apical	Sibilant	Dorsal
voiceless stops	p	t	ts	k
voiceless fricatives	f	θ	s	x
voiced fricatives	v	ð	z	ɣ

The criterion employed here by Trubetzkoy can also be viewed in terms of symmetry, or economy. Since the feature [continuant] is required in any case to distinguish between /k/ and /x/, using it also for /p/ ~ /f/ and /t/ ~ /θ/ results in a minimal feature set and a more symmetrical inventory.

In French, however, Trubetzkoy argues for a split labial series: 'For in the entire French consonant system there is not a single phoneme pair in which the relation "spirant : occlusive" would occur in its pure form' (p. 126). Trubetzkoy argues that place should take priority over occlusion in this type of case. Indeed,

9 I substitute phonetic transcription for Trubetzkoy's Greek letters.

he follows this analysis to its logical conclusion that there is no opposition between occlusives and spirants in French, because degree of occlusion cannot be regarded independently of position of articulation (n. 93). As Trubetzkoy does not give a chart, I adapt the one in (11) from Martinet (1964: 65), whose analysis is clearly influenced by Trubetzkoy.

(11) French: minor place, voicing > occlusion

	bilabial	labiodental	apical	alveolar	pre-palatal	dorso-velar
voiceless	p	f	t	s	ʃ	k
voiced	b	v	d	z	ʒ	g

We can express the above analyses formally if Greek and French have different orderings of the occlusion feature, which we can call [continuant], relative to the minor place features that distinguish bilabial from labiodental place:

(12) Variable feature ordering
French: minor place features > [continuant]
Greek: [continuant] > minor place features

Moreover, Trubetzkoy's discussion of these cases suggests a principle that guides the choice of ordering: minor place features take scope over occlusion (German *h*, French labials) *unless* an occlusion contrast is needed as evidenced by neutralization (Czech *ɦ*) or by the principle of parallelism (Greek labials).

3.3.3.2 Vowel systems

Polabian In his discussion of the Polabian (Lechitic West Slavic, extinct) vowel system, Trubetzkoy (1969: 102–3) observes that a 'certain hierarchy existed' whereby the back ~ front contrast is higher than the rounded ~ unrounded one, the latter being a subclassification of the front vowels. Trubetzkoy's analysis suggests that the features are ordered into the (partial) hierarchy: [low] > [back] > [round]; under this analysis, the vowel system is as in (13).[10]

10 Because the point here is to present Trubetzkoy's analysis of Polabian, the vowels in (13) are as presented by Trubetzkoy. See Polański (1993) for a more recent account. According to Polański, the Polabian non-nasal, non-reduced monophthongs are: high vowels /i, ü, u/; 'closed

(13) Polabian vowel system (based on Trubetzkoy 1969)

	[−back]		[+back]
	[−round]	[+round]	
	i	ü	u
	ê	ö	o
[−low]	e		α
[+low]		ɑ	

Trubetzkoy's rationale for this analysis is that, in Polabian, palatalization in consonants is neutralized before all front vowels and before 'the maximally open vowel ɑ which stood outside the classes of timbre' (p. 102). The chart in (13) captures the notion that ɑ 'stood outside the classes of timbre' by ordering [low] before [back]: thus, ɑ has no contrastive value for [back] or [round]. Trubetzkoy cites, as further evidence, the fact that the oppositions between back and front vowels are constant, but those between rounded and unrounded vowels of the same height are neutralizable after *v* and *j* to the unrounded vowels *i* and ê. Because [back] is ordered ahead of [round], 'The properties of lip participation were phonologically irrelevant for the back vowels'. That is, they have no contrastive value for [round].

Though Trubetzkoy does not say so explicitly, it is clear that the vowels /ü/ and /u/ do not form a bilateral opposition in this analysis of Polabian. Since /u/ is not contrastively round, the two vowels have in common only the feature [high], which they share with /i/. The phonemes /i/ and /ü/, however, do form a bilateral opposition based on the shared features [–back] and [+high].

Triangular vowel systems with an 'indeterminate vowel' In the case of Polabian, Trubetzkoy explicitly refers to a hierarchy of contrasts. In other cases, he makes clear that different bilateral oppositions may be assigned to inventories that look the same. This notion is not compatible with pairwise comparisons, which would yield the same result for such inventories.

Consider first triangular vowel systems that have a central 'indeterminate vowel', as in (14). If we assume that the relevant distinctive features are [low],

é /è/; mid vowels /e, ö, o/; and low vowels /a, å/. The vowel written α in (13) may correspond to Polański's å, which he characterizes as the rounded counterpart of *a*.

[high], [round] and [front], then the fully specified specifications of the vowels in (14) are as in (15).

(14) Triangular system with schwa

i u

ə

e o

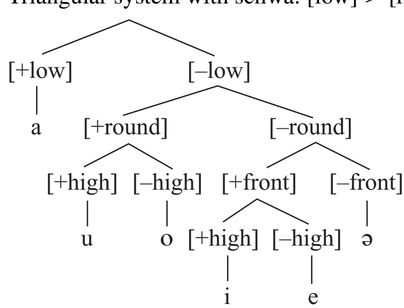
a

(15) Distinctive features of vowels in (14)

	i	e	ə	a	o	u
front	+	+	–	–	–	–
round	–	–	–	–	+	+
low	–	–	–	+	–	–
high	+	–	–	–	–	+

Reading the feature values off the chart in (15), we can see that /a/ and /ə/ are a (potential) minimal pair. However, Trubetzkoy (1969: 113) explicitly states that ‘the “indeterminate vowel” does not stand in a bilateral opposition relation with any other phoneme of the vowel system’ in the usual case, but is ‘characterized only negatively’. We can instantiate this idea in terms of a contrastive hierarchy ordered [low] > [round] > [front] > [high] (among other possibilities), as in (16).

(16) Triangular system with schwa: [low] > [round] > [front] > [high]

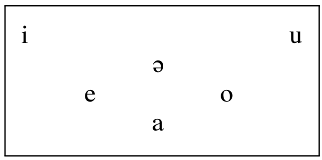


In the contrastive specifications that result from the hierarchy in (16), /a/ is specified only as [+low], and so shares no other feature(s) exclusively with /ə/. Therefore, there is no bilateral opposition between /a/ and /ə/.

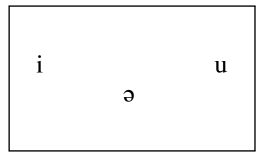
Though Trubetzkoy considers the above to be the usual analysis for vowel systems like (14), he also allows for other analyses. Because the indeterminate /ə/ vowel is ‘outside the system of timbre’, it can enter into a relation with the maximally sonorous vowel in a triangular system, which is also outside the timbre system (1969: 114). An example of such a system is Bulgarian (17).

(17) Bulgarian (Trubetzkoy 1969)

a. Stressed syllables

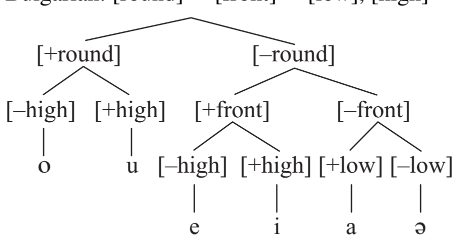


b. Unstressed syllables



An opposition is *proportional* if the relation between its members is identical to the relation between the members of another opposition. Otherwise, it is *isolated*. With respect to Bulgarian, Trubetzkoy writes (1969: 114), ‘It would hardly be possible to assume a pure opposition of timbre between Bulgarian ə and *o*, or between ə and *e*. But the proportions *o*:*a* = *u*:ə, *e*:*a* = *i*:ə, and the proportion *u*:*o* = *i*:*e* = ə:*a* deduced therefrom may well be established.’ We can interpret these remarks in terms of contrastive divisions by ordering [round] and [front] before any height contrasts. In this way we divide the vowel space into three vertical sets, as in (18).[11] Trubetzkoy’s evidence for this analysis is the pattern of neutralization in unstressed syllables (17b), where /u/ and /o/ neutralize to /u/, /i/ and /e/ neutralize to /i/, and /ə/ and /a/ neutralize to /ə/.

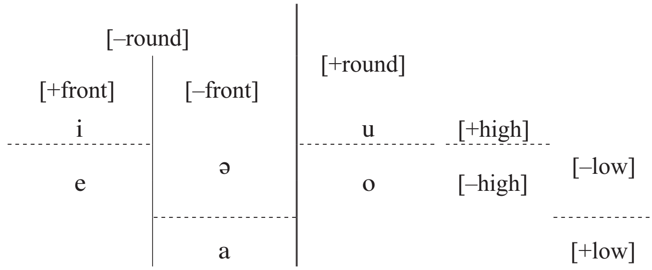
(18) Bulgarian: [round] > [front] > [low], [high]



This way of cutting up the vowel space can be schematically represented as in (19). It may be the ordering of all place features before height that makes this a ‘rare case’ for Trubetzkoy.

11 Alternatively, a single height feature [high] or [low] can be employed if we interpret height features in a more relative manner than is the usual practice.

(19) Bulgarian: [round] > [front] > [low], [high]



Five-vowel systems The last set of cases I will discuss here is Trubetzkoy's analysis of five-vowel systems. He assigns such systems a variety of contrastive relations in a way that can be modelled using a contrastive hierarchy, but not by the pairwise method.

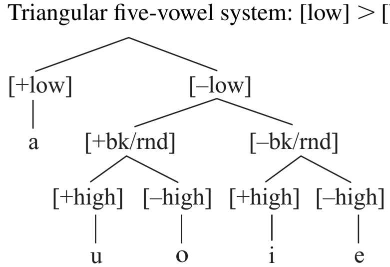
In a typical five-vowel system of the form /i, e, a, o, u/, the non-low vowels are opposed in both place and rounding. Although the low vowel in such systems is (typically) phonetically [+back] and [–round], it is not contrastively so, according to Trubetzkoy. That is to say, the low vowel is not in the contrastive scope of these features, even though they could in principle be relevant to it. An example is Latin (20).

(20) Latin

i u
e o
a

In the contrastive hierarchy of a language like Latin (21), therefore, the feature [low] must take wider scope than the other features.

(21) Triangular five-vowel system: [low] > [back/round], [high]



Though the above analysis is the typical case, Trubetzkoy (1969: 100) also finds some triangular systems that have different contrastive structures, which he deduces from the distribution of the allophones or ‘from the circumstances surrounding the neutralization of the various oppositions’.

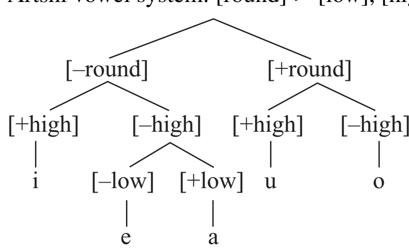
In Artshi, a language of Central Daghestan, certain consonants are divided into a rounded and unrounded variety. This contrast is neutralized before and after the rounded vowels /u/ and /o/: ‘As a result, these vowels are placed in opposition with the remaining vowels of the Artshi system, namely, with unrounded *a*, *e*, and *i*. This means that all vowels are divided into rounded and unrounded vowels, while the back or front position of the tongue proves irrelevant’ (Trubetzkoy 1969: 100–1). He finds further support for this view in the fact that *u*, *o* and *a* are fronted in specific environments.

This analysis of the Artshi vowels can be represented as in (22). The low vowel is contrastively specified for [round], a result that follows from ordering [round] over [low] in the contrastive hierarchy for Artshi, as shown in (23).

(22) Artshi (East Caucasian)

[–round]	[+round]
i	u
e	o
a	

(23) Artshi vowel system: [round] > [low], [high]



Trubetzkoy argues that neutralization of the opposition between palatalized and non-palatalized consonants before *i* and *e* in Japanese shows that these vowels are put into opposition with the other vowels /a, o, u/, and that the governing opposition is that between front and back vowels, lip rounding being irrelevant. Because of the phonological activity of the front vowels, they are the marked ones.

(24) Japanese

[+front]	[–front]
i	u
e	o
	a

Again the low vowel is included in the scope of the front ∼ back contrast, which would here be modelled by placing the feature [front] (or [coronal]) at the top of the contrastive hierarchy.[12]

In these examples there appears to be no alternative to a hierarchical analysis: the pairwise method would yield the same results for all the five-vowel systems discussed above.

3.3.4 Summary

I have argued that Trubetzkoy, despite his many contributions to our understanding of how contrasts work in a phonological system, did not explicitly work out a procedure for determining which features of a phoneme are contrastive and which are redundant. When we try to deduce what he had in mind from the particular analyses presented in the *Grundzüge*, we find that his results sometimes appear to presuppose a procedure involving pairwise comparisons of phonemes. We have seen that this method is not an adequate way to determine contrastive features. It is interesting that Trubetzkoy's most obvious applications of this method occur early in the book when he is discussing oppositions in the abstract. In these cases he brings forward no empirical evidence that the oppositions are in fact the way he proposes. Hence, I conclude that his analyses of these cases, such as French *n*, are incorrect.

Where he wishes to account for actual phonological patterning, however, Trubetzkoy's analyses usually imply a feature ordering approach to determining contrastive specifications. Indeed, in hindsight, one could see Trubetzkoy's work as the beginning of an effort to develop criteria governing the formulation of contrastive hierarchies for particular languages. However, phonological theory did not develop in this way.

12 See Hirayama (2003) for an analysis of Japanese vowels. Her proposal agrees with that of Trubetzkoy's to the extent that the vowels /i/ and /e/ must be specified for [+front] ([coronal], in her analysis). While lip rounding ([peripheral], in her analysis) is not contrastive in the underlying representation of /u/, she proposes that it is specified for /o/ (in contrast to /a/). She proposes further that postlexical processes require the specification of additional features.

Trubetzkoy's failure to arrive at a consistent point of view concerning how to determine contrasts was to be repeated many times in the history of phonology. Particular analyses of Trubetzkoy were discussed and debated in subsequent years – his analysis of French was to give rise to a recurring debate – but little more was said about the criteria he proposed, or about the hierarchies they imply. In fact, the principles governing the selection of relevant contrasts became more obscure in subsequent work, as we shall see in the following sections. Thus, despite its drawbacks, Trubetzkoy's work on contrast attained a level of insight that remained unequalled in the phonological literature.

In the rest of this chapter I will consider three structuralist analyses of the French consonant system: Martinet (1964), Jakobson and Lotz (1949) and Hockett (1955). Each takes a different position on what the relevant contrasts are; each is also crucially incomplete as a theory of contrast. However, we can see in these works the central role that contrast played in phonological theory.

3.4 Martinet: French contrasts based on place

Martinet's *Éléments de linguistique générale,* first published in 1960 and translated into English by Elisabeth Palmer (Martinet 1964), follows in the Prague School tradition of phonological analysis, which gives a central role to contrast:

> The aim of phonological analysis is to identify the phonic elements of a language and to classify them according to their function in that language. Their function is distinctive or oppositional when they contribute to the identification, at one point of the spoken chain, of one sign as opposed to all the other signs which could have figured at that point if the message had been a different one. (Martinet 1964: 53)

Martinet begins his analysis of the French consonants by looking at all the consonants 'which appear or may appear before -*ouche*' (1964: 64).[13] Grouping together segments characterized by a relevant feature, Martinet arrives at the sets in (25). He puts the names of the features in quotation marks to emphasize

13 The idea is to ground the contrastive set by an objective criterion, by limiting the comparison set to segments that can occur in this one context. However, the contextual criterion is loosened considerably when it includes not only those consonants which actually appear in a given context but also those that 'may appear', since no further criteria are given as to what may appear in a given context.

that these are not intended as exhaustive phonetic descriptions, but rather as phonological contrastive categories.[14]

(25) Contrastive sets of French consonants (Martinet 1964)

'unvoiced'	/p, f, t, s, ʃ, k/	'voiced'	/b, v, d, z, ʒ, g/
'non-nasal'	/b, d, j/	'nasal'	/m, n, ɲ/
'lateral'	/l/	'uvular'	/r/
'bilabial'	/p, b, m/	'labio-dental'	/f, v/
'apical'	/t, d, n/		
'hiss'	/s, z/	'hush'	/ʃ, ʒ/
'palatal'	/j, ɲ/	'dorso-velar'	/k, g/

As we observed in section 2.2 looking only at the nasal consonants, these specifications are consistent with contrastive specifications derived by pairwise comparisons. The 'voiced' consonants do not include any sonorants, which are also phonetically voiced, but only obstruents that participate in a minimal pair with an 'unvoiced' phoneme. Similarly, 'non-nasal' phonemes are only those that have a minimally distinct 'nasal' partner.

The other features can be considered as representing nine values of a single place feature. These include values that are not usually thought of as place features, such as 'lateral' as well as 'hiss' and 'hush'. However, this interpretation would account for why there is a 'lateral' category but no 'non-lateral' category. Since Martinet does not suggest that there are any internal groupings among these features (unlike Trubetzkoy, who views bilabial and labiodental as having a special relationship, for example), we will simply treat all these values on a par.

The choice of 'uvular' is unexpected given Martinet's (1964: 54) assertion that /r/ does not always have a uvular pronunciation in French. According to Trubetzkoy's criteria, variation in place of /r/ indicates that [uvular] should not be the defining characteristic of the phoneme.

Though he does not refer to Trubetzkoy's discussion of French consonants, Martinet is clearly following Trubetzkoy in distinguishing stops and fricatives by place of articulation rather than by occlusion. Thus, occlusion is redundant in Martinet's analysis, though he gives no argument for choosing this approach over one that makes it the relevant contrast, an alternative he does not consider.

14 'Hiss' is the translator's rendering of Martinet's term 'sifflant', and 'hush' translates 'chuintant'. I use current IPA symbols in place of some of Martinet's.

Martinet represents some of these features and phonemes in tabular form, shown in (26).

(26) French consonants (Martinet 1964: 65)

	'bilabial'	'labiodental'	'apical'	'hiss'	'hush'	'palatal'	'dorso-velar'
'voiceless'	p	f	t	s	ʃ		k
'voiced'	b	v	d	z	ʒ		g
'nasal'	m		n			ɲ	
						j	

3.5 Jakobson and Lotz (1949): French contrasts based on continuousness

A paper on the Standard French phonemic pattern by Jakobson and Lotz was published in a volume in honour of Henri Muller. The phonemes of Standard French are analysed into six features. Phonemes are assigned one of the following values for each feature: +, if a phoneme has the feature contrastively; −, if a phoneme contrastively lacks a feature; ±, if a phoneme has a contrastive intermediate value of the feature; or nothing, if a phoneme lacks a contrastive value for the feature.

In (27) I present a chart of the specifications proposed by Jakobson and Lotz; # is a 'special value' of [vocality] that Jakobson and Lotz reserve for the 'zero phoneme' /ə/.

(27) Standard French specifications (Jakobson and Lotz 1949)

	d	t	z	s	b	p	v	f
Vocality	−	−	−	−	−	−	−	−
Nasality	−	−	−	−	−	−	−	−
Saturation	−	−	−	−	−	−	−	−
Gravity	−	−	−	−	+	+	+	+
Tensity	−	+	−	+	−	+	−	+
Continuousness	−	−	+	+	−	−	+	+

	g	k	ʒ	ʃ	n	m	ɲ	r	l
Vocality	–	–	–	–	–	–	–	±	±
Nasality	–	–	–	–	+	+	+		
Saturation	+	+	+	+	–	–	+		
Gravity					–	+			
Tensity	–	+	–	+					
Continuousness	–	–	+	+				–	+

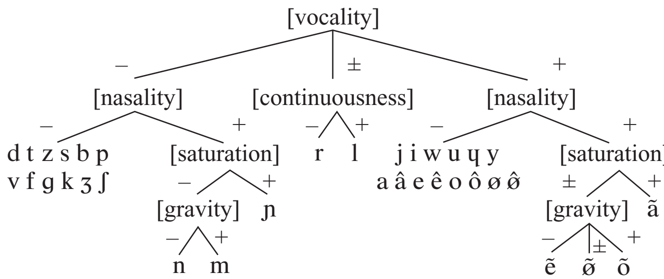
	j	i	w	u	ɥ	y	a	â	e	ê
Vocality	+	+	+	+	+	+	+	+	+	+
Nasality	–	–	–	–	–	–	–	–	–	–
Saturation	–	–	–	–	–	–	+	+	±	±
Gravity	–	–	+	+	±	±			–	–
Tensity	–	+	–	+	–	+	–	+	–	+
Continuousness										

	o	ô	ø	ø̂	ã	ẽ	õ	ø̂	ə
Vocality	+	+	+	+	+	+	+	+	#
Nasality	–	–	–	–	+	+	+	+	
Saturation	±	±	±	±	+	±	±	±	
Gravity	+	+	±	±		–	+	±	
Tensity	–	+	–	+					
Continuousness									

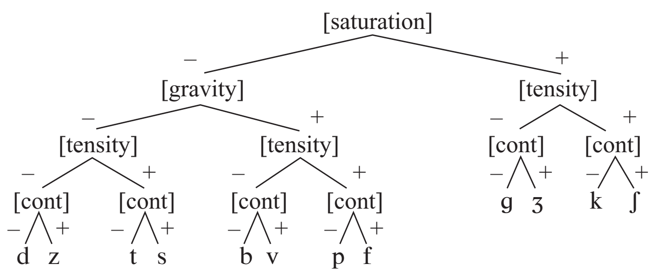
Jakobson and Lotz do not discuss the method whereby they arrive at these specifications. Though they do not explicitly refer to feature ordering, it is clear that the specifications in (27) follow from a hierarchical approach to contrastive specification in which the features are ordered as in the chart. That is, the specifications can be converted into the tree in (28).

(28) Contrastive hierarchy for Standard French

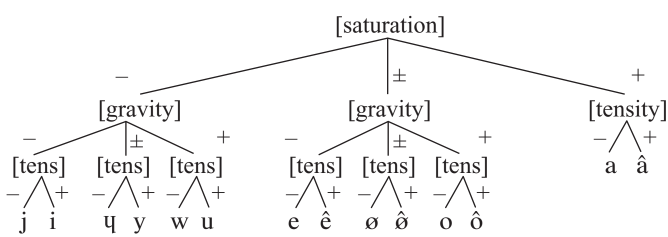
a. Top of the hierarchy: [vocality] > [nasality]



b. Consonants [–vocality, –nasality]



c. Vowels [+vocality, –nasality]



The subtree in (28a) shows the expansion of the top two features, [vocality] and [nasality]. Every phoneme receives a value for [vocality]: – for consonants, ± for liquids, + for vowels and glides, and the special value # for the ‘zero phoneme’ /ə/ (not depicted in this tree). As /ə/ is now uniquely specified, it receives no other features. The liquids need only be further distinguished from each other, in this feature system by [continuousness]. All the other phonemes participate in a nasal/oral contrast. The nasal phonemes are completed as shown in (28a).

Subtree (28b) depicts the contrasts among the consonants specified [–vocality, –nasality] and (28c) shows the remaining vowels specified [+vocality, –nasality].

Whereas Martinet (1964) stipulates that place is the main contrastive dimension for French consonants, with occlusion playing no contrastive role at all, Jakobson and Lotz make occlusion the main contrastive dimension, and collapse the place distinctions into four contrastive places, demarcated by the features [saturation] and [gravity]. They propose that saturated phonemes have longer duration, higher perceptibility and greater resistance to distortion than non-saturated (= diluted) phonemes; saturated consonants (palatals and velars) have a widened front resonator and reduced volume of the back resonator in comparison to diluted consonants (labials and dentals). The distinction between grave and non-grave (= acute) consonants is relevant only among consonants that are specified [–saturation], and opposes labials, with a predominant lower formant, to dentals, with a predominant upper formant.[15]

Jakobson and Lotz present some empirical evidence in favour of their analysis, based on the adaptation of foreign sounds, as well as on language-internal alternations. They observe:

> the difference between velar and palatal is irrelevant in French phonemics . . . These contextual variations do not hinder French speakers from rendering the English velar *ŋ* through the French palatal *ɲ* . . . or the German ‘ich-Laut’ through *ʃ*. The advanced articulation of *k g* before *j* or *i*, as well as the existence of *ŋ* instead of *ɲ* before *w* . . . illustrates the unity of saturated consonants in French. (1949: 153)

Jakobson and Halle (1956) return to the problem of the structure of the French obstruent system, updating the Jakobson and Lotz feature system, but keeping to the same basic analysis. Instead of saturation they use diffuse/compact. A chart based on their proposal is given in (29).

15 The limitation of the grave/acute distinction to diluted consonants is inconsistent with other definitions of gravity, where palatals, which are saturated, are acute, not grave; compare the analysis of Serbo-Croatian by Jakobson (1949) presented in the next chapter.

(29) French consonants, based on Jakobson and Halle (1956)

	diffuse				compact	
	grave		acute			
oral	tns	lax	tns	lax	tns	lax
discontinuous	p	b	t	d	k	g
continuant	f	v	s	z	ʃ	ʒ
nasal	m		n		ɲ	

Jakobson and Halle also provide arguments for choosing their analysis over one which bases the contrasts on point of articulation. They argue (1956: 46) that this solution is 'the unique solution' on the grounds that it is optimal in terms of the number of binary decisions that have to be made. They state that if point of articulation rather than (dis)continuousness were distinctive, 'then the six French voiceless consonants . . . would require, for their identification, fifteen distinctions instead of three, according to the elementary mathematical formula cited by Twaddell (1935)'.[16] Second, the narrower differences in point of articulation are 'minute' and 'hardly recognizable' by themselves. Third, they find that the distinctions between /s/ ~ /f/ and /t/ ~ /p/ involve the same contrast, as do /k/ ~ /t/ and /ʃ/ ~ /s/.[17]

3.6 Hockett: French contrasts and the 'odor of pure game-playing'

C. F. Hockett's *A manual of phonology* (1955) is an outstanding example of late Bloomfieldian American structuralism. This theory is characterized by an unyielding empiricism with respect to both science and psychology. In terms of science, Hockett (1955: 2) writes at the outset, 'it would be well to state explicitly that our view will be empiric: I accept Bloomfield's assertion that "the only useful generalizations about language are inductive generalizations".' With respect to psychology, Hockett inherited Bloomfield's extreme antimentalism that rules out any role for the mind in a scientific description. Together,

16 This calculation assumes the worst case, that the narrower points of articulation are not generated by binary features along the lines used in their own solution.

17 Note that, whereas Jakobson and Lotz's (1949) arguments primarily involve phonological patterning and phonological activity, Jakobson and Halle (1956) focus less on activity and more on arguments of economy, part of a shift in emphasis from phonological activity to economy that will be discussed in the next chapter.

this scientific empiricism and psychological behaviourism keep any theorizing about the grammar confined to rather narrow limits.

3.6.1 Hockett's approach in principle

A good example of this theoretical stance is its application to the familiar problem of the French obstruent inventory. Hockett observes (1955: 173) that it is possible to regard each consonant phoneme as 'a bundle of three coequal ultimate constituents: a voicing-term (voiceless or voiced), an occlusion term (stop or spirant), and one of three positions (say front, central, and back)'. This is essentially the analysis of Jakobson and Lotz (1949), shown in (30a). Hockett also considers the analysis of Trubetzkoy and Martinet, shown in (30b).

(30) Two decompositions of French obstruents (Hockett 1955: 173)

a. Major places only, stop vs spirant, voiceless vs voiced

		'front'	'central'	'back'
'stop'	'voiceless'	p	t	ʃ
	'voiced'	b	d	ʒ
'spirant'	'voiceless'	f	s	k
	'voiced'	v	z	g

b. Major and minor places, voiceless vs voiced

	'bilabial'	'labiodental'	'apico-dental'	'apico-dent. rill'	'lamino-alveolar'	'dorso-velar'
'voiceless'	p	f	t	s	ʃ	k
'voiced'	b	v	d	z	ʒ	g

After considering the merits of each approach, Hockett states his own view (1955: 173) that 'Both of these decompositions of the French obstruents have the odor of pure game-playing, an odor which is seemingly appetizing to some linguists'. He argues that if our sole guide in decomposition is retaining predictability of the omitted features, then far more drastic solutions are available. He notes that any system of sixteen phonemes can be assigned values of four 'determining' (i.e., contrastive) features, as in (31). All other features are then 'determined' (i.e., redundant).

(31) Assigning 'determining' features to a set of sixteen phonemes

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
A	x	x	x	x	x	x	x	x								
B	x	x	x	x					x	x	x	x				
C	x	x			x	x			x	x			x	x		
D	x		x		x		x		x		x		x		x	

The chart in (31), which represents an arbitrary binary coding, is for Hockett (1955: 174) 'a psychologic reductio ad absurdum . . . The opposite of this sort of game-playing is what I mean by "hugging the phonetic ground closely".'[18]

Of course, the chart in (31) assumes that the *only* criterion for assigning contrastive features is minimality of feature specifications, and that there are no testable empirical consequences of any particular feature assignment. Neither of these assumptions is necessary. This train of reasoning, however, leads Hockett to reject the possibility of making any distinctions between contrastive and redundant features:

> Furthermore, it turns out that in general we cannot divide the ostensible ultimate phonologic constituents of a system neatly into 'determining' and 'determined', assigning the latter some sort of secondary status. In the actual complexity of speech, a given feature or difference turns up in some contexts as of primary relevance, in other contexts as subsidiary . . . Thus, for French obstruents, we have no choice but to recognize (1) two voicing terms; (2) two occlusion terms; (3) six combinations of articulator, point of articulation, and contour of articulator – ten features in all. (1955: 174–5)

The fact that this system could lead one to expect twice as many phonemes as there are – for example, a bilabial spirant or a labiodental stop – 'is simply a limitation on privilege of occurrence'.

Against Hockett's arguments in this section, one can observe first that it is incorrect to reduce any example of distinguishing between a contrastive and redundant feature to an arbitrary binary coding. One could have empirical, non-arbitrary reasons for deciding that some features are contrastive. Moreover, contrastive features need not be the mathematically minimal set. Second, Hockett adduces no empirical evidence in favour of his own analysis of the French obstruents. Thus, we have no reason to think it is superior to either of the two analyses he rejects.

18 This chart, and the related discussion, may have been inspired by Cherry, Halle and Jakobson (1953); see further section 4.4.

3.6.2 Hockett's approach in practice

Third, and most important, Hockett's conclusion is not consistent with his own practice in the rest of the *Manual*. If we can indeed make no distinctions between 'determining' and 'determined' features, it would be difficult to assign phonemic symbols to a set of allophones, let alone arrange them into neat schematic diagrams. But this Hockett does consistently in his presentation of types of vowel and consonant systems.

For example, he observes (1955: 84) that a 2×2 type of vowel system is widespread. He portrays such a system with the diagram in (32).

(32) A 2×2 vowel system (Hockett 1955)

i	o
e	a

As examples, Hockett cites Rutul (Caucasian), in which the high back vowel is sometimes rounded, sometimes not, depending on environment; Fox and Shawnee (Algonquian), where the low back vowel is usually unrounded, though rounded in certain environments; and a number of other languages. He adds that 'we class Fox as a two-by-two system despite the fact that the vowel classed as low back, /a/, is typically lower than that classed as low front, /e/'. Though he lists no features, thus leaving open whether the relevant contrast is one of roundness, backness or both, the arrangement in (32) can only mean that these dimensions, as well as a single height contrast, are the relevant (determining) ones. In particular, it is not relevant that /o/ may be phonetically lower than /i/, and /a/ lower than /e/; indeed, the choice of these symbols suggests that /o/ and /e/ might be at the same height phonetically, though functioning phonemically at different heights, whereas /i, o/ and /e, a/ show the reverse. If it were not so, we would have to diagram the vowel system as in (33).

(33) A different four-vowel system

i		
e		o
	a	

Thus, the schematization in (32) does not 'hug the phonetic ground' as closely as it might; on the contrary, it appears to be specifically chosen to show how the contrastive structure of a vowel system can *differ* from its surface phonetic appearance.

Hockett (1955: 84) admits that in his survey he may have made some 'arbitrary' decisions. Thus, he observes that he has assigned three vowel heights rather than two to systems like /i, y, u, e, a, o/ and /i, ɨ , u, e, a, o/ (34), because 'the /a/ in any such case is typically lower than the /e o/'.

(34) Six-vowel systems: three heights

a. Sixth vowel /y/

i	y		u
e			o
		a	

b. Sixth vowel /ɨ/

i	ɨ	u
e		o
	a	

He goes on, 'Yet such a minor difference in height is not always decisive', as, for example, in Fox. Though Hockett would no doubt deny it, it appears clear that his decisions here are not arbitrary (though they may be incorrect), but are based on his understanding of how these systems function.

Hockett makes decisions like these throughout his survey of vowel and consonant systems. To take one more example involving vowels, he writes (1955: 84–5) that a 3+1 system 'is reported for Amahuaca' (35a), 'though the /ɨ/ may be lower than /i, u/, placing Amahuaca rather with Ilocano and others' (35b). He observes that in the Filipino (Austronesian) languages represented by (35b), /ə/ has fronted variants, and also higher central or back unrounded variants.

(35) Vowel systems: 3+1 vs 2+1+1 (Hockett 1955)

a. Amahuaca

i	ɨ	u
	a	

b. Ilocano

i		u
	ə	
	a	

It is not important, for the purposes of this discussion, whether Amahuaca (a Panoan language of Peru and Brazil) is as in (35a) or (35b). What is important is that Hockett believes it is meaningful to assign it to one or the other. If there is indeed no way to distinguish between determined and determining features, we could not represent Ilocano as in (35b), since this diagram implies that the determining features of /ə/, for example, are that it is central and mid, even though it has variants that are front and others that are high. Similarly, Amahuaca could not be represented as in (35a) if /ɨ/ is phonetically lower than /i, u/ to any extent, because that means making a decision that its centrality is the determining feature and its lower height is the determined feature.

Hockett's discussion of consonant systems is also at odds with his stated theoretical position. He begins his discussion of classification of consonants by observing, 'It does not seem feasible to handle them as wholes for constitutional classification; it seems better to develop some manner of breaking them up into subsystems' (1955: 95). The breaking up into subsystems begins with a binary split, typically dividing the obstruents from the sonorants. To illustrate, he presents the table of Ossetic (Iranian) shown in (36).[19]

(36) Consonant system of Ossetic (Hockett 1955)

p		t			c	tʃ	k	
p^h		t^h			c^h	$tʃ^h$	k^h	
b		d			ɟ	dʒ	g	ʔ
f					s	ʃ	x	
v					z	ʒ	ɣ	
m		n						
w	j		r	l				

As to what counts as an obstruent, Hockett writes, 'We include among the obstruents not only all stops and affricates and most spirants, but also, in some cases, a [y]-like, [w]-like, or [l]-like consonant *if it fits neatly into the scheme*, and if distributional facts do not militate against such a treatment' (1955: 96, italics added).

Though the main split in the consonant system is often obstruent ~ sonorant, 'In a few cases all the consonants, apart from manner consonants (if any), must be otherwise dichotomized in the first instance, the classification into obstruent and sonorant coming second'. He illustrates with Lifu (Malayo-Polynesian), in which the voicing contrast 'is operative throughout the system except for /h/; the latter is a manner consonant and pairs off with the whole set of voiceless obstruents and sonorants'.

(37) Consonant system of Lifu (Hockett 1955: 96)

p	t		ʈ	tʃ	k	
f	θ	s		ʃ	x	h
m̥	n̥			ɲ̊	ŋ̊	
w̥	l̥					
b	d		ɖ	dʒ	g	
v	ð	z		ʒ	ɣ	
m	n			ɲ	ŋ	
w	l					

19 I have replaced some of Hockett's symbols with their IPA counterparts in this and the following chart.

Hockett calls the laryngeals in Ossetic and Lifu ‘manner consonants’ because ‘they match one or another of the styles of delivery found for obstruents’. He also considers Russian /y/ (= IPA /j/) a manner consonant, because it ‘matches the whole set of palatal consonants as over against the plain consonants’.

Hockett’s analysis translates easily into a contrastive hierarchy. Indeed, it is exactly a hierarchy of successive binary splits, but not carried out all the way or entirely explicit. Ossetic first splits into obstruent ~ sonorant sets. Then /ʔ/ splits from the rest of the obstruents. We could interpret Hockett’s designation of ‘manner consonant’ to mean that /ʔ/ is placeless, being characterized only by the fact it is an obstruent. The rest of the obstruents are then split into place and manner categories. The diagram suggests that the coronal sonorants do not have a definite place. In Lifu, the first division is voiceless ~ voiced. Then /h/ splits away from the voiceless set, and can be interpreted along the lines of Ossetic /ʔ/.

Hockett’s grounds for adopting these diagrams are distributional. With respect to Lifu, he notes (1955: 96) that ‘Distributional classification supports the constitutional grouping indicated above: only voiced consonants (and all of them except /w ŋ ɡ ɣ v/) occur finally.’ Thus, there are empirical grounds, based on phonological patterning, for organizing segment inventories one way rather than another. We see, then, that Hockett’s practice throughout the *Manual* is inconsistent with his discussion of the French obstruents.

One could argue that Hockett viewed his diagrams simply as ways of expressing distributional or other generalizations about inventories, and not as expressing any ‘correct’ analysis of the grammar. Ultimately, it is not really important if Hockett thought of his classifications as real or as useful fictions. What is important is his practice, which is reminiscent of Trubetzkoy’s in this sense: when he considers the problem of contrast in the abstract, he arrives at conclusions that are not consistent with what he does when he needs to make sense of actual data; in such cases, where there is real empirical evidence for doing things a certain way, he assigns contrastive features in terms of a language-particular contrastive hierarchy.

3.7 Prolegomenon to a theory of contrastive specification

In this chapter we have surveyed the work of some leading phonologists from the structuralist period (Roman Jakobson will be discussed further in the next chapter). I have argued that issues of phonological contrast were central to their thinking. Matters are somewhat more obscure when we try to isolate an explicit or consistent approach to assigning contrastive features. However, when we put together their most insightful and empirically supported analyses we can begin to distil the main principles of a theory of contrastive specification.

Of the various features that characterize a phoneme, we must distinguish between those that are contrastive and those that are redundant. We do so by ordering the features into a contrastive hierarchy, and assigning features to phonemes in order until each phoneme has been uniquely distinguished from every other one. The contrastive features make up the 'phonemic content' of a phoneme. There are in principle different ways of ordering the features for a given set of phonemes, and each ordering corresponds to a particular set of contrastive specifications. In this sense, the contrastive structure of a language is a function of 'point of view', that is, of a particular way of ordering the features.

This variability gives rise to a fundamental question: how do we know what the particular ordering is in any given case? The works surveyed in this chapter suggest a general answer to this question. We can recognize the phonemic content of a phoneme by its 'patterning', that is, by the way it behaves. A common assumption of the authors surveyed above is that the behaviour of a phoneme is a function of its contrastive features. Reviewing the cases discussed above, we can compile a list (38) of diagnostics used in this chapter for identifying contrastive features.

(38) Diagnostics used in identifying contrastive features
A phoneme φ has contrastive feature F if:

a. φ enters into an alternation or neutralization that is best explained if F is part of φ (cf. Sapir (1); Trubetzkoy, Czech /ɦ/, Polabian front vowels (13), Bulgarian vowels (19)).
b. φ causes other phonemes to alternate or neutralize in a way that is best explained if F is part of φ (Trubetzkoy, Polabian (13), Artshi round vowels (22), Japanese front vowels (24)).
c. φ participates in a series with other phonemes, Φ, with respect to phonotactic distribution, where F is required to characterize Φ in a general way (Sapir (1); Trubetzkoy, Greek (10)).
d. the set of allophones which make up φ all have F in common (Trubetzkoy, German /x/ and German, Czech and Gilyak /r/; Hockett, Rutul, Fox, Shawnee back vowels (32)).
e. speakers adapt a sound from another language in a way that can be explained by supposing that they assign F to the foreign sound (Jakobson and Lotz, English /ŋ/ adapted as French /ɲ/; also Jakobson (1962b [1931]), Slovak, Russian and Czech ability to pronounce foreign front rounded vowels).

To the extent they are not present or contradicted, some of these diagnostics can be used conversely to identify system-redundant features (39).

(39) Diagnostics used in identifying system-redundant features
Feature F is system redundant (noncontrastive) in phoneme φ if the set of allophones which make up φ do not have F in common:
a. Trubetzkoy: dorsal and palatal features not contrastive in German /x/.
b. Trubetzkoy: place of articulation not contrastive in German /r/.
c. Hockett: [round] not contrastive in Rutul, Fox, Shawnee back vowels (32).

We have also seen some arguments for ordering features in a given system (40).

(40) Diagnostics for ordering features
If a phoneme φ has two features, F and G:
a. F is ordered above G if F is contrastive in φ and G is system redundant in φ (based on the general assumption in all authors above that phonologically irrelevant features do not participate in phonemic content).
b. if F and G are both contrastive, F is ordered above G if G is neutralizable and F is not (Trubetzkoy, Polabian (13)).
c. F is ordered above G if F applies to a wider range of phonemes than G (Hockett, Ossetic [sonorant] > [voiced] (36), Lifu [voiced] > [sonorant] (37)).

The above make up a preliminary set of principles that may contribute to a theory of contrastive specification. It is not clear that all of the above diagnostics are valid, or valid in all circumstances, and we will try to refine these principles in later chapters. Much has also been left open. To take one important issue, we have not attempted to clarify the relationship between contrast and specification. A natural assumption is that contrastive features are specified in a phoneme and redundant features are unspecified; but this is not a necessary position to take. As Stephen Anderson (1985) and Calabrese (1995) have emphasized, distinguishing between contrastive and redundant features does not necessarily imply that the latter must be absent from representations. In a theory where all features are specified, for example, we can still designate some as contrastive. In that kind of theory, the above remarks about phonemic content and related matters still hold, but must be understood not as making a distinction between features that are present and those that are entirely absent, but rather as distinguishing between features that are present as contrastive features and those that are not. However, this and other issues must be deferred to chapters 7 and 8.

To conclude this section, let us first subsume the various diagnostics in (38) under a general term, and look at a sample example. We will say that a feature that exhibits one of the characteristics of (38) is *phonologically* *active*. Conversely, a feature that does not have one of these characteristics is *phonologically inert*. Now the various diagnostics for identifying contrastive features follow from the hypothesis in (41).

(41) Contrast and phonological activity (preliminary hypothesis)
Only contrastive features are active in the phonology. System-redundant features are inert.

D. C. Hall (2007: 20) calls this idea the *Contrastivist Hypothesis*, which he formulates as in (42).

(42) The Contrastivist Hypothesis (D. C. Hall 2007)
The phonological component of a language L operates only on those features which are necessary to distinguish the phonemes of L from one another.

Let us consider again the Artshi example (22) as analysed by Trubetzkoy. The vowels /u/ and /o/ cause the distinction between rounded and unrounded consonants to be neutralized, when they are adjacent to these consonants. Since the affected feature involves rounding, it is reasonable to suppose that the neutralization is caused by the feature [+round] associated with these vowels. That is, the feature [round] is active in these vowels. By hypothesis (41), only contrastive features are active. Therefore, the feature [round] must be contrastive in /u/ and /o/. Since there is no evidence that the feature [back] is contrastive in these vowels, we assume that this feature is system redundant, because no contrastive hierarchy exists that would assign these vowels the contrastive feature [+back] if they are already assigned [+round].

Note a certain benign circularity here: we hypothesize that only contrastive features are active, and then we decide that the feature [round] is contrastive in /u/ and /o/ because it is active in these vowels. This circularity is 'benign' because it is the typical circularity characteristic of scientific explanation. To explain why objects fall to the earth with a certain acceleration, we posit a force of gravity; evidence for this force is the fact that objects fall with a certain acceleration. What is important is that contrast and activity are not defined in terms of each other: activity is not part of the definition of contrast, and contrast is not part of the definition of activity. Activity is not defined in terms of contrast because the various manifestations of activity do not refer to the contrastive status of features. Contrast is not defined in terms of activity because the notions of feature ordering and the SDA do not refer to activity. Moreover, there are situations where we must designate features as contrastive to differentiate between phonemes in the absence of any evidence of activity.

At a practical level, the hypothesis that only contrastive features are active can be easily falsified. If in Artshi, for example, we found that the features [round] and [back] were both active in /u/ and /o/, and that [+low] was active in /a/, this result would not be consistent with the hypothesis in (41). For there is no feature ordering that would make all these features contrastive at the same time. Conversely, this hypothesis is supported to the extent that we find cases where the active features are consistent with orderings that make them contrastive.

To sum up, based on the work surveyed in this chapter, we have sketched the beginnings of a theory of phonology that assigns a central role to contrastive feature specifications. So far, this theory has two main tenets: (1) only contrastive feature specifications are active in the phonology (the Contrastivist Hypothesis), and (2) contrastive features are assigned by ordering the features and applying the SDA.

In the work surveyed above, the Contrastivist Hypothesis was much in evidence in practice, if not in name, but the contrastive hierarchy was not clearly understood. In the next chapter we will see that the contrastive hierarchy was promoted to a leading place in phonological theory by Roman Jakobson and his colleagues; its connection to phonological activity, however, was loosened and eventually lost, along with the Contrastivist Hypothesis.

4 *The rise and fall of the contrastive hierarchy*

4.1 Introduction: Jakobson and his collaborators

The work of Roman Jakobson and his colleagues merits a separate discussion. Many of the main ideas in Trubetzkoy's *Gründzuge* were worked out in collaboration with Jakobson. To Jakobson is due the notion of the distinctive feature. Whereas Trubetzkoy allowed features to be gradual (multivalued), Jakobson eventually proposed that all features are binary. The notion that language and cognition crucially involve binary dichotomies became central to Jakobson's thinking, and this idea was later taken over into generative phonology.

In the 1950s and early 1960s, Roman Jakobson, Morris Halle and their colleagues wrote a series of publications that laid the groundwork for the next phase of distinctive feature theory and what eventually became the theory of generative phonology. Much critical attention, then and afterwards, was directed at various aspects of their theory that struck observers as being the most controversial and innovative, such as the nature of distinctive features, the relation of phonemes to allophones, and the organization of the grammar.

What is of greatest interest for our present purposes is an aspect of their work that attracted relatively less attention. This is their development and utilization of a contrastive hierarchy of distinctive features. We have seen that such a hierarchy is already implicit in much of Trubetzkoy (1939), and it could be that this was one source of the idea. It also fits well with Jakobson's general emphasis on dichotomies as a fundamental aspect of cognition: if the distinctive feature is grounded in a binary discrimination, then it makes sense that an inventory is built up via a succession of such binary splits.

Although Jakobson and his collaborators assumed that feature ordering is the correct way to arrive at contrastive features, the concept remained oddly unfocused. Like Trubetzkoy, Jakobson and Halle continued to oscillate between feature ordering and pairwise comparisons in working out the distinctive features for a language. The notion of feature ordering was never, to my knowledge, the subject of debate and discussion in the way that other aspects of linguistic theory were. When it was used it was without explicit defence or recognition that there were other methods; similarly, when it was not used, its absence was not commented on.

Most important, for our purposes, is that the earlier association between contrastive specification and phonological activity was not developed in Jakobson's work in the 1950s, but was gradually dropped as a prime motivation for assigning contrastive specifications. As the contrastive hierarchy became detached from the Contrastivist Hypothesis, it was deprived of its most important empirical motivation. This left the notion of feature ordering, and of contrastive specification more generally, vulnerable to arguments that it plays no useful role in phonological theory.

In this chapter I will trace the rise of the contrastive hierarchy in the early 1950s to its position as 'the pivotal principle' of linguistic structure, through its gradual dissociation from the Contrastivist Hypothesis, to its ultimate disappearance from phonological theory in the 1960s.

4.2 The dichotomous scale

Jakobson, Fant and Halle propose that the distinctive features into which the phonological systems of all the world's languages can be analysed are drawn from twelve binary oppositions (1).

(1) Universal set of distinctive features (Jakobson, Fant and Halle 1952)

1. vocalic/non-vocalic	7. compact/diffuse
2. consonantal/non-consonantal	8. grave/acute
3. interrupted/continuant	9. flat/plain
4. checked/unchecked	10. sharp/plain
5. strident/mellow	11. tense/lax
6. voiced/voiceless	12. nasal/oral

They propose that listeners identify phonemes by distinguishing them from every other phoneme in the system. These distinctions are effected by making a series of binary choices that correspond to the oppositions active in the language: 'The dichotomous scale is the pivotal principle of the linguistic structure. The code imposes it upon the sound' (1952: 9). There is no clear statement as to whether there is a universal hierarchical ordering of the features and, if so, what it is. But it is possible to infer from their discussion that they believe that there are some universal patterns and some tendencies for features to be ordered in a certain way. They write (1952: 40–1) that there are 'laws of implication' that are 'universally valid or at least have a high statistical probability: X implies the presence of Y and/or the absence of Z'. These laws limit the possible variety of phonological systems to 'a limited set of structural types'.

Jakobson, Fant and Halle (1952: 10) claim that the dichotomous scale of distinctive features, in the context of the patterning of the linguistic code, influences perception of speech sounds:

> Therefore, a monolingual Slovak identifies the rounded front vowel /ø/ of the French word *jeu* as /e/, since the only distinctive opposition in his mother tongue is acute (front) vs. grave (back) and not flat (rounded) vs. plain (unrounded). A monolingual Russian, on the contrary, perceives the same French vowel as /o/ because his native tongue possesses only the one of the two oppositions in question, namely, flat vs. plain.[1]

It is clear that reference here is not to phonetic contrasts, but to the contrastive features that are active in the phonologies of these languages.

4.3 The dichotomous scale as an acquisition sequence

The contrastive hierarchy is featured prominently in Jakobson and Halle's influential *Fundamentals of language* (1956). Jakobson and Halle (1956: 47) refer to this hierarchy as the 'dichotomous scale'. In reply to Chao (1954), who asked if the dichotomous scale is a principle imposed by the analyst or inherent in the structure of language, they adduce 'several weighty arguments' in support of this hierarchical approach to feature specification. First, they claim (1956: 47) that such a system is an 'optimal code' for speech participants who have to encode and decode messages; I will discuss this claim further in section 4.4.

Their second argument involves language acquisition (cf. Jakobson 1941). They suggest that distinctive features are necessarily binary because of the way they are acquired, through a series of 'binary fissions'. They propose (1956: 41) that the order of these contrastive splits is partially fixed, thereby allowing for certain developmental sequences and ruling out others.

1 The rationale behind this analysis is provided in Jakobson (1962b [1931]), discussed in section 1.1.

The sequence in (2), for example, concerns oral resonance (primary and secondary place) features. The decimals indicate precedence relations: if one decimal sequence is entirely contained in another sequence, then the contrast corresponding to the former must precede the acquisition of the latter contrast. In (2), wide (equals compact) vowels are low vowels of high sonority; narrow (equals diffuse) vowels are high vowels of low sonority; palatal vowels are front, and velar vowels are back. Thus, (2) predicts that a height contrast between a low and a high vowel (0.11) must precede the emergence of a contrast between a front and a back vowel (0.111), since 0.11 is contained in 0.111. Further, the latter contrast should emerge in the high vowels (0.111) before it can emerge in the low vowels (0.1111). At this point also, a rounding contrast can arise in the high vowels (0.1112), and this must precede the development of a rounding contrast in the low vowels (0.11121). These precedence relations are meant to account for why /i, a, u/ is a typical vowel inventory, but /i, æ, ɑ/ and /i, y, a/ are not.

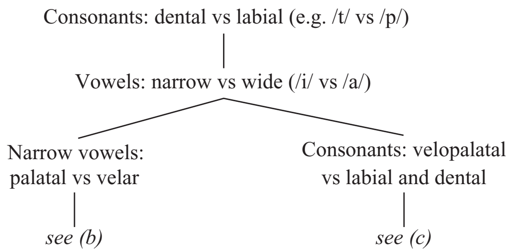
(2) Predicted acquisition sequences (Jakobson and Halle 1956: 41)

Consonants: dental vs labial	..... 0.1
Vowels: narrow vs wide	..... 0.11
Narrow vowels: palatal vs velar	..... 0.111
Wide vowels: palatal vs velar	..... 0.1111
Narrow palatal vowels: rounded vs unrounded	..... 0.1112
Wide palatal vowels: rounded vs unrounded	..... 0.11121
Velar vowels: unrounded vs rounded	..... 0.1113
Consonants: velopalatal vs labial and dental	..... 0.112
Consonants: palatal vs velar	..... 0.1121
Consonants: rounded vs unrounded or pharyngealized vs non-pharyngealized	..... 0.1122
Consonants: palatalized vs non-palatalized	..... 0.1123

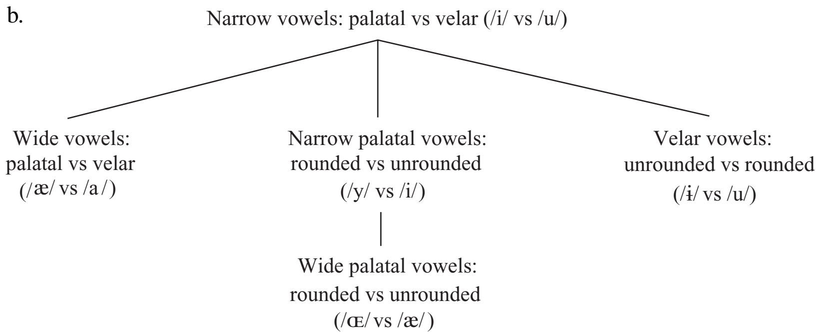
By the same token, (2) predicts that the first contrast to be established in a phonological system is between dental and labial consonants (e.g. /t/ ∼ /p/). The next consonantal contrast adds a velopalatal consonant (e.g. /k/). This contrast (0.112) is predicted to follow the acquisition of contrast 0.11, since 0.11 is contained in 0.112, but no implicational relation is predicted to hold between it and contrast 0.111, since 0.111 is not contained in 0.112. There is no significance to the fact that one decimal is a smaller number than the other; thus, the chart in (2) is equivalent to the branching tree diagram in (3).

(3) Predicted acquisition sequences: oral resonance features (Jakobson and Halle 1956: 41)

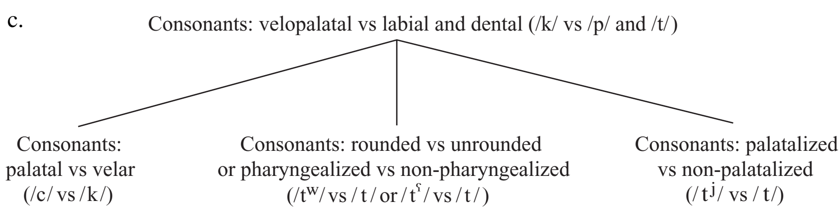
a.



b.



c.



Language acquisition will be discussed further in section 7.8. For now, it is important to observe that the arguments from perception and acquisition discussed above provide empirical evidence in support of the contrastive hierarchy. Over time, however, the emphasis began to shift toward non-empirical considerations of efficient coding and reducing redundancy, and this emphasis eventually led away from the Contrastivist Hypothesis. These developments are discussed in the following sections.

4.4 Feature ordering and efficient coding

Cherry, Halle and Jakobson (1953) consider some properties of phonemic structure in terms of mathematical concepts relevant to the then-emerging field of statistical communication theory. In their article, they explicitly adopt feature ordering: 'For the purpose of identifying one particular phoneme out of the set employed by the language, the distinctive features may be regarded as questions to be answered yes or no.' They make it clear that these questions must be asked in an order, and point out that the logic employed is three-valued: yes (+), no (–), and zero, which they take to mean that either answer can be given.

The authors make some observations about the number of feature specifications required to differentiate any number of phonemes. For example, they observe that a set of eight phonemes can be fully distinguished with maximum economy by three binary features without incurring zeros.[2]

(4) Logical identification of objects in a set of eight

A	B	C	D	E	F	G	H
−	−	−	−	+	+	+	+
−	−	+	+	−	−	+	+
−	+	−	+	−	+	−	+

In contrast to the hypothetical chart in (4), which splits the members of the set into equal groupings so as to produce an optimally efficient coding, Cherry, Halle and Jakobson observe that phonemes in real languages are not so evenly divided. The consequence of this is that more than the logical minimum number of features has to be employed, and each phoneme requires more than the minimum possible number of bits of information. As an example they consider the phonemes of Russian (discussed further in section 4.6 below). In their analysis there are forty-two phonemes which require eleven distinctive features, rather than the six features which make up the logical minimum needed to fully distinguish a set of forty-two members.

2 This diagram and the related discussion may have been the trigger for Hockett's (1955) attack on arbitrary binary coding of phonemes; see section 3.6.1 above. They make clear, however, that the table in (4) is meant only as an illustration of the logic of successive binary choices when dealing with 'objects without linguistic significance'. Their discussion of Russian demonstrates that real phonemes do not typically split up so as to produce an optimally efficient binary coding.

Continuing their earlier attention to the question of the efficiency of the coding obtained by their procedure, Cherry, Halle and Jakobson ask if it is possible to remove all or many of the zero signs in their table by reordering the features. They observe that no simple reordering will permit them to remove all the zeros or push them to the end of each phoneme column. They go on to propose that gains can be achieved if one gives up a fixed feature order, and allows the order of the features to vary in different branches of the decision tree. In their focus on the problem of zero specifications, Cherry, Halle and Jakobson (1953) foreshadow what was soon to become a major concern in early generative phonology.

4.5 Inconsistent approaches to contrastive specification

Though the contrastive hierarchy was well established in the work of Jakobson and his colleagues from the late 1940s on, it was not always strictly employed in working out phoneme specifications. While many publications incorporate it, some others do not, for reasons that are not made explicit. One can speculate that the authors were somehow not satisfied with the results of feature ordering in particular cases. It is possible that they wanted to specify certain phonemes in ways not permitted by strict feature ordering. Another possibility is that they were unable to arrive at a universal feature hierarchy that gave satisfactory results for every language they studied.

In this section I will briefly review some case studies where Jakobson and his collaborators deviate from feature ordering. The existence of these analyses illustrates that the concept of the contrastive hierarchy was not as securely rooted as the comments cited earlier in this chapter might have led one to believe.

4.5.1 Jakobson (1949): Serbo-Croatian

Jakobson's analysis of Serbo-Croatian appeared in the same year as the paper by Jakobson and Lotz on the phonemes of Standard French. We have seen that the specifications in the latter paper are consistent with feature ordering. The same is not the case for Jakobson's analysis of Serbo-Croatian. It can be shown that Jakobson does not use feature ordering. Rather, his results are fairly consistent with what would be given by the pairwise method, once various adjustments and allowance for errors are made.

Jakobson (1949) observes that Standard Serbo-Croatian has twenty-nine phonemes that are analysed into eight 'dichotomous properties'. Six of these are qualitative or inherent features, and two are prosodic features (high tone and length). Jakobson gives the values of each phoneme for the inherent features in the chart in (5).[3] For some features he allows a ternary choice (not counting a lack of specification): + indicates the presence of a feature; – indicates its 'distinctive absence' (i.e., not its absence by virtue of redundancy); and ± indicates a 'complex combining both opposite terms'.

(5) Feature specifications: Serbo-Croatian (Jakobson 1949)

	t	d	c	s	z	p	b	f	v
Vocality									
Nasality		–					–		
Saturation	–	–	–	–	–	–	–	–	
Gravity	–	–		–	–	+	+	+	+
Continuousness	–	–	±	+	+	–	–	+	+
Voicing	–	+		–	+	–	+	–	+

	ć	đ	č	ǧ	š	ž	k	g	x
Vocality									
Nasality		–							
Saturation	+	+	+	+	+	+	+	+	+
Gravity	–	–			–		+	+	+
Continuousness	–	–	±	±	+	+	–		+
Voicing	–	+	–	+	–	+	–	+	

	n	m	ń	r	c	c’	i	u	e	o	a
Vocality				±	±	±	+	+	+	+	+
Nasality	+	+	+								
Saturation	–	–	+		–	+	–	–	±	±	+
Gravity	–	+					–	+	–	+	
Continuousness				–	+	+					
Voicing											

3 I retain here Jakobson's use of the Croatian spelling form for the phonemes.

Simple inspection of the chart shows immediately that no feature ordering yields the given values. Recall that feature ordering requires that some feature be at the top of the hierarchy, and that an initial division based on this feature will result in all phonemes receiving a value for that feature.[4] We observe, however, that there is no row in which all phonemes have a value. Certainly, the ordering of the features cannot reflect their ordering in any contrastive hierarchy, since the feature listed first, vocality, is unspecified for all of the obstruents, and the following feature, nasality, is specified for only six consonants.[5]

A look at the specifications of the oral and nasal stops for [voicing] and [nasality] shows us that Jakobson must have been working from fully specified minimal pairs. The relevant specifications are shown in (6). I have rearranged the rows and columns to bring out the relevant comparisons more clearly.

(6) Specifications of oral and nasal stops

	t	d	n	p	b	m	ć	đ	ń	k	g
Voicing	−	+		−	+		−	+		−	+
Nasality		−	+		−	+		−	+		

In each box in (6) we can see the characteristic pattern of specifications driven by minimal pairs: in the coronals and labials, the voiced stops receive two specifications because they participate in two minimal pairs, whereas the voiceless stops and the nasal each receive one specification because they participate in one minimal pair with the voiced stop. As there is no velar nasal, the velar stops /k, g/ lack specifications for nasality.

Applying the pairwise method to the full set of Serbo-Croatian phonemes in (5) gives the results in (7). Any divergence from Jakobson's assignments is shaded.

4 This is not the case for privative features, but it is clear from Jakobson's remarks that he is not using privative features here, because the 'distinctive absence' of a feature is indicated by minus.

5 It is interesting that the procedure of van den Broecke (1976), an elaborate version of the pairwise method discussed in section 2.5.5, resulted in only a single consonant, /ð/, being specified for [sonorant], which plays a similar role to that assumed by [vocalic] in Jakobson's analysis. This is further evidence that Jakobson must have been using some form of pairwise comparison here.

(7) Serbo-Croatian via pairwise method

	t	d	c	s	z	p	b	f	v
Vocality		–			–				–
Nasality		–					–		
Saturation	–	–	–	–	–	–	–	–	
Gravity	–	–		–	–	+	+	+	+
Continuousness	–	–	±	+	+	–	–	+	+
Voicing	–	+		–	+	–	+	–	+

	ć	đ	č	ǧ	š	ž	k	g	x
Vocality						–			
Nasality		–							
Saturation	+	+	+	0	+	+	+	+	+
Gravity	–	–			–		+	+	+
Continuousness	–	–	±	±	+	+	–		+
Voicing	–	+	–	+	–	+	–	+	

	n	m	ń	r	l	l’	i	u	e	o	a
Vocality				±	±	±	+	+	0	0	0
Nasality	+	+	+								
Saturation	–	0	+		–	+	–	–	±	±	+
Gravity	–	+					–	+	–	+	
Continuousness				–	+	0					
Voicing											

These results come fairly close to Jakobson’s specifications. There are ten discrepancies in all. Some of them may be due to simple mistakes, or typos in Jakobson’s article.

Three discrepancies between my application of the pairwise method and Jakobson’s specifications involve consonants where Jakobson has a value where I derive a 0. These cases are: /m/ and /ǧ/ ([saturation]) and /l’/ ([continuousness]). In these cases, the phoneme in question forms a class with two other phonemes: /m, n, ń/ are the only nasals, /c, č, ǧ/ are the only affricates and /r, l, l'/ are the only liquids. The middle term in each set participates in two minimal pairs; the peripheral phonemes each participate in one (8).

(8) Nasal and liquid contrasts via minimal pairs

a. Nasals

	m	n	ń
Saturation	−	−	+
Gravity	+	−	

b. Affricates

	c	č	ǧ
Saturation	−	+	+
Voicing		−	+

c. Liquids

	r	l	l'
Saturation		−	+
Continuousness	−	+	+

These may be simple mistakes in Jakobson's chart.

The remaining discrepancies have to do with [vocality]. From the point of view of the pairwise method, /d/ and the voiced continuants are underspecified in Jakobson's chart while the vowels are overspecified for this feature.

According to the full feature specifications I have assumed, /d/ and the voiced continuants /z/, /v/ and /ž/ form minimal pairs based on [vocality] with /r/, /l/, /u/ and /l'/, respectively. Therefore, according to the pairwise method, they must be specified [–vocality]. On the other hand, the vowels /e, o/ need not be specified for [vocality]. The mid vowels are unique in being specified [±saturation], so they need only be distinguished from each other by one more feature.

The phoneme /a/ is a different case. It forms minimal pairs only with /u/ and /o/, from which it is distinguished by [saturation]. Its lone specification [+saturation] is indeed insufficient to distinguish it from any other segment that is [+saturation], but it is too far from them to form a minimal pair that can be recognized by the pairwise method. Therefore, the lack of specification of further features in /a/ results from a failure in the mechanical operation of this approach. Perhaps because of this Jakobson departed from strict minimal pair comparisons to guide the specification of [vocality] in the vowels, which may have resulted in the apparent over- and under-specifications.

4.5.2 Halle (1954): Standard Literary German

Though Halle (1954) refers to the ‘dichotomous scale’, he does not appear to utilize feature ordering consistently in the one example for which detailed specifications are given, Standard Literary German, shown in (9).[6] Neither feature ordering nor pairwise comparisons yield exactly these specifications.

(9) Feature specifications: Standard Literary German (Halle 1954)

	m	p	b	f	v	pf	n	t	d	s	z	ts	k	g	x	ʃ
Vocalic vs non-vocalic	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
Consonantal vs non-cons.	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Compact vs diffuse	0	−	−	−	−	−	0	−	−	−	−	−	+	+	+	+
Grave vs acute	+	+	+	+	+	+	−	−	−	−	−	−	+	+	+	−
Flat vs plain	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Nasal vs non-nasal	+	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−
Continuant vs interrupted	0	−	−	+	+	−	0	−	−	+	+	−	−	−	+	0
Strident vs mellow	0	−	0	0	0	+	0	−	0	0	0	+	0	0	0	0
Tense vs lax	0	+	−	+	−	0	0	+	−	+	−	0	+	−	0	0

	r	l	u	o	a	ü	ö	i	e	æ	h
Vocalic vs non-vocalic	+	+	+	+	+	+	+	+	+	+	−
Consonantal vs non-cons.	+	+	−	−	−	−	−	−	−	−	−
Compact vs diffuse	0	0	−	±	+	−	±	−	±	+	0
Grave vs acute	0	0	+	+	+	−	−	−	−	−	0
Flat vs plain	0	0	+	+	0	+	+	−	−	0	0
Nasal vs non-nasal	0	0	0	0	0	0	0	0	0	0	0
Continuant vs interrupted	−	+	0	0	0	0	0	0	0	0	0
Strident vs mellow	0	0	0	0	0	0	0	0	0	0	0
Tense vs lax	0	0	0	0	0	0	0	0	0	0	0

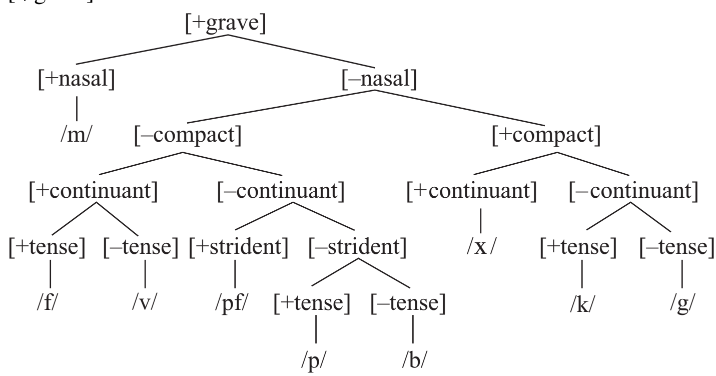
Looking at the consonants, it is possible to derive almost all the specifications via feature ordering if we slightly reorder the features. To come close to the

6 I have replaced Halle’s *f̂* by *pf*, and *ŝ* by *ts*. Halle notes that he has not indicated vowel length contrasts, and that the distinction between /e/ and /æ/ is not phonemic among short vowels.

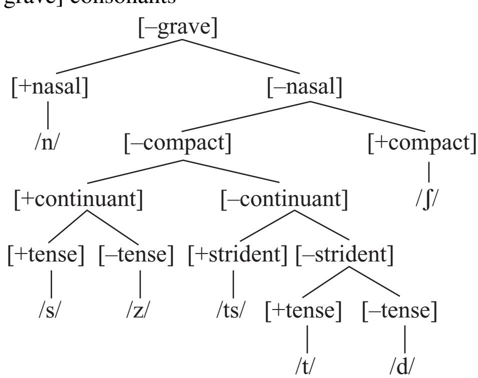
listed specifications we must move [compact] down below [grave] and [nasal] ([flat] is not relevant to consonants here). The order [vocalic] > [consonantal] > [grave] > [nasal] > [compact] > [continuant] > [strident] > [tense] is illustrated in (10).

(10) Contrastive hierarchy: Standard Literary German consonants

a. [+grave] consonants



b. [−grave] consonants



In Halle's chart the segments /b, d/ lack specifications for [strident]. In the tree in (10) they must both be specified [–strident]. A [strident] specification can be omitted if we reorder the feature [tense] higher than [strident]; but then /pf, ts/ would have to receive a specification for [tense]. The possible specifications of /b, p, pf/ for the features [strident] and [tense], assuming the other values as listed, are shown in (11). I have shaded the specifications that are zeros in Halle's chart. Evidently, Halle has derived these specifications using minimal pairs.

(11) Specifications of [–compact, +grave] segments by the SDA

a. strident > tense

	b	p	pf
strident	–	–	+
tense	–	+	0

b. tense > strident

	b	p	pf
tense	–	+	+
strident	0	–	+

However, it could not be the case that minimal pairs were used to derive most of the specifications. For example, all the [–vocalic, +consonantal] non-nasal phonemes are specified [–nasal], though only /b/ and /d/ are distinguished from nasals solely by that feature. Thus, the chart appears to have been basically derived by feature ordering, and then perhaps some adjustments were made to the specifications. The discrepancy between the order of features in the table and the order required to derive the specifications may be due to a wish to maintain the general order used for other languages in tension with an order that appeared to be more appropriate for this particular language.

4.6 Ordered rules and the contrastive hierarchy

Halle's *The sound pattern of Russian* (1959) is a major work in early generative phonology. The book is in two parts: the first part presents a phonological analysis of Russian, and the second, longer, part is devoted to acoustics. Halle sets out a number of principles that a phonological theory should observe. Prominent among these are that features must be ordered into a hierarchy.

Halle's first argument in support of this principle is based on the need to keep segments properly distinct. To this end he proposes that feature specification must adhere to the Distinctness Condition (see section 2.5.1). He argues that specifying phonemes by 'branching diagrams' is the only way to ensure that phonemes meet this condition.[7]

Halle also requires that the number of specified features be minimal, in keeping with the requirement that they contrast. He argues that this condition requires feature ordering. He discusses some examples of branching diagrams, showing how a different ordering of features can produce different specifications. He observes that it is possible that features could apply in a

7 Halle (1959: 34 n. 26) mentions that his discussion of specification via branching diagrams 'is based in part on an unpublished paper by N. Chomsky and M. Halle "On the Logic of Phonemic Description," presented at the M.I.T. Conference on Speech Communication, June 16, 1956'. I have found no other mention of this paper. One can only speculate how generative grammar might have developed had this topic been pursued.

different order in different parts of the branching diagram, though he writes that he does not know if such cases arise in natural languages.[8]

There are limits to the extent to which Halle expects the feature hierarchy to have consequences for the workings of the phonology. He does in places refer to empirical reasons for ordering the features in a particular way. Thus, he observes (1959: 37) that intermediate nodes in the tree represent classes of segments (essentially the Prague School archiphonemes), and that these nodes play an important role in the functioning of a language. Again, in discussing the features [vocalic] and [consonantal], he observes (1959: 52) that these features 'show a high degree of negative correlation' and one might suppose that they should be replaced by less redundant features. However, he counters that these features establish the major classes of Russian, and allow for the simple characterization of phonological constraints on the construction of Russian phonemes. Therefore, in these instances he appeals to the patterning of Russian phonology to support a particular feature hierarchy, even at the expense of minimality. Elsewhere, however, minimality plays a more important role. Thus, he chooses to order [strident] ahead of [continuant] because this ordering results in fewer specifications.

Halle (1959) presents one of the first analyses to use synchronic rule ordering to convert underlying lexical forms to surface phonetic forms.[9] In contrast to the prevailing neo-Bloomfieldian theory, Halle does not split the phonology into morphophonemic and phonemic components, where the first converts morphophonemes into phonemes, and the latter accounts for the allophonic realizations of phonemes. Rather, Halle views the conversion of underlying to surface forms as occurring gradually in the course of the derivation. Thus, he makes no fundamental distinction between rules that change a specified feature value into another value (a plus into a minus or vice versa) and rules that assign a plus or minus to a zero value.

Halle argued that the seamless conversion of underlying to surface representations allowed one to capture generalizations that would be lost in a theory that separates morphophonemic from phonemic rules. However, this same reasoning comes into conflict with the Contrastivist Hypothesis.

This conflict can be illustrated by the case of Russian voicing assimilation, the subject of Halle's famous argument against what Chomsky later called

8 In fact, he uses different orders in different parts of the tree in his specification of Russian several pages later.

9 Chomsky (1951) had applied a series of ordered rules in his analysis of Hebrew.

the 'taxonomic phoneme' (see section 5.1). Halle (1959: 45) proposes the specifications in (12) for Russian.[10] The features apply in the order indicated.

(12) Feature specifications: Russian (Halle 1959)

	t	d	t^j	d^j	n	n^j	ts	s	z	s^j	z^j
vocalic	–	–	–	–	–	–	–	–	–	–	–
consonantal	+	+	+	+	+	+	+	+	+	+	+
diffuse	0	0	0	0	0	0	0	0	0	0	0
compact	–	–	–	–	–	–	–	–	–	–	–
low tonality	–	–	–	–	–	–	–	–	–	–	–
strident	–	–	–	–	–	–	+	+	+	+	+
nasal	–	–	–	–	+	+	0	0	0	0	0
continuant	0	0	0	0	0	0	–	+	+	+	+
voiced	–	+	–	+	0	0	0	–	+	–	+
sharped	–	–	+	+	–	+	0	–	–	+	+
accented	0	0	0	0	0	0	0	0	0	0	0

	p	b	p^j	b^j	m	m^j	f	v	f^j	v^j
vocalic	–	–	–	–	–	–	–	–	–	–
consonantal	+	+	+	+	+	+	+	+	+	+
diffuse	0	0	0	0	0	0	0	0	0	0
compact	–	–	–	–	–	–	–	–	–	–
low tonality	+	+	+	+	+	+	+	+	+	+
strident	–	–	–	–	–	–	+	+	+	+
nasal	–	–	–	–	+	+	0	0	0	0
continuant	0	0	0	0	0	0	0	0	0	0
voiced	–	+	–	+	0	0	–	+	–	+
sharped	–	–	+	+	–	+	–	–	+	+
accented	0	0	0	0	0	0	0	0	0	0

10 The specifications are very similar to those proposed by Cherry, Halle and Jakobson (1953). As before, the transcription is slightly modified to reflect contemporary usage.

	tʃ	ʃ	ʒ	k	kʲ	g	x	j	r	rʲ	l	lʲ
vocalic	−	−	−	−	−	−	−	−	+	+	+	+
consonantal	+	+	+	+	+	+	+	−	+	+	+	+
diffuse	0	0	0	0	0	0	0	0	0	0	0	0
compact	+	+	+	+	+	+	+	0	0	0	0	0
low tonality	−	−	−	+	+	+	+	0	0	0	0	0
strident	0	0	0	0	0	0	0	0	0	0	0	0
nasal	0	0	0	0	0	0	0	0	0	0	0	0
continuant	−	+	+	−	−	−	+	0	−	−	+	+
voiced	0	−	+	−	−	+	0	0	0	0	0	0
sharped	0	0	0	−	+	0	0	0	−	+	−	+
accented	0	0	0	0	0	0	0	0	0	0	0	0

	e	é	o	ó	a	á	i	í	u	ú
vocalic	+	+	+	+	+	+	+	+	+	+
consonantal	−	−	−	−	−	−	−	−	−	−
diffuse	−	−	−	−	−	−	+	+	+	+
compact	−	−	−	−	+	+	0	0	0	0
low tonality	−	−	+	+	0	0	−	−	+	+
strident	0	0	0	0	0	0	0	0	0	0
nasal	0	0	0	0	0	0	0	0	0	0
continuant	0	0	0	0	0	0	0	0	0	0
voiced	0	0	0	0	0	0	0	0	0	0
sharped	0	0	0	0	0	0	0	0	0	0
accented	−	+	−	+	−	+	−	+	−	+

As can be seen from the charts, voicing is a contrastive feature in Russian that distinguishes pairs of obstruent phonemes: /t/ and /d/ have opposite specifications for the feature [voiced], as do /k/ and /g/, /s/ and /z/, and so on. Of special interest are the Russian obstruents that do not have voiced counterparts, /ts, tʃ, x/. Halle (1959: 22–3) observes that these phonemes participate in voicing alternations just the way other obstruents do. In particular, there is a rule of regressive voicing assimilation (RVA) that assimilates all obstruents in a cluster to the voicing of its final obstruent. Halle's formulation is given in (13).

(13) Regressive voicing assimilation (Halle 1959: 64)
Rule P 3a. If an obstruent cluster is followed by a [phrase] boundary or by a sonorant, then with regard to voicing the cluster conforms to the last segment.

Halle observes that in the prevailing neo-Bloomfieldian theory the rule of RVA would have to apply twice: once in the morphophonemic component, where the result is an existing phoneme (14a); and again at the phonemic level to create voiced allophones of the unpaired phonemes /ts, tʃ, x/ (14b).

(14) Russian regressive voicing assimilation applying twice

a. Morphophonemic voicing		
Morphophonemes	//mók bi//	//ʒétʃ bi//
RVA	móg bi	–
(Taxonomic) phonemes	/móg bi/	/ʒétʃ bi/
b. Allophonic voicing		
(Taxonomic) phonemes	/móg bi/	/ʒétʃ bi/
RVA	–	ʒédʒ bi
Systematic phonetic form	[móg bɨ]	[ʒédʒ bɨ]

We find [móg bɨ] 'were (he) getting wet', where *k* voices to *g* before voiced obstruent *b* (compare [mók l^ji] 'was (he) getting wet?', with a *k* preceding the sonorant *l*). The rule that changes *k* to *g* changes one phoneme to another, and so it must be a morphophonemic rule (14a). This result is forced in any phonemic theory that observes the constraint that allophones of different phonemes may not overlap: in this case, [k] may not be an allophone of both /k/ and /g/. But the very same process produces [ʒédʒ bɨ] 'were one to burn', where *dʒ* is the voiced counterpart of *tʃ* (compare [ʒétʃ l^ji] 'should one burn?', with voiceless *tʃ* before l^j). Because [dʒ] is not a phoneme in its own right, but exists only as an allophone of /tʃ/, this application of voicing is an allophonic rule, and must be assigned to the component that maps phonemic forms into phonetic forms (14b).[11]

Since the alternations between [k] ∼ [g] and [tʃ] ∼ [dʒ] occur under the exact same conditions, the analysis in (14) must state the same generalization twice. By giving up the condition on non-overlapping, and hence giving up a

11 The surface vowel in [bɨ] is the result of a rule backing /i/ following a 'hard' (non-palatalized) consonant. This rule is discussed further in section 8.3.

level of taxonomic phonemes that observes this and other such conditions, the rule of RVA can be stated once for both types of case.

The unpaired phonemes /ts, tʃ, x/ function as regular obstruents not only as targets of regressive voicing assimilation: when they are in a position to trigger the rule (when they are final in an obstruent cluster), they function as ordinary voiceless consonants, not as consonants with no specification for voicing. To capture this result, Halle proposes the rule in (15), which applies before RVA.[12]

(15) Specification of unpaired segments (Halle 1959: 63)
Rule P 1b. Unless followed by an obstruent, /ts/, /tʃ/ and /x/ are voiceless.

Halle's derivation of the voicing alternations discussed above is presented in (16). The phoneme /tʃ/, unspecified for [voiced] at the underlying level, is converted to [–voiced] early in the derivation, making it equivalent to phonemes with underlying specifications for this feature with respect to all further voicing rules. Thus, /tʃ/ behaves just like /k/.

(16) Derivation of Russian voicing

a. Systematic phonemes	/mók bi/	/ʒétʃ bi/	/mók lʲi/	/ʒétʃ lʲi/
[voiced]	0 – +	+ 0 +	0 – 0	+ 0 0
b. Rule P 1b (specify *tʃ*)				–
c. Rule P 3a (RVA)	+	+		
d. Other rules	[móg bɨ]	[ʒédʒ bɨ]	[mók lʲi]	[ʒétʃ lʲi]

Though it is unclear to what extent any of the spirit of the Contrastivist Hypothesis still animated thinking about phonology, one can see that the introduction of a seamless set of ordered rules is a major blow against it. Recall that the Contrastivist Hypothesis posits a major distinction between contrastive and redundant features: only the former are computed by phonological operations. But if unspecified values can be freely changed to specified ones in the course of a derivation, then the distinction becomes unimportant: a redundant feature that is filled in acts just like a contrastive feature with respect to all subsequent rules.

This may be why the question of whether the voicing feature is really non-contrastive for the unpaired phonemes did not arise. In the specifications proposed by Halle, these phonemes have no specification for [voiced], just like the sonorants. This result follows from the ordering, which puts place features above manner features. A different result would follow if the feature [voiced]

12 As throughout, I have replaced Halle's symbols by the IPA symbols, and changed his {} to slant brackets, as per current practice.

were placed higher in the order. This issue is important with respect to the hypothesis that only contrastive features are active: the fact that /tʃ/ triggers devoicing, like all the other voiceless obstruents, suggests that the voicing feature is active, hence by hypothesis contrastive, in this segment. Halle (1959), however, did not assume this kind of connection between contrast and activity; by an early rule (P 1b) the feature [voiced] became *as if* contrastive in /tʃ/ with respect to the rule of RVA (rule P 3a). Therefore, there was no compelling reason to question or revise the feature hierarchy on these grounds. Rather, minimality of specification became the most compelling criterion for ordering features.

Minimality also dictated that feature specifications made predictable by sequential (contextual) constraints should also be removed from underlying forms, and this, too, led to a blurring of the specifications derived by feature ordering. To account for specifications made predictable in this way, Halle proposes a set of morpheme structure rules. For example, if a Russian word begins with a glide (G), the following segment must be a vowel (V), as there are no initial clusters GC where a consonant (C) follows a glide. Therefore, to satisfy minimality, Halle (1959) proposes that the predictable specifications for [vocalic] and [consonantal] could be omitted from a vowel that follows an initial glide.

Contextual underspecification greatly complicates the nature of underlying representations. Now, it is not just feature ordering that governs specification, but a large array of morpheme structure rules that refer to sequences. As with paradigmatic underspecification, there are many ways of arriving at contextual underspecification, for there are many possible contexts one can consider.

4.7 Summary: the contrastive hierarchy on the edge

The developments sketched above show that, by the 1960s, the contrastive hierarchy had lost much of its *raison d'être*. In the earlier work of Jakobson, the main motivation for contrastive specification was the patterning of phonological systems, assuming some version of the Contrastivist Hypothesis. The contrastive hierarchy as a method for arriving at contrastive specifications remained an obscure principle until the early 1950s, when 'branching diagrams' were viewed as 'the pivotal principle' of linguistic structure by Jakobson, Fant and Halle (1952). All the ingredients existed, at that moment, for the development of a theory in which the Contrastivist Hypothesis and the contrastive hierarchy played central roles.

Phonological theory, however, did not advance in that direction. Rather, considerations of efficient coding and minimality came to the fore. Failure to find a universal feature order may have led to a certain loss of faith in the utility of feature ordering. Finally, the advent of a single gradual derivation from lexical to phonetic representations blurred the distinction between contrastive and redundant specifications, resulting in the total eclipse of the Contrastivist Hypothesis. Lacking this empirical motivation, the stage was set for the contrastive hierarchy to be ushered out of phonological theory when underspecification itself came under attack.

4.8 Stanley (1967) and the end of zeros in generative phonology

Stanley's famous article, 'Redundancy rules in phonology' (1967), is well known in the phonological literature for its attack on the potential for misuse of zero values in binary feature matrices. This article convinced phonologists that the phonological component should be limited to working with fully specified matrices. In place of filling in empty values by redundancy rules, Stanley proposed the adoption of morpheme structure constraints. The main proposals of this article were adopted by Chomsky and Halle (1968 [*SPE*]) and other influential works and effectively put an end to underspecification in phonology for the next fifteen years.

Stanley's general arguments against underspecification would have sufficed to put an end to any sort of underspecification. However, he also made some arguments explicitly against the 'branching diagrams', that is, the contrastive hierarchy, as used by Halle (1959). We will see that the main arguments against contrastive specification rest on the assumption that underspecification of a feature is merely a matter of notation and should have *no* empirical consequences for how the phonology works.

4.8.1 Zeros and ternary power

Following Lightner (1963), Stanley (1967: 413) points out that the presence of unspecified feature values in the phonology raises formal questions about how rules apply in such situations. He shows that, whatever the definition, there is the danger that binary features will be used in a ternary manner, whereby 0 contrasts with both + and −.

He gives the example of the initial matrices in (17a). Phoneme A has no specification for feature f, B is [+f] and C is [−f]. In the view of redundancy rules commonly held at the time, the omission of a feature value in A is simply supposed to be a way of reducing the information content of a segment, as indicated by the number of symbols that need to be specified for it. Since f is a binary feature, it should only be able to make a binary distinction. Thus, A has either the specification [+f] or [–f], but this specification is omitted in underlying representation because it is redundant in this segment.

(17) Rule application: sub-matrix interpretation (Stanley 1967: 413)

a. Initial matrices

$$\underset{\textstyle\mathrm{A}}{\begin{bmatrix} \\ \vdots \end{bmatrix}} \quad \underset{\textstyle\mathrm{B}}{\begin{bmatrix} +\mathrm{f} \\ \vdots \end{bmatrix}} \quad \underset{\textstyle\mathrm{C}}{\begin{bmatrix} -\mathrm{f} \\ \vdots \end{bmatrix}}$$

b. Phonological rules

i. [ ] → [–g] ii. [+f] → [+g]

iii. [–f] → [+g] iv. [–g] → [+f]

c. Derived matrices

$$\underset{\textstyle\mathrm{A}}{\begin{bmatrix} +\mathrm{f} \\ -\mathrm{g} \\ \vdots \end{bmatrix}} \quad \underset{\textstyle\mathrm{B}}{\begin{bmatrix} +\mathrm{f} \\ +\mathrm{g} \\ \vdots \end{bmatrix}} \quad \underset{\textstyle\mathrm{C}}{\begin{bmatrix} -\mathrm{f} \\ +\mathrm{g} \\ \vdots \end{bmatrix}}$$

Consider now the ordered phonological rules in (17b). Rule (17bi) states that a segment unspecified for g is specified [–g]. The next two rules apply to segments that are specified [+f] and [–f], respectively. The formal question that arises is whether a segment that is unspecified for f should undergo these rules or not. On the sub-matrix interpretation of rule application, a segment is subject to a rule only if the rule is a sub-matrix of the feature matrix of the segment. On this interpretation, segment A does not undergo rules (ii) and (iii). But, then, if we add a redundancy rule like (17biv), all three segments A, B and C will end up being distinct, though they started out distinguished only by the feature f.

Stanley shows that similar results can be obtained if we adopt a different convention of rule application. On the distinctness interpretation of rule application, a rule applies to a matrix if the matrix is non-distinct from the feature matrix of the segment. In that case, a segment unspecified for f will undergo any rule that mentions either [+f] or [–f], and segments specified for f undergo rules that do not mention f, if their matrices are otherwise non-distinct from the structural description of the rule. No matter what the default or redundant value of f turns out to be, all three segments end up distinct in their values for features g and h, and again a binary feature has functioned in ternary fashion (18). Stanley (1967: 414) argues that this is an ‘improper use of blanks’ that gives ternary power to a binary system.

(18) Rule application: distinctness interpretation (Stanley 1967: 414)

a. Initial matrices

$$\underset{\text{A}}{\begin{bmatrix} \\ \vdots \end{bmatrix}} \quad \underset{\text{B}}{\begin{bmatrix} +\text{f} \\ \vdots \end{bmatrix}} \quad \underset{\text{C}}{\begin{bmatrix} -\text{f} \\ \vdots \end{bmatrix}}$$

b. Phonological rules

i. [ ] → [–g, –h] ii. [+f] → [+g]

iii. [–f] → [+h] iv. Default: [ ] → [+f]

c. Derived matrices

$$\underset{\text{A}}{\begin{bmatrix} +\text{f} \\ +\text{g} \\ +\text{h} \\ \vdots \end{bmatrix}} \quad \underset{\text{B}}{\begin{bmatrix} +\text{f} \\ +\text{g} \\ -\text{h} \\ \vdots \end{bmatrix}} \quad \underset{\text{C}}{\begin{bmatrix} -\text{f} \\ -\text{g} \\ +\text{h} \\ \vdots \end{bmatrix}}$$

On logical grounds, Stanley's argument is perfectly correct. However, it rests on the empirical assumption that distinctive features must function in a strictly binary way. Although this assumption was adopted by almost all generative phonologists in the subsequent decade, it is actually inconsistent with the Contrastivist Hypothesis, which posits that the phonology treats contrastive specifications differently from noncontrastive specifications.

Stanley (1967: 416) considers the case of a language like Russian, where all [–consonantal] segments are [+voiced]. By redundancy (and by feature ordering, assuming [consonantal] is ordered higher than [voiced]), we would therefore not specify such segments as [+voiced]. Suppose now that the language also has a rule (cf. RVA in Russian) that applies to the class of voiced obstruents, [+consonantal, +voiced]. Stanley observes that a phonological rule could just mention [+voiced], taking advantage of the fact that [–consonantal] segments are not specified at the point in the phonology where this rule applies.[13] He writes that 'this is a specious simplification, an improper use of blanks' (1967: 416).

What Stanley considered specious was later considered a result! Thus, we can *explain* why rules spreading voicing values do not affect sonorants, and are not triggered by sonorants, if sonorants indeed lack contrastive specifications for [voiced], and are barred (at least in a certain component of the grammar) from taking any on.[14]

13 Stanley assumes here the sub-matrix interpretation of rule application.

14 We may still want to rule out some of the situations Stanley wished to bar, not by banishing blanks from the phonology, but through other means, by limiting the power of phonological operations (see Dresher 1985 for further discussion).

4.8.2 Arguments against 'branching diagrams'

Stanley (1967: 407) opens his discussion of the 'branching diagrams' (contrastive hierarchy) of Halle (1959) by recalling Halle's (1959: 35) observation that 'not all ways of choosing the non-redundant feature values leave open the possibility of constructing a branching diagram'. Stanley remarks that the required distribution of non-redundant features is not always found. This comment appears to assume that branching trees are constructed working back from feature specifications, rather than looking at it the other way round – that specifications are the consequence of the branching tree. Stanley (1967: 408) notes that generative grammars have been chosen that have a branching diagram of this kind, and summarizes the reasons why it was regarded as important to have such a diagram:

> (1) Giving segments in a branching diagram appears to be the most direct means, involving the fewest feature specifications, which will guarantee that each pair of segments is DISTINCT in the sense that for any pair of different segments there is at least one feature f such that one member of the pair is specified +f and the other –f. (2) The branching diagram gives a hierarchy of features which can be interpreted as meaning that the features at high nodes (such as Consonantal) are in some sense more basic than the features at low nodes (such as Voiced). (3) The branching diagram gives a way of formalizing the notion of archiphoneme.

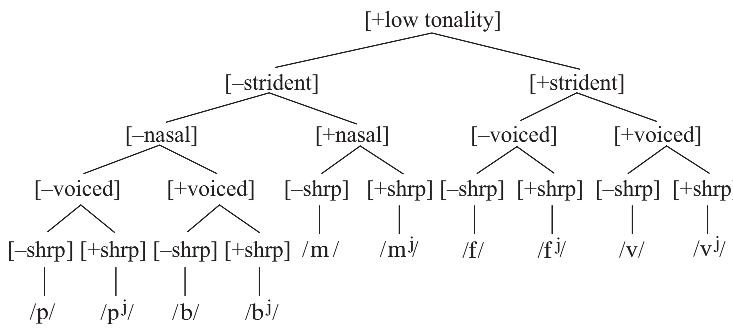
Stanley proceeds to take issue with each of these assumptions. He argues that the Distinctness Condition is both too strong and too weak. It is too weak because it does not prevent 'specious simplifications' of rules obtained by improper use of blanks, an issue we considered above. The Distinctness Condition is too strong, in Stanley's view, because it does not omit all logically redundant features. This argument would go through only if it were indeed a requirement of the theory that all logical redundancies be expressed by blank specifications. But we have seen in chapter 2 that it is neither necessary nor possible for a theory to express all logical redundancies. If we take specification by feature ordering as the fundamental way to express contrasts in a system, then a certain amount of underspecification follows from the contrastive hierarchy, if we interpret redundant features as unspecified. However, underspecification, on this view, is a consequence of establishing contrasts, not an end in itself.

Stanley's second argument against feature ordering is that it is 'somewhat strange' to capture the feature hierarchy in a branching diagram because many different such diagrams can be constructed for a given set of phonemes. Any such hierarchy must represent more than 'a vague set of intuitions' as to which features are more basic than others, which is a danger in the absence of evidence that a certain hierarchy is required by the facts of a particular language.

Stanley is correct that one must have evidence to support the proposed hierarchies. He even suggests (1967: 408) where such evidence could possibly be found: ‘perhaps . . . in terms of the different ways in which different features behave in the P[honological] rules or in the MS rules’. At the time, however, the difference between specified and unspecified feature values was not considered to have empirical consequences, and this avenue remained unexplored.

The third argument is that branching diagrams do not adequately formalize the notion of the archiphoneme. By ‘archiphoneme’, Stanley means a representation of the neutralization of a contrast, or a number of contrasts. For example, in Halle (1959), the feature [sharped] is at the bottom of the hierarchy for consonants. Thus, pairs of plain and sharped (palatalized) consonants are sisters at the bottom of the branching tree. A part of the tree showing the labial consonants is given in (19).

(19) Contrastive hierarchy for Russian labial consonants (Halle 1959)



Suspending these final branches results in ‘archiphonemes’ that represent consonants without regard to the value of the feature [sharped], but with all other contrastive features expressed. Here we have the archiphonemes /p, p^j/, /b, b^j/, /m, m^j/, /f, f^j/ and /v, v^j/. More extended archiphonemes can be represented by stopping the tree at higher nodes.

Stanley observes that not all archiphonemes that we may want to represent can be read off the tree. Thus, we may want to represent consonants without regard for their value for [voiced], but still retaining their values of [sharped], neutralizing the contrast between /p, b/, /p^j, b^j/, /f, v/ and /f^j, v^j/, for example. Yet these neutralizations cannot be read off the same tree; we would require a different tree in which [sharped] is higher in the contrastive hierarchy than [voiced]. Therefore, Stanley argues, the branching tree incorrectly limits the possible archiphonemes (neutralizations) that can be represented, and thus fails to capture the relevant notion of archiphoneme.

Stanley is correct if we look only at the tree structure without considering the representations derived from the trees. The contrastive hierarchy of which the tree in (19) is a part yields the representations in (20). Reading off the table, we see that voiced and voiceless counterparts are indeed minimal pairs, distinguished only by the feature [voiced]. Therefore, to the extent that representations are important apart from the contrastive hierarchy that generates them, the branching tree does not limit possible neutralizations as much as Stanley claimed it does.

(20) Feature specifications: Russian labial consonants (Halle 1959)

	p	b	p^j	b^j	m	m^j	f	v	f^j	v^j
vocalic	−	−	−	−	−	−	−	−	−	−
consonantal	+	+	+	+	+	+	+	+	+	+
compact	−	−	−	−	−	−	−	−	−	−
low tonality	+	+	+	+	+	+	+	+	+	+
strident	−	−	−	−	−	−	+	+	+	+
nasal	−	−	−	−	+	+	0	0	0	0
continuant	0	0	0	0	0	0	0	0	0	0
voiced	−	+	−	+	0	0	−	+	−	+
sharped	−	−	+	+	−	+	−	−	+	+

Of course, any choice of a contrastive hierarchy does limit possible neutralizations beyond what we can obtain with full specification. For example, the representations in (20) do not allow us to group /m/ with /b/ against /p/, since /m/ has no value for [voiced]. But this limitation is undesirable only if there is evidence in the language that /b, m/ act together as being [+voiced]. In Russian, however, sonorants do not participate in rules of voicing or devoicing, so the representations in (20) are supported in this respect.

4.8.3 Summary: 'there is obviously some kind of hierarchical relationship'

In sum, Stanley's arguments against the branching trees, and hence against the contrastive hierarchy, rest mainly on logical, not empirical, criteria. The logical points are correct as far as they go; however, his main empirical assumptions have become less relevant as phonological theory has developed. Nevertheless, his arguments carried the day, and branching trees disappeared from mainstream generative phonology for the rest of the century.

Still, the intuition that, as Stanley put it, ‘there is obviously some kind of hierarchical relationship among the features which must somehow be captured in the theory’ (1967: 408), continued to haunt generative phonological theory, and took a number of different forms. Three of them coexisted uneasily at the same time at the heart of mainstream generative phonology: markedness theory, underspecification theory and feature geometry. All bear interesting, though seldom discussed, affinities with the contrastive hierarchy, and these form the subject matter of the next chapter.

5 *Generative phonology: contrast goes underground*

5.1 Chomsky and Halle's revolution in phonology

Chomsky and Halle's approach to phonological theory, as represented in their major work *The sound pattern of English* (Chomsky and Halle 1968, henceforth *SPE*), represented a sharp break with the main currents of American linguistics that immediately preceded them. Nevertheless, some elements of the Jakobson–Halle approach to phonology were continued in generative phonology. In general terms, Chomsky and Halle's theory of generative phonology was a synthesis of Jakobson and Halle's theory of distinctive features and phonemic analysis, revised in the light of Chomsky's emphasis on formal explicitness, simplicity, and abstractness and autonomy of mental representations (Dresher 2005).

One aspect of Jakobson and Halle's theory that did not make it into *SPE* was the contrastive hierarchy. As discussed in the previous chapter, the atmosphere was not conducive to assigning a special role to contrastive specifications. One of the major points of difference between the older structuralist approach and generative phonology was the status of the taxonomic phoneme (Chomsky 1964), alluded to above in section 4.6. Classical generative phonology posits only two significant levels in the phonology: the underlying lexical, or 'systematic phonemic' level (known as the 'morphophonemic' level in structuralist theory), and the surface phonetic level (a level that was not recognized as a systematic level in structuralist phonology, following the arguments of Bloomfield 1933). Instead of a phonetic level, the neo-Bloomfieldians posited what Chomsky called the 'taxonomic phonemic' level, a level defined largely in terms of surface-oriented criteria designed to simplify the acquisition problem (see Dresher 2005 for discussion). Chomsky and Halle argued that this level was superfluous, causing a loss of generalizations by requiring certain phonological rules, such as Russian RVA, to be split across two components, as demonstrated in section 4.6 in the previous chapter.

As I observed there, the Russian example does not necessarily show that the contrastive status of a phoneme is irrelevant to the functioning of RVA. However, it may have seemed so at the time. More generally, an emphasis on a single seamless derivation from lexical to phonetic representations is not compatible with dividing up the phonology into a component that is concerned with contrastive representations and a component that allows for redundant features as well; such a proposal would have looked like a move to reestablish the taxonomic phoneme.

Therefore, the basic theory of *SPE* grants no special role to contrastive feature specifications, and in this respect amounts to a complete rejection of the Contrastivist Hypothesis implicit in much early phonological theory. However, generative phonology found over time that it could not get along without something like the Contrastivist Hypothesis and a contrastive feature hierarchy, and several subtheories arose within generative phonology to fill this gap. In this chapter, I focus on the treatment of contrast within 'classical' generative phonology, beginning with the first attempt, in *SPE* itself, to remedy some shortcomings brought about, in part, by its rejection of the Contrastivist Hypothesis. I will look at three subtheories that arose with generative phonology: markedness (section 5.2), underspecification (section 5.3), and feature geometry (section 5.4). I will concentrate on ways in which they are similar to, and differ from, the contrastive hierarchy.

5.2 Markedness

5.2.1 Chapter 9 of SPE

Chapter 9 of *SPE* opens with a dramatic statement (*SPE*: 400): 'The entire discussion of phonology in this book suffers from a fundamental theoretical inadequacy.' This inadequacy consists of *SPE*'s 'overly formal' approach to features, which does not take into account their intrinsic content. If all segments are fully specified in terms of the same features, and the evaluation measure counts only the number of symbols, then there is no basis for evaluating any segment as being more or less costly, or complex, than any other. As a result, the *SPE* theory as presented in the previous eight chapters could not account for why certain segments are more common than others, or why certain segmental inventories are common and others are not.

In addition, Chomsky and Halle observe that the problem is not limited to inventories, but has implications for the functioning of phonological rules. They present a number of examples, some of which are reproduced in (1), of pairs of rules to which a purely formal evaluation measure would assign incorrect relative costs.[1]

(1) Pairs of rules (*SPE*: 401)

a.	i. i → u	ii. i → ɨ
b.	i. t → s	ii. t → θ
c.	i. $k \rightarrow t\int / \_\_\_ \begin{bmatrix} -\text{cons} \\ -\text{back} \end{bmatrix}$	ii. $t\int \rightarrow k / \_\_\_ \begin{bmatrix} -\text{cons} \\ +\text{back} \end{bmatrix}$

In (1a), rule (i) is formally more complex than (ii) because it changes an extra feature, [round], in addition to a change in [back]: *i* is [−back, −round], among other features not in play here; *ɨ* is [+back, −round]; and *u* is [+back, +round]. Nevertheless, rule (i) is much more common than rule (ii). This is presumably because [u] is a more common segment than [ɨ], a fact not recognized by the *SPE* theory, which only counts the number of symbols in a rule, and does not evaluate their content. In this case the cost of the extra feature in rule (i) is outweighed by the more natural result. Or, rather, the cost of the extra feature is only apparent.

The case of (1b) is similar. Of the features relevant to this rule, *t* is [−continuant, −strident], θ is [+continuant, −strident] and *s* is [+continuant, +strident]. Rule (1bii) appears to be the cheaper rule because it changes only [continuant] in the *SPE* feature system, whereas rule (i) also changes [strident]. Again, the formally simpler rule is less common than the more complex one, because [s] is a more common segment than [θ].

The example in (1c) looks a bit different, but ultimately falls under the same heading in Chomsky and Halle's analysis. They point to an asymmetry in the direction of a rule: whereas the change of /k/ to [tʃ] before a front vowel is quite common,[2] the converse change of /tʃ/ to [k] before a back vowel is not. Since the two rules are formally very similar, once again the formal simplicity criterion fails to distinguish the common rule from the rare one.

Contrast by itself does not account for why [u] and [s] are more common than [ɨ] and [θ], respectively; but contrast is relevant to the observations in (1). Thus, rule (i) in (1a) is more common than rule (ii) only in vowel systems that lack a three-way contrast between /i, ɨ, u/. In inventories with all three vowels, the result of simply backing /i/ would indeed be *ɨ*; to arrive at *u*, it would be necessary to change not just [back], but also [round]. In such inventories, in other words, the formal simplicity measure gives the right answer. Similar

1 In this chapter I have modified the phonetic symbols in the sources to reflect IPA values.

2 Actually, the fronting of /k/ to [tʃ] most usually occurs before /i/ and sometimes also /e/, but rarely before the low front vowel /æ/.

considerations hold for the change of /t/ to [s]: rule (i) in (1b) is only more common than rule (ii) in inventories that lack a phoneme /θ/.

To address these problems, Chomsky and Halle (*SPE*) appeal to a modified version of the Prague School notion of markedness. The *SPE* markedness theory classifies segments as either unmarked (u) or marked (m) for particular features. A set of *marking conventions* converts markedness values into +/− values before the phonology proper begins. Markedness conventions can nevertheless interact with the phonology via the concept of *linking*. I will first briefly review some of the marking conventions, and then look at how linking works. I will show that these devices do some of the work of specification by a contrastive hierarchy, but run into problems because they do not take account of language-particular contrasts.

5.2.2 SPE *marking conventions and feature hierarchy*

Chomsky and Halle (*SPE*: 404–7) present a 'tentative statement' of the marking conventions. To illustrate the form of these statements I reproduce the conventions for vowels in (2).

(2) Some markedness conventions for vowels (*SPE*: 405)

$$\text{a. V} \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \end{bmatrix} \rightarrow \begin{bmatrix} -\text{ant} \\ -\text{strid} \\ +\text{cont} \\ +\text{voice} \\ -\text{lateral} \\ \textit{etc.} \end{bmatrix}$$

$$\text{b. VI} \quad [\text{u low}] \rightarrow \left\{ \begin{array}{ll} [+\text{low}] / \begin{bmatrix} \overline{} \\ \text{u back} \\ \text{u round} \end{bmatrix} & \text{(i)} \\ [-\text{low}] & \text{(ii)} \end{array} \right\}$$

$$\text{c. VII} \quad [+\text{low}] \rightarrow [-\text{high}]$$

$$\text{d. VIII} \quad [\text{u high}] \rightarrow [+\text{high}]$$

$$\text{e. IX} \quad [+\text{high}] \rightarrow [-\text{low}]$$

$$\text{f. X} \quad [\text{u back}] \rightarrow [+\text{back}] / \begin{bmatrix} \overline{} \\ +\text{low} \end{bmatrix}$$

$$\text{g. XI} \quad [\text{u round}] \rightarrow \left\{ \begin{array}{ll} [\alpha\text{round}] / \begin{bmatrix} \overline{} \\ \alpha\text{back} \\ -\text{low} \end{bmatrix} & \text{(i)} \\ [-\text{round}] / \begin{bmatrix} \overline{} \\ +\text{low} \end{bmatrix} & \text{(ii)} \end{array} \right\}$$

$$\text{h. XII} \quad [\text{u tense}] \rightarrow [+\text{tense}]$$

The marking conventions are rather heterogeneous. Some take as their inputs an unmarked feature value, [u F], whereas others have +/− values as their input. These latter are statements of necessary universal implications, such as that [+low] vowels are necessarily [−high] (VII), or that [+high] vowels are necessarily [−low] (IX). Most rules converting u values to +/− values have contexts specified only with +/− values, or variables ranging over these values (V, X, XI), but some conventions also contain u/m values in their contexts (only VI in the conventions shown, but the full set of markedness conventions contains further examples).

Not all features are governed by markedness conventions in all contexts. For example, convention X states that the unmarked value of [back] is + in the context [+low] (low vowels are preferably back: /a/ is preferred to /æ/), but there is no corresponding convention for [back] in the context [−low]. Therefore, non-low vowels have no u/m value for [back], but only a +/− value.

No marking conventions convert a marked feature into +/− values. If for some feature, F, a rule converts [u F] into [αF] in some context C, it follows logically that a segment specified [m F] in the same context must receive the opposite value, [−αF]. To make this explicit, Chomsky and Halle (*SPE*: 403) propose the interpretive convention in (3). Following Kean (1980) I will refer to the rule replacing [m F] as the *complement* of the rule replacing [u F] in the same context.[3]

(3) Interpretive convention for marking statements
Each schema [u F] → [αF] / X _____ Y, where α = + or −, and X and Y may be null, is interpreted as a pair of rules, the first of which replaces [u F] by [αF] in the context X _____ Y and the second of which replaces [m F] by [−αF] in the context X _____ Y.

Of special interest to our topic is the relationship of the markedness conventions to feature hierarchies. Though not completely ordered, we find that the markedness conventions do encode a partial ordering of features.

At the top of the hierarchy are the major class features [consonantal] and [vocalic]. They are at the top of the list because their markedness conventions (not shown here) apply first. These conventions are different from most of the others in that they apply recursively. Also, there is no hierarchical relation between these two features: sometimes [vocalic] presupposes [consonantal]

3 Kean's (1980) Complement Convention subsumes more than this, however. See the next section for further discussion.

and sometimes the opposite is the case. Thus, no well-formed branching tree can model these markedness conventions.

The recursive manner of application is limited to the conventions for these features. Chomsky and Halle propose (*SPE*: 408) that the remaining conventions apply only once and in order. We will investigate the relationship between this order and feature ordering with respect to the markedness conventions for the vowels.

The conventions for the main vowel features, [low], [high], [back] and [round], apply in this order, which might at first suggest that this is their order in the hierarchy. The conventions depart from a branching tree in that the convention for [low] (VI) mentions [u back] and [u round]; in particular, the unmarked value of [low] is + for a vowel that is unmarked for [back] and [round]. This statement is inconsistent with a contrastive hierarchy, in that a higher-order feature refers to one ordered later. Nevertheless, the vowel conventions V–XII correspond most closely to a hierarchy in which [low] is higher than [back] and [round]. The convention for [back] (X) presupposes specification of [+low], and the convention for [round] (XI) presupposes +/− values of both [low] and [back]. The relationship between [low] and [high] also departs from a consistent hierarchy: [+low] leads to the specification of [−high] (VII), presupposing that [low] is specified first, but IX states that [+high] implies [−low], presupposing the opposite.

Chomsky and Halle perceive a hierarchy among the features, though they modify the markedness theory for vowels somewhat to achieve it. The immediate reason for modifying the results of the markedness conventions is their dissatisfaction with the predicted vowel inventories that emerge from these conventions. Assigning a cost of 0 to a *u* and 1 to *m*, + or −, with the total cost of each segment being its *complexity*, the theory assigns the feature representations of some typical vowels the complexity measure shown in (4).

(4) Markedness matrices for vowels (*SPE*: 409)

	a	i	u	æ	ɒ	e	o	y	ɨ	Œ	ø	ʌ
low	u	u	u	m	m	u	u	u	u	m	u	u
high	u	u	u	u	u	m	m	u	u	u	m	m
back	u	−	+	m	u	−	+	−	+	m	−	+
round	u	u	u	u	m	u	u	m	m	m	m	m
Complexity	0	1	1	2	2	2	2	2	2	3	3	3

The conventions predict, correctly, that the simplest possible three-vowel system is /a, i, u/. However, for five-vowel systems, the theory assigns the same complexity score to every system containing /a, i, u/ and any two of /æ, ɒ, e, o, y, ɨ/. For example, the commonly found and presumably optimal five-vowel system /a, i, u, e, o/ receives the same complexity score as the unlikely /a, i, u, y, ɨ/.

Rather than modify the markedness conventions, Chomsky and Halle propose to add two general conditions on lexical representations, as in (5).

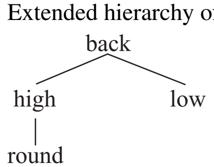
(5) Conditions on lexical representations (*SPE*: 410)

a. No vowel segment can be marked for the feature ‘round’ unless some vowel segment in the system is marked for the feature ‘high’.

b. Other things being equal, a system in which more features have only the specification *u* is preferable to a system in which fewer features have only the specification *u*.

Chomsky and Halle (*SPE*: 410) observe that condition (5a) ‘establishes a hierarchy in the availability of features for marking vowels in the lexicon. There are doubtless other conditions of this sort. Thus one would expect a hierarchy in which the feature “segment” is above “consonantal” and “vocalic,” and the latter two are above the features listed in [(4)].’ They go on to say that (5a) should be extended so that features [high] and [low] would be available for marking only after [back] has been marked, resulting in the hierarchical structure in (6).

(6) Extended hierarchy of vowel features (*SPE*: 410, ex. (11))



The hierarchy in (6) refers to marked values, and is not equivalent to a developmental path specified by a contrastive hierarchy. Thus, (6) does not require that the first contrast in a vowel system be between [+back] and [−back] (a position that would contradict the view of Jakobson and Halle 1956; compare Jakobson and Halle’s proposed acquisition sequence given in section 4.3). The maximally unmarked vowel remains /a/, and the unmarked three-vowel system remains /a, i, u/. The vowels /i, u/ are considered to be marked for [back],[4] so

4 Recall that /i, u/ are specified [−back] and [+back], respectively, in (4), rather than by *u* or *m*. These specifications count as marking, since they contribute to a segment’s complexity score.

other vowels in the inventory could be marked for [high] or [low]. Of the other vowels in (4) that have complexity 2, /ɒ, y, ɨ/ are marked for [round], so cannot be added to an inventory containing /a, i, u/, as none of these vowels are marked for [high].

This still leaves vowels /æ, e, o/ which have complexity 2. Now the vowel system /a, i, u, e, o/ is preferred to /a, i, u, æ, o/ by (5b) because in the former only the features [high] and [back] are marked, whereas in the latter [high], [back] and [low] are all marked. Thus, condition (5b) builds in a symmetry requirement, since symmetrical systems will tend to have marked values restricted to fewer segments than unsymmetrical ones. In this case, symmetrical /e, o/ are both marked for [back] and [high], whereas in unsymmetrical /e, æ/, /e/ is marked for [back] and [high] and /æ/ is marked for [back] and [low].

Thus, it is clear that Chomsky and Halle were moving in the direction of imposing a consistent feature hierarchy on the markedness statements of *SPE*, though this goal remains unachieved in *SPE* itself. A revision that completed this change was undertaken a few years later by Mary-Louise Kean.

5.2.3 Kean's (1980) revision of the marking conventions

Kean's 1975 MIT dissertation, published as Kean (1980), is a revision of the *SPE* theory of markedness. Kean proposes that markedness statements are restricted in format, so that for every feature, F, there is a *markedness convention* of the form in (7) that specifies a + or − value of the unmarked specification of F, [u F].

(7) Markedness convention (Kean 1980: 11)

$$[\text{u F}] \rightarrow [\alpha\text{F}] \,/\, X \qquad \text{where X is} \quad \begin{bmatrix} \overline{} \\ \beta_1 G_1 \\ \vdots \\ \beta_n G_n \end{bmatrix}$$

F, $G_1, \ldots, G_n$, are features, and $\alpha, \beta_1, \ldots, \beta_n$, are + or −.

For example, Kean proposes the markedness conventions in (8) for the main vowel features [back], [low], [lab] and [high].[5]

5 Kean employs a unified set of features that apply to vowels as well as consonants. For illustrative purposes, I have extracted the features that are most relevant to vowels. The feature [lab] is the same as [round] when applied to vowels. The Roman numeral indicates the ordering of the convention with respect to the entire list.

(8) Some markedness conventions (Kean 1980: 23–5)

a. V $[\text{u back}] \rightarrow [+\text{back}] \,/\, \begin{bmatrix} \_\_\_ \\ -\text{ant} \end{bmatrix}$

b. VI $[\text{u low}] \rightarrow [+\text{low}] \,/\, \begin{bmatrix} \_\_\_ \\ -\text{cons} \\ +\text{back} \end{bmatrix}$

c. VII $[\text{u lab}] \rightarrow [+\text{lab}] \,/\, \begin{bmatrix} \_\_\_ \\ -\text{cons} \\ +\text{back} \\ -\text{low} \end{bmatrix}$

d. XI $[\text{u high}] \rightarrow [-\text{high}] \,/\, \begin{bmatrix} \_\_\_ \\ \alpha\text{cons} \\ \alpha\text{ant} \\ \beta\text{cor} \\ -\beta\text{lab} \\ -\beta\text{spr} \end{bmatrix}$

Kean proposes that every markedness convention of the form in (7) projects a set of *markedness rules* according to the Complement Convention (9). In (9), [m F] is the marked value of feature F, X is the context as in (7), and $\bar{X}$ is the set of environments that make up the complement of X.

(9) The Complement Convention (Kean 1980: 11)

a. $[\text{u F}] \rightarrow [\alpha\text{F}] \,/\, X$
b. $[\text{m F}] \rightarrow [-\alpha\text{F}] \,/\, X$
c. $[\text{u F}] \rightarrow [-\alpha\text{F}] \,/\, \bar{X}$
d. $[\text{m F}] \rightarrow [\alpha\text{F}] \,/\, \bar{X}$

To illustrate, the markedness convention (8a) generates the markedness rules in (10). All vowels are [−anterior], so according to (8a) (= (10a)) it is unmarked for a vowel to be [+back], and thus it follows (10b) that it is marked for a vowel to be [−back]. Consonants that are [−anterior] (palatals, velars and other back consonants) are likewise unmarked if [+back] (velars) and marked if [−back] (palatals). The [+anterior] consonants (labials and prepalatal coronals) are the other way around, following (10c, d): their unmarked value is [−back], and marked value (if it can exist at all) is [+back].

(10) Markedness rules for [back] (Kean 1980: 11–12)

a. $[\text{u back}] \rightarrow [+\text{back}] \,/\, \begin{bmatrix} \_\_\_ \\ -\text{ant} \end{bmatrix}$

b. $[\text{m back}] \rightarrow [-\text{back}] \,/\, \begin{bmatrix} \_\_\_ \\ -\text{ant} \end{bmatrix}$

c. [u back] $\rightarrow$ [−back] / $\left[\frac{\quad}{+\text{ant}}\right]$

d. [m back] $\rightarrow$ [+back] / $\left[\frac{\quad}{+\text{ant}}\right]$

In this case the context of the markedness convention is simple, consisting only of a single feature, so the complement of the context is also simple. When the context is more complex, the complement rules multiply. For example, the complement of the environment [−cons, +back] consists of the three contexts [+cons, +back], [+cons, −back] and [−cons, −back].

The limitation on markedness conventions to a single context sometimes results in odd specifications. For example, it is unmarked for all vowels to be [+high], because vowels occur in the complement of the context of the markedness convention for [high] (8d). However, this results in the vowel /a/ receiving the specification [m high], though its unmarked value is [+low]. As Kean (1980: 28) points out, no vowel could be simultaneously [+low, +high]; once a vowel is specified [+low], it has no choice but to be [−high]. As there is no choice in the matter, [−high] cannot be a marked value, for that would imply that it is possible for the unmarked counterpart to exist. Rather than complicate the markedness conventions, Kean proposes to have a set of implications that capture co-occurrence restrictions on feature specifications. Some examples of such rules are listed in (11).

(11) Implicational co-occurrence restrictions (Kean 1980: 29–30)
a. [−cons] $\supset$ [+son]
b. [−cons] $\supset$ [−ant]
c. [+low] $\supset$ [−high]
d. [+lat] $\supset$ [+cor]

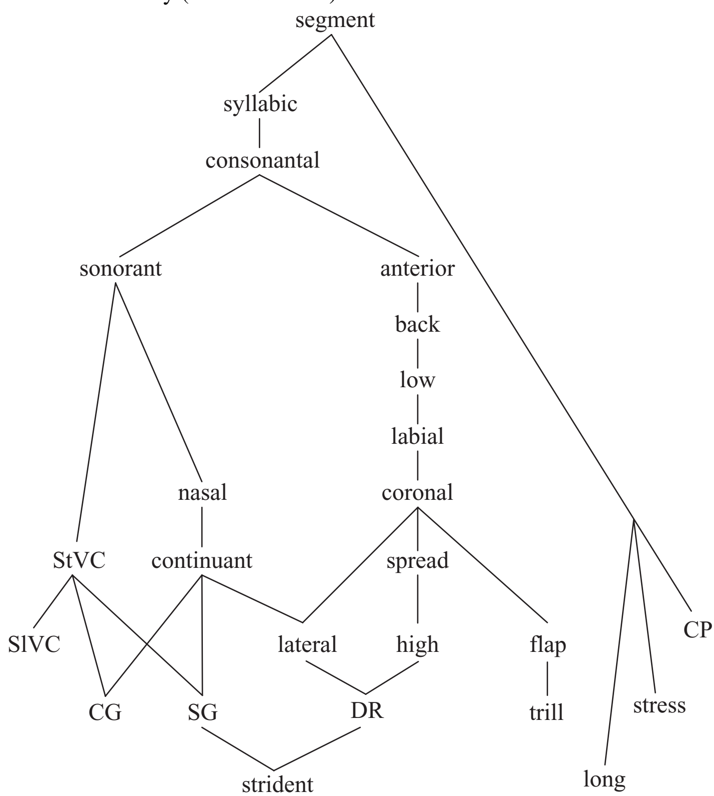
Kean proposes the principle in (12) to be part of the procedure for specifying feature values.

(12) Specification of co-occurrence restrictions (Kean 1980: 30)
Whenever a segment is specified to be [αF], where F is a feature and α is + or −, all implications whose antecedents are satisfied apply to that segment.

5.2.4 Markedness and feature hierarchy

It is evident even from the few examples in (8) that the markedness conventions are intrinsically ordered. The rule for [back] presupposes a value for [anterior]; the rule for [low] presupposes [back] and [consonantal], [labial] presupposes these two and [low], and [high] presupposes [labial], among others. Kean has devised the markedness conventions so the features fall into a partially ordered hierarchy, which she presents as in (13). In the diagram, StVC = stiff vocal cords; SlVC = slack vocal cords; CG = constricted glottis; CP = constricted pharynx; SG = spread glottis; and DR = delayed release.

(13) Feature hierarchy (Kean 1980: 26)



The markedness conventions thus presuppose a feature hierarchy. This hierarchy, however, functions quite differently from the contrastive hierarchy because it is not driven by language-particular contrasts, and does not distinguish between contrastive and redundant feature specifications in particular languages.[6]

6 One generative approach to markedness that did incorporate a contrastive hierarchy is the Markedness Theory of Syllable Structure (MTSS) proposed by Cairns (1988). This theory combines markedness, underspecification and a contrastive hierarchy, which Cairns calls a 'coding tree', to arrive at feature specifications. The MTSS determines contrasts narrowly by position, rather than over the language as a whole.

Consider, for example, the markedness matrices for vowels that emerge from the markedness conventions and implicational restrictions proposed by Kean, shown in (14). I omit features for which all vowels have the same specifications.

(14) Markedness matrices for vowels (Kean 1980: 31)

	i	e	æ	y	ø	œ	ɨ	ə	a	u	o	ɒ
back	m	m	m	m	m	m	u	u	u	u	u	u
low	u	u	m	u	u	m	m	m	u	m	m	u
lab	u	u	u	m	m	m	m	m	u	u	u	m
high	u	m	u	u	m	u	u	m	u	u	m	u
Complexity	1	2	2	2	3	3	2	3	0	1	2	1

If we assume that languages favour segments with low complexity values, it would follow from (14) that the least marked three-vowel inventory should contain /a/, the least marked vowel (with 0 marked specifications), and two of /i/, /u/ and /ɒ/. The latter would be an unexpected choice, since the most common three-vowel inventory is /i, a, u/, not /i, a, ɒ/ or, even less probably, /a, ɒ, u/. Whichever vowels we choose, there is no notion that some feature values are contrastive whereas others are not.

5.2.5 Linking to the phonology

In the *SPE*–Kean approach to markedness, the markedness statements are largely insulated from the phonology proper. Recall that the motivation for this was to enable the phonology to operate with fully specified +/− feature values, and to avoid bringing into the phonology any underspecified feature values, or features specified in terms of the *u*/*m* notation of markedness theory. The disadvantage of this strategy, however, is that the phonology cannot refer to either markedness or underspecification in cases where such reference would be useful in capturing generalizations about phonological processes. In effect, a theory in which the active phonological component is entirely insulated from the markedness component embodies the claim that markedness is relevant only to underlying phonological inventories and phonotactics, but ceases to play any role in the phonology proper.

For example, recall the pairs of rules in (1a, b). These are listed by Chomsky and Halle as being examples where the rule in (i) of each pair is formally more complex than that in (ii), though more natural and more expected. The markedness theory they propose has the effect of making the derived segment in the rule in (i) less marked than the one in (ii). But the rule component manipulates only +/− values, not *m*/*u* values. It remains to show how markedness can influence the operation of phonological rules.

Chomsky and Halle accomplish this by means of the device of *linking*. Essentially, when a phonological rule changes the value of a feature, any markedness conventions that mention this feature value are 'linked' to the rule, meaning that these markedness conventions can be brought into play. Linking the markedness conventions to phonological rules allows the phonology to refer to markedness in a limited way.

Chomsky and Halle state the linking convention formally as in (15).

(15) Linking convention (*SPE*: 420)
Suppose that the phonology contains the rule (i) and that there exists markedness convention (ii), where α, β = + or −, Y, Z, Q, W may be null, and the feature G is distinct from F.

$$\text{i.}\quad X \rightarrow [\alpha F] \;/\; Y \begin{bmatrix} \_\_ \\ Q \end{bmatrix} Z$$

$$\text{ii.}\quad [u\,G] \rightarrow [\beta G] \;/\; \begin{bmatrix} \_\_ \\ \alpha F \\ W \end{bmatrix}$$

If the segment to which (i) has applied meets the condition W of (ii), then the feature specification [βG] is automatically assigned to that segment.

For example, consider again the rules in (1a), whereby /i/ becomes [+back]. *SPE* would fully specify /i/ as in (16a).[7] Let us write the rule in question as in (16b). Changing the specification of [back] leads to (16c), which, if left undisturbed, would yield [ɨ], not [u]. However, there exists a markedness convention that mentions the specification [+back] and whose other specifications are also met by (16c), namely vowel convention XI, listed above as (2gi) and repeated here as (16d). Linking this convention to the rule means in effect that the value of the feature [round] in the segment produced by the backing rule must revert to its unmarked value, which is [+round] in the context of [+back, −low], yielding (16e), which is [u], not [ɨ].

7 Full specification is a relative matter, since there are further features that could apply here ([ATR], [tense], [nasal], etc.). Additional features will not affect the basic point.

(16) Linking in the change of /i/ to [+back]

a. Specification of /i/: [−low, +high, −back, −round]

b. Backing rule: $V \rightarrow [+\text{back}]\ /\ \begin{bmatrix} \overline{} \\ -\text{low} \end{bmatrix}$

c. Change by rule (b): [−low, +high, +back, −round]

d. Linked markedness convention (2gi):

$$\text{(XI) [u round]} \rightarrow [\alpha\text{round}]\ /\ \begin{bmatrix} \overline{} \\ \alpha\text{back} \\ -\text{low} \end{bmatrix}$$

e. Change by linking: [−low, +high, +back, +round]

Linking crucially depends on the feature hierarchy inherent in the marking conventions, for the only feature values that can be affected by linking are those that are ordered after the feature directly changed by the rule. This result would be expected given specification governed by the contrastive hierarchy. If a potentially contrastive feature, like [round] in the above example, is not in fact contrastive in a segment, it must be that some other contrastive feature that draws a related contrast, in this case [back], is ordered above it in the hierarchy. The noncontrastive features are left unspecified and take on their default values, given the specified (contrastive) features.

Moreover, Chomsky and Halle (*SPE*: 423) observe that for the phonology to benefit fully from linking, it must allow the successive application of all markedness conventions that apply to a given rule as well as to any markedness convention that links to that rule. That is, if a rule, R, changes a feature value to [αF], and a marking convention, C1, assigns [uG] to [βG] in the context of [αF], then C1 links to R. But then so, too, does marking convention, C2, that assigns [uH] to [γH] in the context of [βG] and other relevant features, and so on.

In other words, when a rule changes a feature value, all features ordered below it revert to their unmarked values, *as if* they had never been specified. In this way, linking effects a temporary despecification of subordinate features. Moreover, in the usual case, these features are also noncontrastive in the relevant domain. In this way, the linking convention mimics contrastive specification via a contrastive hierarchy.

As mentioned, however, the *SPE*–Kean approach to markedness does not really do the same work as the contrastive hierarchy, because it does not take account of language-particular contrasts. One disadvantage of this kind of approach to markedness is that it treats markedness as universal, whereas there is evidence that it is relative to a particular inventory. We will return to this subject in chapter 7, where I will argue that the ‘complexity’ of a segment is not fixed universally, but varies with the particular contrasts in play in an inventory.

As we will see, /ɨ/ is more marked than /i/ in the sense that the latter occurs more frequently in vowel inventories; but when both /i/ and /ɨ/ occur in the same inventory, it is /ɨ/ that functions like the less marked vowel. This kind of fact cannot be captured by a markedness scale like that in (14).

5.3 Contrast and underspecification

In the early 1980s the old arguments against underspecification began to lose their force, and underspecification was reintroduced to the theory. However, it did not come back in the form of a contrastive hierarchy. I will argue in what follows that, like the pre-generative phonologists before them, underspecification theorists associated with one of the two main theories of underspecification, Radical Underspecification, did not pay sufficient attention to the notion of contrast. This was a point also made by advocates of the other prominent approach to underspecification, Contrastive Specification; I will argue that the latter, however, did not pay sufficient attention to how contrasts are established. The lack of an adequate theory of contrast hindered the subsequent development of underspecification theory and left it vulnerable to critiques that underspecification has no principled basis and is empirically flawed.

5.3.1 Lexical Phonology and Structure Preservation

In a series of groundbreaking papers in the early 1980s, Paul Kiparsky was able to overcome Stanley's arguments and reintroduce underspecification into generative phonology. Kiparsky achieved this by means of a number of theoretical innovations. Kiparsky (1982) proposed that Stanley's objection against underspecification on the grounds that it allows for the ternary use of binary features can be answered if the phonology is restricted to specifying only one lexical feature value, usually the marked one, in any given context. Thus, if [+F] is specified, [–F] is barred, and vice versa. The result is that there is an underlying binary contrast between [+F] and 0 or [–F] and 0, but no ternary contrast between [+F], [–F] and 0 in any given context.

To complete the argument, it is also necessary to prevent the creation of a ternary contrast from arising in the course of the derivation. In the *SPE* conception of phonology this poses a problem, because unspecified values must be filled in at some point in order to arrive at a phonetic representation. Kiparsky drew together diverse streams of research in morphology and phonology into a theory he called Lexical Phonology, also known as Lexical Phonology and Morphology (LPM, Pesetsky 1979; Kiparsky 1982, 1985; Kaisse and Shaw 1985; Mohanan 1986). LPM posits that there is a fundamental distinction between *lexical* and *postlexical* phonology. Lexical phonology interacts with the morphology and the lexicon, and exhibits properties such as cyclic application, restriction to derived environments, and exceptions. Postlexical phonology follows the lexical phonology and does not observe the above restrictions, having rather properties one would associate with phonetic rules.[8]

Most important with respect to our topic, Kiparsky (1982) argued that the lexical phonology is the domain in which restrictions on feature specification hold. For example, Kiparsky (1985) observes that voicing in English is distinctive for obstruents but not for sonorants.[9] Further, sonorants do not trigger or undergo rules of voicing or devoicing in the lexical phonology of English, though they may do so in the postlexical phonology (e.g. *r* may be devoiced in words like *cry*).

Kiparsky (1985) proposes that these facts are connected. He suggests that there exists a marking condition in English that prohibits voicing from being marked on sonorants in the lexicon (17a). He proposes further to extend this prohibition throughout the lexical phonology, a constraint he calls *Structure Preservation* (17b).

(17) Underspecification and Structure Preservation (Kiparsky 1985: 92)
  a. Marking condition: *[αvoiced, +sonorant] in the lexicon.
  b. Structure Preservation: marking conditions such as (17a) must be applicable not only to underived lexical representations but also to derived lexical representations, including the output of word-level rules.

The marking condition in (17a), together with Structure Preservation (17b), accounts for the fact that English lexical voicing assimilation is triggered by and applies to obstruents, not sonorants. The marking condition (17a) is possible because voicing on sonorants is predictable.

The rationale for this proposal is that there is a deep connection between predictability of feature specifications and phonological behaviour: predictable features are not present in the phonology, and thus the phonology has no access

8 Dresher (1983) argues that Biblical Hebrew postlexical phonology (in the sense of phonology that takes place in a domain larger than the word – for example, the phonological phrase) exhibits properties more usually associated with lexical phonology. This observation does not affect the distinction between a phonological and phonetic component.

9 The example of [voiced] being predictable given [sonorant] is perhaps the oldest and most common example of underspecification in the literature (cf. Stanley 1967). Nevertheless, it may not be a good example if, as has been argued, sonorants do not have the same voicing feature as voiced obstruents (Piggott 1992, Rice 1993, Avery 1996, Boersma 1998). For purposes of this discussion, I will assume that sonorants do potentially bear a feature [voiced] that is also carried by voiced obstruents. What is crucial here is the logic of the argument, whether or not sonorant voicing is in fact a good exemplar of it.

to them. The fact that voicing rules do not need to explicitly exclude sonorants is thus a real generalization, not a ‘specious’ one, as Stanley (1967) claimed.

In connecting predictability to underspecification, and enforcing this underspecification throughout the lexical phonology, LPM comes close to incorporating the Contrastivist Hypothesis, because the features in play in the lexical phonology are contrastive, to a first approximation. However, as with many other theories that involve contrast, the basis for determining which features are contrastive was not clear, and this led to serious ambiguities.

Consider again the marking statement in (17a), which states that no value of [voiced] may be assigned to sonorants. We might think that this condition follows directly from the fact that voicing is predictable on sonorants. However, the matter is more complicated than that. Kiparsky (1985) contains an extended analysis of ‘the notorious problem of voicing assimilation in Russian’ (cf. the rule of RVA in section 4.6). The part of interest to the present discussion is his treatment of the Russian sonorants and the ‘unpaired’ obstruents /f, ts, tʃ, x/. Recall that both the sonorants and the unpaired obstruents have predictable voicing; but whereas the sonorants do not participate in RVA, the unpaired obstruents do. Kiparsky proposes the marking statements in (18) for the sonorants (a) and the non-coronal unpaired obstruents (b).

(18) Russian lexical conditions on marking (Kiparsky 1985: 108)

a. Marking condition on sonorants

$$* \begin{bmatrix} \alpha\text{voiced} \\ +\text{sonorant} \end{bmatrix}$$

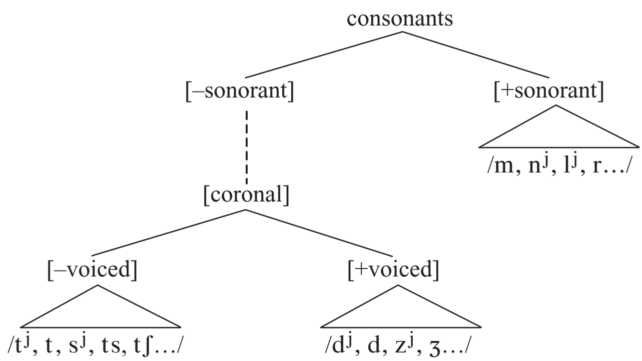
b. Marking condition on /f, x/

$$* \begin{bmatrix} +\text{voiced} \\ -\text{coronal} \\ +\text{continuant} \end{bmatrix}$$

The condition on sonorants (18a) is the same as for English sonorants (17a). But the condition on the obstruents is crucially different: rather than prohibiting them from bearing *any* specification for [voiced], it excludes only [+voiced]. Kiparsky (1985: 136 n. 13) notes that the condition is restricted in this way to allow the obstruents to function as [−voiced] segments in the lexicon. Kiparsky’s analysis thus mirrors that of Halle (1959) in some respects, while retaining something of the older distinction between the morphophonemic and phonemic components. As in Halle’s account, unpaired obstruents are unspecified in the lexicon for [voiced] because they are predictably voiceless, but are allowed to take on the value [−voiced] in time to participate in RVA. With the structuralists, however, Kiparsky proposes that their voiced allophones are prohibited from the lexical phonology, but are derived by a postlexical application of RVA.

The question remains, however, as to what principle explains the difference in the marking conditions in (18). It does not follow from predictability, because it is equally predictable that sonorants are voiced and that unpaired obstruents are voiceless. Therefore, something beyond predictability must be involved. A contrastive feature hierarchy provides a rationale: if the Russian feature hierarchy has [sonorant] ordered ahead of [voiced], and [voiced] is ordered ahead of the features that isolate the phonemes /f, ts, tʃ, x/ from voiced obstruents, then the two groups of consonants can be seen to stand in a very different relation to the feature [voiced], as shown in (19) for the sonorants and coronal obstruents.

(19) Russian consonants: partial feature hierarchy



The partial feature hierarchy in (19) shows that sonorants have no contrastive specification for [voiced]; the unpaired obstruents, however, are contrastively specified [−voiced], just like every other voiceless obstruent. These different scope relations provide a principled basis for distinguishing between the sonorants and unpaired obstruents.[10]

Kiparsky (1985: 135 n. 3) comes close to relating the acquisition of marking conditions to a feature hierarchy, commenting that 'universal grammar will constrain this process by a hierarchy of features which defines their accessibility

10 Structure Preservation remains an unclear principle, however, because it is not clear what 'structures' must be preserved. Given a contrastive hierarchy as in (19), one could interpret Structure Preservation to require that no combination of features may be formed that do not correspond to an underlying segment. On this interpretation, [+voiced] may not be assigned to /ts/ because there is no underlying /dz/. Structure Preservation can be interpreted more liberally, however, as prohibiting the assignment of a noncontrastive feature to a segment. Thus, no value of [voiced] may be assigned to a sonorant, but /ts/ would be able to receive [+voiced]. See, further, section 8.3 and the discussion of systematic and accidental gaps in section 8.6.3.

to marking (Jakobson 1941; Kean [1980])'. This idea was adopted by the markedness theory of Calabrese (see section 8.6); however, underspecification theory did not take the path to which this note points.

5.3.2 Radical Underspecification

The general approach to underspecification pioneered by Kiparsky was taken up and elaborated by Archangeli (1984) and Pulleyblank (1986) under the name Radical Underspecification (RU); see also Archangeli (1988), Abaglo and Archangeli (1989) and Archangeli and Pulleyblank (1989).

Radical Underspecification allows for two main types of rules. *Default rules* capture universal restrictions on feature co-occurrence, and are context dependent. Because voicing on sonorants is completely predictable in most languages, the feature [voiced] is introduced by a default rule of the form in (20):

(20) Default rule for sonorants
$[+\text{sonorant}] \rightarrow [+\text{voiced}]$

The second rule type proposed in RU, *complement rules*, are context-free rules that introduce the opposite feature value on a language-specific basis if a default rule is not available to supply the feature value. Suppose, for instance, that there are two series of stops in a language, voiced stops and voiceless stops. Voicing is clearly distinctive in this language. Following Kiparsky (1982), only a single value of a distinctive feature can be marked. Two possible markings are possible for any single feature, as in (21).

(21) Possible specifications for [voiced] in RU

	a. p	t	k	b	d	g	b. p	t	k	b	d	g
voiced				+	+	+	−	−	−			

Either the plus value of the feature is marked, as in (a), or the minus value of the feature is marked, as in (b), but not both. It is suggested that one of the markings (in this case (a)) will be preferred by universal grammar but that language-particular considerations can override this, yielding (b) as a second possibility.

5.3.2.1 Radical Underspecification and the Contrastivist Hypothesis

RU has one characteristic that accords with the Contrastivist Hypothesis, as opposed to the *SPE*–Kean markedness theory. RU relies to some extent on phonological activity in arriving at the choice of underlying specifications. Because patterns of activity differ across languages, RU therefore posits that languages can differ in their choice of underlying specifications. As evidence in establishing underlying specifications for vowel systems, Archangeli (1984) is guided by the identity of the epenthetic vowel, on the assumption that it is the least specified vowel. For example, Spanish, Japanese and Telugu all have five-vowel systems that can be represented by the phonemes /i, e, a, o, u/. In Spanish the epenthetic vowel is /e/, in Japanese it is /i/, and in Telugu it is /u/. To account for these different patterns, Archangeli (1984) proposes that the underlying representations of the vowels differ in each language.

RU does not fully adopt the Contrastivist Hypothesis, for it places severe limits on the extent to which aspects of phonological activity besides epenthesis are considered to be evidence for underlying specifications. In RU, the notion of contrast is not explicitly mentioned in the characterization of either default or complement rules, nor is there explicitly a contrastive hierarchy. Most central to the Contrastivist perspective, it can be shown that RU does not systematically distinguish between contrastive and redundant feature values.

Consider, for example, the vowel system of Spanish, which has the feature specifications shown in (22a). In Spanish the epenthetic vowel is /e/. One way of guaranteeing that /e/ is completely unspecified is to disallow specification of any values that /e/ has in its full specification. In this case, these values are all [–F] for every feature, F, in (22). A consequence is that these values cannot be specified on any other vowel, either. The result of removing these values is (22b). Each removal requires the formation of a complement rule (22c).[11]

(22) Spanish: /e/ is unspecified (Archangeli 1984)

a. Full specifications

i	e	a	o	u	
+	−	−	−	+	high
−	−	+	−	−	low
−	−	+	+	+	back
−	−	−	+	+	round

b. Removing values of /e/

	i	e	a	o	u
high	+				+
low			+		
back			+	+	+
round				+	+

c. Complement rules

i. [ ] → [−high] ii. [ ] → [−low]
iii. [ ] → [−back] iv. [ ] → [−round]

11 The complement rules in (22c) are all context free. There are other ways to ensure that /e/ is unspecified. One could, for example, adopt context-sensitive complement rules that more closely resemble the markedness rules discussed in the previous chapter (see, further, Mohanan 1991).

Since underlying representations consist essentially of privative features, the Distinctness Condition is not in effect. Once it has been decided that /e/ is the unspecified vowel and its features have been removed, the main principle governing the specification of the rest of the phonemes is minimality of specification. Looking at (22b), we see that the features [high] and [low] divide the vowels into three sets: /i, u/, /e, o/, and /a/. It remains to make a further division in each of /i, u/ and /e, o/. Either feature [back] or [round] will suffice to do this. Archangeli chooses to dispense with [back], presumably because /a/, which requires no specifications beyond [+low], has a specification for [back] but not for [round]; therefore, [round] results in fewer specifications, respecting minimality. This leaves us with the representations in (23a), the complement rules in (23b), and the default rules in (23c).[12]

(23) Spanish vowels: RU underlying features (Archangeli 1984: 58–9)

a. Underlying specifications

	i	e	a	o	u
high	+				+
low			+		
back					
round				+	+

b. Complement rules

i. [ ] → [−high] ii. [ ] → [−low]

iii. [ ] → [−round]

c. Default rules

DR1 [ ] → [+back, −round] / [________, +low]

DR2 [ ] → [αback] / [________, −low, αround]

It can be shown that the complement rules do not distinguish between contrastive and redundant values. Complement rules always fill in *all* values [–αF] of a feature for which [αF] is specified in underlying representation. By any definition of contrast, some of these complement values are contrastive, and some are redundant. Consider the complement rule that introduces the feature [−low]. To gauge whether particular values are contrastive, let us look at how the SDA would treat this feature. Since feature ordering makes a difference to the SDA, consider the two orderings in (24): in (24a), the complement value [−low] is contrastive for all of /i, e, o, u/; in (24b), it is contrastive only for /e/, and redundant for /i, o, u/.

12 Since both values of [back] are removed from underlying representations, Archangeli dispenses with the complement rule for [back] (22ciii), and uses instead the default rules in (23c).

(24) Spanish vowels: contrastive features by the SDA

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

	i	e	a	o	u
high	+	−		−	+
low	−	−	+	−	−
back					
round	−	−		+	+

b. Ordering: [high] > [round] > [low] > [back]

	i	e	a	o	u
high	+	−	−	−	+
low		−	+		
back					
round	−	−	−	+	+

Similar results can be shown to be true of the other complement rules.[13] If the distinction between contrastive and redundant is important in phonology, then the complement rules of RU fail to capture it.

With respect to phonological activity, there is no rationale for assigning a privileged status to epenthesis as an indication of the underlying unspecified vowel. In the case of Japanese, for example, Archangeli (1984) supposes that the epenthetic vowel is /i/, and therefore assigns this vowel no underlying specifications. By RU hypothesis, [−high], [+low] and [+back] must be specified in Japanese, as shown in (25).

(25) Japanese underlying specifications (Archangeli 1984: 59–60)

	i	e	a	o	u
high		−		−	
low			+		
back				+	+

We observed in section 3.3, however, that there are reasons, based on the way vowels affect consonants, to suppose that Japanese /i/ is specified [+front] (equivalently, [−back]), contrary to (25). Also contrary to (25), Hirayama (2003) argues that /o/ must be contrastively specified for a feature indicating lip rounding. Moreover, Hirayama (2003: 117–19) observes that *u* and *o*, as well as *i*, can appear as epenthetic vowels in Japanese; she argues that *u* is the default epenthetic vowel, with *o* and *i* occurring only in specific environments. By looking at a range of phenomena that attest to phonological activity, we can

13 The default rules in (23c) also do not distinguish contrastive from redundant values. DR2 assigns values of the feature [back], which are all redundant by any definition of contrast applicable to Archangeli's analysis. But DR1 also fills in [−round] on /a/; as is evident in (24), this specification may be contrastive (b) or redundant (a).

arrive at a better picture of how phonemes pattern than by limiting the relevant evidence to epenthesis.

5.3.2.2 The Redundancy-Rule Ordering Constraint

The underlying representations posited by RU are quite minimal, and so we might imagine that they represent strong claims about the phonology of a given language. This would be the case if unspecified features could not be filled in in the course of the phonology, or could only be filled in in circumscribed situations. If the phonology could be properly characterized with only a minimal set of feature markings, that would constitute strong empirical support for the analysis.

However, RU allows both + and − values to be filled in freely throughout the phonology. Although the preference is for unspecified features to be filled in as late as possible, they may be inserted whenever needed, in accord with the Redundancy-Rule Ordering Constraint (RROC).

(26) The Redundancy-Rule Ordering Constraint (Archangeli 1984: 85)
A redundancy rule assigning ‘α’ to F, where ‘α’ is ‘+’ or ‘–’, is automatically ordered prior to the first rule referring to [αF] in structural description.

The RROC guarantees that there will be no empirical evidence that can potentially disconfirm the choice of underlying feature specifications. Consider, for example, the underlying vowel system of Yowlumne Yokuts (formerly Yawelmani: Kuroda 1967, Newman 1944), a Penutian language of California. The underlying vowels can be represented as /i, u, a, o/, and their phonological patterning suggests that they are characterized in terms of the features [round] and [high], as in (27).

(27) Yowlumne vowels

[–round]	[+round]	
i	u	[+high]
a	o	[–high]

Because the epenthetic vowel is /i/, Archangeli chooses to specify [−high] and [+round] underlyingly, as in (28). The representations in (28) limit the underlying vowel feature specifications to [−high] and [+round].

(28) Yowlumne vowels (Archangeli 1984: 75–6)

a. Underlying specifications

	i	a	o	u
high		–	–	
round			+	+

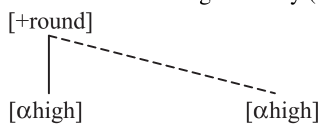
b. Complement rules

i. [ ] → [+high]

ii. [ ] → [−round]

Yowlumne has a rule of rounding harmony whereby an *i* changes to *u* when a *u* stands in the preceding syllable, and *a* changes to *o* following a preceding *o*. That is, the rule of rounding harmony is limited to affecting vowels in the same height class. Archangeli formulates it as a spreading rule as shown in (29).

(29) Yowlumne Rounding harmony (Archangeli 1984: 79)



The rule makes reference to the feature [αhigh]. According to the RROC, the reference to [αhigh] triggers all rules that fill in both values of [high], if they have not yet applied, before the spread of [round] can occur. Since rounding harmony is a fairly early rule, there are no empirical consequences of the decision to specify only [−high] in underlying forms. As far as the phonology is concerned, both values of [high] may as well have been specified.

5.3.2.3 Summary

Although Radical Underspecification takes phonological activity into account to the extent that it bases the choice of underlying specifications on the identity of the epenthetic vowel, this is the only aspect of phonological activity it looks at. Considerations of minimality trump any other criteria, and the RROC systematically removes phonological activity as a possible determinant of feature specification by insulating the theory from any possible negative consequences of underspecification. As D. C. Hall (2007: 24–6) puts it, Radical Underspecification predicts that features may be absent only if they are redundant, whereas the Contrastivist Hypothesis says that they may be present only if they are contrastive.

5.3.3 Contrastive Specification

Steriade (1987) proposes an alternative to Radical Underspecification called Contrastive Specification (CS) (see also Clements 1987 and Christdas 1988). Steriade argues that we must distinguish between two rule types. First, there are redundancy, or R-rules. These are similar to the default rules of RU, introducing a redundant value within a class for which that value is fully predictable. The familiar voicing on sonorants provides the prototypical example of an R-value introduced by an R-rule: the feature [voiced] is introduced on sonorants by a redundancy rule, as in (20). Sonorants are thus underlyingly unmarked for voicing.

The second class of rules, D-rules, introduces D-values, which are contrastive feature values that distinguish between segments. In a language with voicing contrasts among obstruents, the feature [voiced] is distinctive in the obstruent inventory, and so is a D-value for the obstruents. Steriade proposes that only R-values are underspecified in the sense of RU; contrary to RU, she proposes that both D-values (+ and –) are specified underlyingly. Thus, in a language with both voiced and voiceless stops, the voiced stops are specified as [+voiced] and the voiceless stops as [−voiced].

5.3.3.1 Contrastive Specification and the contrastive hierarchy

By making a basic distinction between contrastive and redundant values, Steriade's proposal puts the notion of contrast at the centre of underspecification theory. It thus becomes crucial to know whether a feature value is redundant or contrastive in any particular instance. Steriade does not supply an explicit mechanism for determining contrasts, but rather provides evidence for particular analyses that can be viewed as consistent with an asymmetry between the patterning of redundant and contrastive features. For example, in her analysis of the Pasiego dialect of Montañes, a language with a five-vowel system, Steriade (1987: 343) argues that /a/, which neither triggers nor blocks a rule of height harmony, has no marking for the feature [high]. This is because 'the impossibility of simultaneous [+high, +low] specifications establishes that the height of low vowels is a R-value' (p. 342). This analysis appears to be unproblematic, especially when the vowel system is diagrammed as in (30), where /a/ is the obvious odd man out with respect to the feature [high]:

(30) Pasiego (Steriade 1987)

i		u	[+high]	
				D-class of [high]
e	o		[−high]	
	a			R-class of [high]

As demonstrated earlier, however, these cases are deceptively simple-looking: the basis according to which D-values are determined is not self-evident. The value [−high] is logically redundant for /a/, as it is indeed retrievable from [+low]; however, we have seen that logical redundancy is not an adequate basis for establishing contrastive specifications. From the point of view of feature ordering, the values in (30) correspond to (31a), and require [low] > [high]. If [high] > [low], as in (31b), [high] would be a D-value for /a/.

(31) Pasiego vowels: possible orderings of [high] and [low]

	a. [low] > [high]					b. [high] > [low]				
	i	e	a	o	u	i	e	a	o	u
high	+	–		–	+	+	–	–	–	+
low	–	–	+	–	–		–	+	–	

Similarly, Steriade (1987) argues that the low vowel /a/ is unspecified for [back] in triangular five-vowel systems such as those of Ainu and Tamil. Again, though she does not state this explicitly, the analysis relies either on logical redundancy or on a contrastive feature hierarchy whereby [low] > [back].

Although Steriade (1987) makes no mention of the contrastive hierarchy as a general principle for determining whether feature values are contrastive or redundant, she does propose a hypothesis that translates into a constraint on possible hierarchies. Dividing features into *stricture features* (essentially manner features for consonants and height features for vowels) and *content features* (place features for consonants and timbre features for vowels), she proposes that, in a redundancy rule [αF] → [βG], F may be a stricture feature and G may be a content feature, but not vice versa. Thus, we may have a redundancy rule [+low] → [+back], but not [+round] → [+high]. This is tantamount to claiming that stricture features must always have wider contrastive scope than content features; that is, they are ordered higher in the contrastive hierarchy.

5.3.3.2 Pairwise Algorithm for Contrastive Specification

Given that the distinction between R-values and D-values is central to the theory of CS, it is noteworthy that neither Steriade (1987) nor any other advocate of CS proposed an algorithm to determine which values are R-values and which are D-values. Lacking an algorithm for CS from its proponents, one was suggested by an opponent of the theory.

Archangeli (1988) proposes an algorithm for determining contrasting classes, namely the Pairwise Algorithm discussed in chapter 2. As we have seen, the Pairwise Algorithm instantiates a particular approach to picking out contrastive specifications and cannot be identified with contrastive specification itself. Nevertheless, the Pairwise Algorithm became identified with CS theory, and the many shortcomings of that algorithm became shortcomings of CS.

5.3.4 *Backlash against theories of underspecification*

The lack of an adequate mechanism for deciding which features to omit ultimately led to a backlash against theories of underspecification. Critics such as Mohanan (1991) and Steriade (1995) pointed out various inconsistencies and apparently insoluble conundrums in the practice of underspecification theorists. Mohanan (1991: 306–7) argues that, if the goal of underspecification theories is to remove all predictable information from underlying representations, then this goal is unachievable because of the existence of mutual dependencies.

This is essentially the argument I made in chapter 2 against any procedure that is designed to remove all logical redundancies from specifications. Underspecification that results from contrastive specification by the SDA is not susceptible to this argument, because it is not intended to remove all logical redundancies, nor does it equate system redundancy with predictability.

Steriade (1995: 118–19) observes that a constraint of the form *[+sonorant, −voice]:

> renders predictable not only the voicing of sonorants but also the sonority of voiceless segments. Must we then leave the [–sonorant] value of *p*, *t*, *k* out of the underlying representations? There seems to be little evidence for such a move and we need to ask why, especially as there exists substantial evidence for leaving the sonorants unspecified for [voice] (Kiparsky 1985; Itô and Mester 1986).

Similarly, she observes that either [back] or [round], but not both, can be omitted from the specifications of the non-low vowels in triangular systems such as /i, e, a, o, u/; the choice of which to omit thus appears to her to be arbitrary.

As we have seen, this type of problem is solved by the contrastive hierarchy in conjunction with the Contrastivist Hypothesis. The feature [sonorant] is a major class feature that is typically higher in the hierarchy than [voiced]; in the typical case where all sonorants are voiced, [voiced] is not contrastive in the [+sonorant] set, but is in the [−sonorant] set. The fact that [−voiced] segments are predictably [−sonorant] plays no role in their specification, according to the SDA. To have a situation where [sonorant] is omitted from voiceless consonants, we would need a feature hierarchy with [voiced] > [sonorant]. Such a hierarchy is relatively rare.[14] Similarly, we have seen numerous cases where the choice of [back] or [round] is not arbitrary, but dictated by considerations of phonological patterning and activity. As in other areas of linguistic theory, there may be

14 It may be premature to rule out such systems, however; compare Hockett's analysis of Lifu discussed in section 3.6. It is noteworthy that Lifu has voiceless sonorants, in his analysis. There also appear to be languages where the feature [sonorant] does not play a major role at all; compare the dialect of Inuktitut discussed in section 7.3.

situations in which the feature order is not clear, for lack of decisive evidence pointing to one or another ordering. But such indeterminacy arises for empirical reasons, not because the theory lacks coherence or is inherently arbitrary.

In chapter 7 I will pursue these issues in the framework of Modified Contrastive Specification, a theory that developed partly in response to the shortcomings of these theories of underspecification.

5.4 Theories of feature organization

In addition to markedness theory and underspecification theory, generative phonology produced a third family of subtheories that crucially involve hierarchical relations between distinctive features. These theories posit that features are organized into various kinds of structures. This idea was developed in a number of quite different ways, giving rise to distinct theories of phonological representation. Some examples include what has come to be known as feature geometry (Clements 1985; Sagey 1986; McCarthy 1988; Clements and Hume 1995; Halle 1995), Government Phonology (Kaye, Lowenstamm and Vergnaud 1985), Dependency Phonology (J. M. Anderson and Ewen 1987) and Radical CV Phonology (van der Hulst 1995, 1996, 2005). It is not possible to discuss all these theories here. Rather, I will focus on how a theory of feature organization might be related to the contrastive hierarchy.

5.4.1 Feature geometry

What has come to be known as 'feature geometry' developed as an extension and generalization of the theory of Autosegmental Phonology (Leben 1973; Goldsmith 1976), which starts with the observation that certain features, notably tone, are relatively autonomous of the other segmental features. In their review and synthesis of a number of approaches to feature geometry, Clements and Hume (1995: 245–6) write, 'Earlier theoreticians tended to think of phonemes as unstructured sets of features, or "feature bundles" in Bloomfield's well-known characterization. In accordance with this view, later work in the Jakobsonian and generative traditions treated segments as feature columns with no internal structure.'[15]

One obvious similarity between feature geometry and the contrastive hierarchy is that they both encode hierarchical relations between features. In (32) I present the geometry proposed by Clements and Hume.[16] These trees contain two kinds of nodes. Terminal nodes in square brackets, such as [nasal],

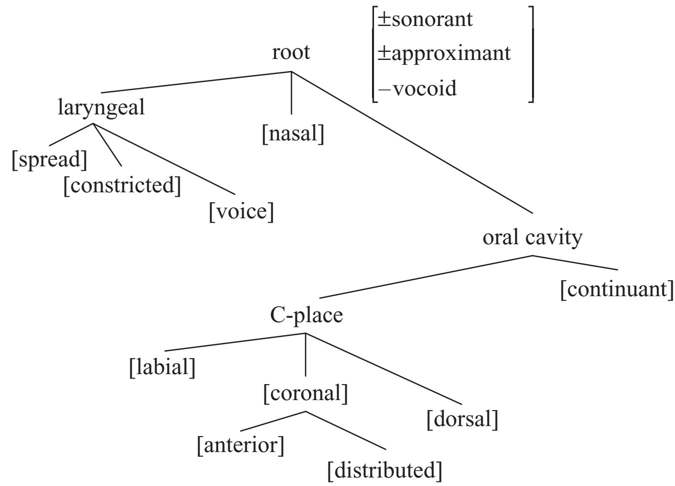
15 As we have seen, this summary is not entirely correct. Jakobson and early generative phonology did experiment with a hierarchical approach to features in terms of the contrastive hierarchy.

16 The feature [vocoid] is the converse of [consonantal], and the class of vocoids includes vowels and glides.

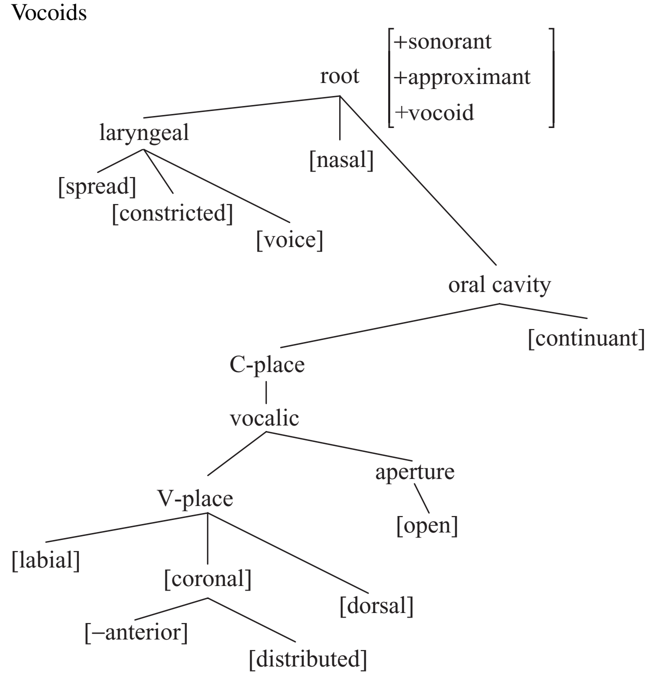
[continuant] and [labial], are features. The other nodes, such as laryngeal, oral cavity and C-place, are class labels that function as organizing nodes, grouping together the features below them. The root node receives special treatment, as it not only is considered to be an organizing node, but also bears the major class features [sonorant], [approximant] and [vocoid].

(32) Feature geometry (Clements and Hume 1995: 292)

a. Consonants



b. Vocoids

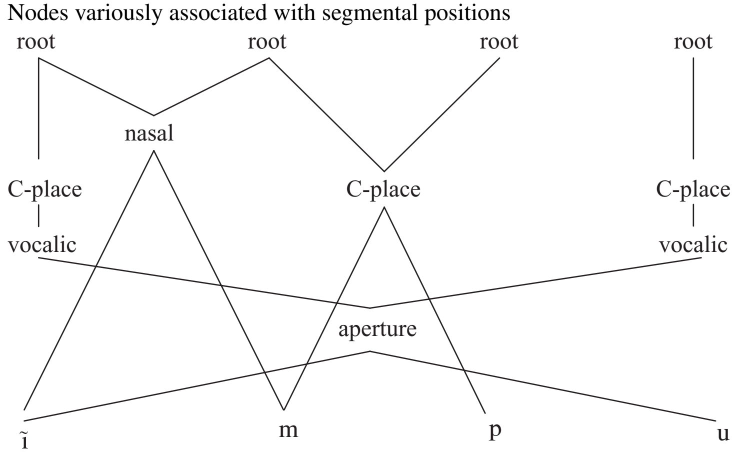


There are a number of differences between the trees in (32) and contrastive branching trees. Most obvious is the presence of class nodes in addition to feature nodes. The motivation for the class nodes is to group together features that function together in the phonology of particular languages. For example, many languages have rules assimilating the place of articulation of neighbouring consonants. In standard binary feature systems, such assimilation can be expressed only by a conjunction of all the features that characterize place of articulation (labial, coronal, anterior, etc.), or else by an ad hoc notation such as [α place of articulation]. This notation is given substance by positing a class node that dominates place of articulation (C-place, in (32)). Similarly, the laryngeal node groups together the laryngeal features that govern processes such as voicing, aspiration, glottalization, and so on, which can pattern together in some languages.

In standard versions of nonlinear phonology, of which feature geometry is a part, class nodes serve not just as abstract superordinate sets, but as real nodes that can spread and be delinked from skeletal positions. In (32), each node is on a separate line, which represents the fact that each node defines its own independent tier.

For example, assimilation in nonlinear phonology is viewed as spreading of a node from a trigger segment (or position) to one or more target segments. Features may be linked to more than one segmental position, either through spreading, or underlyingly. The result is that different nodes may be associated with different stretches of segmental material, as in (33) (not all nodes are shown).

(33) Nodes variously associated with segmental positions



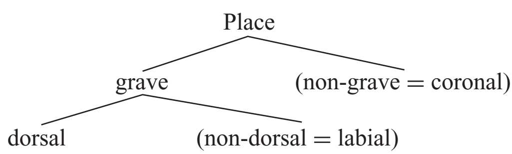
5.4.2 *Terminal features and the contrastive hierarchy*

Let us first consider the terminal features. We observe that there are a number of ways in which their treatment departs from a contrastive hierarchy. However, these characteristics for the most part represent empirical assumptions about feature organization that can be easily modified so as to bring about a rapprochement between feature geometry and a contrastive hierarchy.

First, the major class features tied to the root node, [sonorant], [approximant] and [vocoid], are treated as a bundle, and not hierarchically. These features, however, can be unpacked to reveal internal hierarchical structure. Conversely, the contrastive hierarchy could reflect the organization in (32) by allowing these features to cross-classify at the top of the feature hierarchy.[17]

Second, the tree in (32) is not strictly binary, since the laryngeal and the place nodes have three branches each. We have observed that there is no logical requirement that the contrastive hierarchy be binary. Therefore, the contrastive hierarchy can mirror *n*-ary splits by treating the daughter nodes as different values of a single *n*-ary feature. Conversely, the ternary branchings in (32) can be converted to a series of binary splits. Avery and Rice (1989), for example, propose to revive the Jakobsonian feature [grave], which groups together labial and dorsal sounds (also sometimes called [peripheral]). Thus, the Place node for consonantal place expands in binary fashion as in (34).

(34) Binary expansion of Place node (Avery and Rice 1989)



Avery and Rice assume, in addition, that features are single-valued, reflecting markedness relations. Thus, dorsal consonants, the most marked in this structure, are specified [Place, grave, dorsal]; labials are [Place, grave], reflecting the assumption that they are unmarked among the grave consonants; and coronals, as the least marked consonants with Place, are marked only [Place]. Avery and Rice support this hierarchy of [grave] markedness by pointing to assimilation patterns in Korean, where the labials /m/ and /p/ can assimilate to a following dorsal consonant, but dorsals do not assimilate to labials.[18]

17 Cross-classification is equivalent to treating each feature as if it had wide scope independently of the feature(s) it cross-classifies.

18 In subsequent work Avery and Rice do not hold that these markedness relations are necessarily universal; see Avery and Rice (2004), Rice (2007, forthcoming), and section 7.2 below.

A third difference between the feature geometry in (32) and the contrastive hierarchy is that the former is intended to be universal, and does not allow for language-particular variation. But this difference can also be easily bridged. Feature geometry can be made to allow for language-particular variation (cf. Piggott 1992), just as the contrastive hierarchy can be constrained to require universal ordering, to whatever degree is required in either direction.

A fourth difference is that the various nodes of the feature geometry, in some accounts, are assumed to be fixed and independent of considerations of contrast. This point is non-negotiable from the point of view of contrastive hierarchy theory. A hierarchy that does not encode contrasts is no longer a contrastive hierarchy. However, there is no reason why feature geometry cannot adopt contrast as a requirement for expanding nodes. Indeed, in his seminal paper on feature geometry, Clements (1985) assumes that only contrastive features are specified in feature-geometric lexical representations. Avery and Rice tie feature geometry to contrast (Avery and Rice 1989, Rice 1993, Rice and Avery 1993, Avery 1996). Such an approach is taken implicitly by feature geometry theorists who simply ignore aspects of the feature geometry that do not pertain to the language they are working on.

5.4.3 Class nodes and the contrastive hierarchy

This leaves the class nodes as the major difference between feature geometry and the contrastive hierarchy. To the extent that class node theory is a true empirical property of feature organization, then it must be incorporated into the contrastive hierarchy, and this has been done by Avery and Rice. Incorporating class nodes, or indeed any inherent feature dependencies, leads to certain complications in the procedure of contrastive specification.

In its simplest version, the contrastive hierarchy contains a feature if and only if it is contrastive. When built-in dependencies are added in, this relation between specification and contrast no longer holds. In particular, suppose that a feature, F, is a dependent of a class node, C. This means that F cannot exist without C. Now it could be the case that C is not contrastive, but F is. In that case, F must be specified, a specification that in turn requires specification of the noncontrastive C.

This situation arises commonly. One recurring example concerns the feature [coronal]. In (32), as well as many other versions of feature geometry, [coronal] functions both as a feature and as a class node, as it has dependents, [anterior] and [distributed] in (32). An example is Sanskrit, which has a contrast between plain and retroflex coronals. Assuming that the latter have some marked feature that is a dependent of [coronal], say [−anterior], their representation must include [coronal], as in (35a).

(35) Representation of coronal contrasts

```
a. Retroflex        b. Plain (minimal)     c. Plain (NAC)
     Place                Place                 Place
       |                                          |
   [coronal]                                  [coronal]
       |
  [−anterior]
```

But how should the plain coronal consonants be represented? If we assume that [coronal] is the unmarked place for consonants, there is no need to specify [coronal] for such consonants, and they could be represented like coronal consonants in languages without internal coronal contrasts, as simply having a Place node, as in (35b).

Taking (35b) together with (35a), however, it appears that there are more differences between plain and retroflex consonants than there are. Intuitively, the two segments differ only with respect to coronal-internal specifications; but now they appear to differ also with respect to specification for [coronal] itself. In effect, a single contrastive difference – [−anterior] versus no specification for anterior – has turned into two differences. Moreover, there is no actual contrast between the two segments as regards their coronality, so it is odd that one should be specified for [coronal] while the other is not.

Avery and Rice (1989: 183) propose the Node Activation Condition (NAC) for such situations. The NAC, given in (36), requires that Sanskrit plain coronals be represented with a coronal node, as in (35c).

(36) Node Activation Condition (Avery and Rice 1989)
If a secondary content node is the sole distinguishing feature between two segments, then the primary feature is activated for the segments distinguished. Active nodes must be present in underlying representation.

Though the NAC reduces the difference between plain and retroflex consonants to the minimal feature by which they contrast, it does so at the cost of introducing noncontrastive features and nodes into otherwise contrastive representations. In this way feature geometry introduces ‘contrast from below’: a feature (or node) is introduced not for contrastive purposes (contrast from above), but to serve dependent nodes that require its presence. Whether the NAC is required is ultimately an empirical question. The Contrastivist Hypothesis can accommodate such ‘extra’ specifications if they are warranted.

But developments in feature geometry may make devices like the NAC unnecessary.

5.4.4 Spreading of class nodes?

Our comparison of feature geometry and the contrastive hierarchy reveals that, for the most part, the two either encode similar relations or else can be reconciled so that they do so. This raises the prospect that the two can be identified; put differently, that we can dispense with feature geometry as an independent theoretical entity in favour of the contrastive hierarchy. The major obstacle in the way of doing this is the participation of class nodes in phonological rules. If a class node is targeted by phonological processes, then it cannot be dispensed with, because of the risk of losing the ability to formulate phonological rules that express significant generalizations.

It is thus interesting that some more recent approaches to feature geometry no longer allow class nodes to spread. For example, the Revised Articulator Theory (RAT) (Halle 1995; Halle, Vaux and Wolfe 2000) requires that terminal features spread separately. Further, designated articulators are indicated by features, not by nodes in the geometry. Similarly, Padgett (2002) proposes that feature classes like Place, Colour, Laryngeal, etc., are not nodes in a structure, but features of features, or sets of features. Phonological processes or constraints can mention these classes, but apply directly to the features that make them up. Padgett argues that this approach preserves the advantages feature geometry was supposed to have in being able to refer to certain groups of features, while escaping some of its negative consequences.

If class nodes are not required to participate in phonological operations, then there are no compelling reasons why they are required in phonological representations beyond the needs of contrast. If this is correct, then we can do away with noncontrastive 'specification from below' and retain only phonological features motivated by contrast.

It remains an empirical question whether feature geometry may indeed be subsumed into the contrastive hierarchy, or whether certain dependencies between features must be represented independently of requirements of contrast. In what follows I will assume the former; that is, I will assume that the only hierarchical relations among features that need to be represented are those dictated by the contrastive hierarchy.

5.5 Conclusions

Though early pre-*SPE* generative phonology was characterized by extensive use of underspecification, a heritage from the work of Jakobson and Halle, the 'classical' theory initiated by *SPE* barred underspecification from the phonology. I have tried to show that the distinction between contrastive and non-contrastive feature specifications kept reappearing, sometimes disguised as something else, as did the notion of a feature hierarchy. I have also argued that markedness theory, underspecification theory and feature geometry all have affinities with the Contrastivist Hypothesis and the contrastive hierarchy, but also important differences. I have also contended that all these subtheories suffer from the lack of a clear understanding of how to assign contrastive features.

The result was that, by the mid 1990s, both versions of underspecification theory discussed here, RU and CS, had been largely repudiated by their earlier proponents. Research in the other subtheories continued, but the main excitement in phonological theory came from the advent of Optimality Theory, which turned much thinking about phonology upside-down. How contrast looks in Optimality Theory is the subject of the next chapter.

6 *Contrast in Optimality Theory*

6.1 Introduction

Optimality Theory (OT; Prince and Smolensky 2004) is a radical departure from the derivational model of previous versions of generative phonology. Any new theory puts old questions into a new light, and it is illuminating to explore the relationship between OT and the Contrastivist Hypothesis. I will show that the contrastive hierarchy, being a static set of conditions, lends itself very easily to formulation in terms of a set of constraints, and hence to OT. At the same time, I will argue that OT is not itself a theory of contrast, but is capable of instantiating a wide range of such theories.

The relationship between OT and the Contrastivist Hypothesis is more complex. It appears that the insights of a contrastivist approach can best be captured in a serial (derivational) model of OT in which some of the restrictions of the standard parallel version are relaxed.

I will begin with a brief review of the essentials of OT (section 6.2), and then I will present some early treatments of contrast within OT (section 6.3). I will consider how the contrastive hierarchy might be incorporated into OT in section 6.4. Related issues are discussed in sections 6.5–6.7. Dispersion-theoretic approaches to contrast developed within OT are considered in section 8.4.

6.2 Optimality Theory

OT (Prince and Smolensky 2004, first published 1993) is a theory of constraint interaction that posits violable ordered constraints. For example, to account for the universal preference for syllables to have onsets and to avoid codas, Prince and Smolensky posit the constraints in (1).

(1) Syllable preference (markedness) constraints
- a. ONSET: 'A syllable must have an onset.'
- b. NOCODA: 'A syllable must not have a coda.'

The constraints in (1) can be broadly called markedness constraints, for they jointly amount to a markedness scale for syllable types: CV is the least marked syllable, in that it obeys both constraints; CVC and V each violate one of the constraints; and VC is most marked because it violates both constraints.

An underlying syllable CVC (say, /bak/) violates NoCoda; therefore, it is possible that it will be 'repaired' so as to conform to the constraint. Whether or not this happens, and the nature of the repair, depend on the ranking of the constraints. If NoCoda is ranked relatively high, it will have to be satisfied, meaning that underlying /bak/ will not be permitted to surface as [bak]. The nature of the repair depends on the ranking of other constraints, two of which are shown in (2).

(2) 'Faithfulness' constraints
  a. Max C: 'Preserve an input consonant.'
  b. Dep V: 'A vowel in the output must be present in the input.'

The constraints in (2) belong to the class of 'faithfulness' constraints because they regulate the relation between input and output forms. Constraints of the Max type require that input material be preserved in the output; Dep constraints penalize the insertion (epenthesis) of material not in the input.

Rather than applying constraints in sequence to an underlying form, as in classical generative phonology, OT posits that the optimal output form corresponding to a given input form is selected by an evaluation of competing candidates. The optimal candidate is the one that best satisfies the set of ranked constraints. An example is given in (3).

(3) Constraint evaluation tableau

/bak/	Onset	NoCoda	Max C	Dep V
a. bak		*!		
b. ☞ bake				*
c. ba			*!	
d. ak	*!	*	*	

The tableau in (3) represents a grammar in which the constraints are ranked in the order Onset ≫ NoCoda ≫ Max C ≫ Dep V. The input (underlying) form is /bak/, and the forms in rows (a)–(d) are candidate output forms. Candidate (a) is the 'faithful' candidate, matching the input form most closely. However, it violates NoCoda; as there are two other candidates that do not violate this, or any higher-ranking constraint, this violation is fatal, indicated by !. Candidates (b) and (c) represent two different repair strategies: in (b) a vowel is inserted to convert the coda into an onset, and in (c) the coda consonant is deleted. In the grammar in (3), MAX C is ranked above DEP V; therefore, this grammar prefers the preservation of the underlying consonant to its deletion, and candidate (b) is optimal, indicated by the pointing hand. Candidate (d) is eliminated immediately because it violates the highest-ranking constraint. Shading indicates that a cell plays no role in the selection of the optimal candidate. For example, once candidate (d) fatally violates ONSET, whether it satisfies or violates subsequent constraints has no bearing on the outcome of the evaluation.

In standard OT, there is no cumulative calculation of violations: if a candidate violates the highest-ranking constraint and other candidates do not, it is a loser even if it does better with respect to all the remaining constraints.[1]

Different constraint rankings yield different optimal outputs. In (4), the faithfulness constraints are in reverse order from (3), and sandwiched between the markedness constraints, also in reverse order from the previous example.

(4) Constraint evaluation tableau

/bak/	NOCODA	DEP V	MAX C	ONSET
a. bak	*!			
b. bake		*!		
c. ☞ ba			*	
d. ak	*!		*	*

The optimal candidate in this grammar is (c). Candidate (a) will become optimal in a grammar that ranks the faithfulness constraints over the two markedness constraints. No ranking of these constraints will make (d) a winner, because it violates a superset of the constraints violated by each of candidates (a) and (c). OT picks the candidate that best satisfies a particular set of ranked constraints, not a perfect candidate.

6.3 Early treatment of contrast in OT

6.3.1 Licensing and underspecification (Itô, Mester and Padgett 1995)

Itô, Mester and Padgett (1995) propose that the inertness of redundant features can be accounted for in OT by deriving underspecification not as an input

1 Nothing prevents OT from adopting other forms of constraint evaluation, including stochastic rankings, or weighted constraints that are added cumulatively, and so on. These issues are largely orthogonal to the general relation between OT and contrast.

property, but as an output property. Building on a proposal of Padgett (1991: 56–8), they link redundancy to underspecification via a theory of licensing. For every feature F, there is a constraint of the form (5).

(5) Licensing (Itô, Mester and Padgett 1995)
LICENSE (F): 'The phonological feature F must be licensed.'

Features can be licensed by other features. However, there is a restriction on licensing, called Licensing Cancellation (6).

(6) Licensing Cancellation (Itô, Mester and Padgett 1995)
Licensing Cancellation: If $F \supset G$, then $\neg(F\lambda G)$.
'If the specification [F] implies the specification [G], then it is not the case that [F] licenses [G].'

In Itô, Mester and Padgett's example, the features [sonorant] and [nasal] each imply [voiced] (7); that is, [voiced] is redundant in a sonorant segment and in a nasal segment. It follows that [voiced] is not licensed by [sonorant]. Other things being equal, this implies that a sonorant will be unspecified for [voiced] at the surface.

(7) Some redundancy rules
a. SONVOI: [sonorant] ⊃ [voiced]
b. NASVOI: [nasal] ⊃ [voiced]

Itô, Mester and Padgett's approach implies a hierarchy of features in which [sonorant] and [nasal] are higher than [voiced]. Without such a hierarchy, the theory runs into the problems with logical redundancy that we discussed earlier. Thus, it is also true in most languages that a nasal is predictably sonorant. Does this mean that feature cancellation requires that nasal segments not be specified for [sonorant]? Presumably not, but one would have to explain why. In a contrastive hierarchy approach, it is because [sonorant] generally has wider contrastive scope than [nasal]; thus, nasal segments are a subtype of sonorants, whereas sonorants are not a subtype of nasals. Similarly, we need to know why we do not cancel the licence of Russian affricates to be marked for voice because coronal affricate implies voicelessness. So we still need a theory to tell us when a crucial redundancy exists, just to make licensing cancellation well defined. The contrastive hierarchy is one such theory.

Let us now turn to the relation between contrast and phonological activity. Itô, Mester and Padgett write (1995: 608) that the ranking of licensing constraints and redundancy rules like SonVoi with respect to each other tips the scales in favour of either specification or underspecification. That is, if we rank Licensing above Redundancy, then a sonorant (in this case /m/) will be unspecified for voice (8a). If we rank the other way, with SonVoi above Licensing, then a sonorant will be forced to have a voice specification (8b).

(8) Ranking and underspecification

a. License ≫ Redundancy

Candidate	License	SonVoi
m | [voiced]	*!	
☞ m		*

b. Redundancy ≫ License

Candidate	SonVoi	License
☞ m | [voiced]		*
m	*!	

While this is no doubt true, the conclusion is that this theory makes no predictions about whether a redundant feature will or will not be specified in any given case. Like Radical Underspecification, it predicts only that a feature may be absent if it is redundant. If underspecification is indeed essentially arbitrary, then this is the best we can do. But we could have done the same in a derivational theory as well: we can simply stipulate the underlying specifications of each phoneme, allowing some to be more fully specified and others to be underspecified as required.

6.3.2 Contrast as emergent

Kirchner (1997) diagrams a standard representational model of contrastive specification as in (9).[2]

2 This model is weaker than the Contrastivist Hypothesis, in that it allows some noncontrastive features to be added in the phonological component; it remains an empirical question whether

(9) Representational contrastive specification (Kirchner 1997: 84)

Underlying representation	*(ideally) pure representation of contrast*
↓	
Phonological component	*non-contrastive properties may be filled in, particularly if contrastive in other languages*
↓	
Phonetic component	*remaining non-contrastive phonetic properties, including gradient values, filled in*
↓	
Phonetic representation	*representation of all speaker-controlled phonetic properties of the utterance*

Kirchner (1997: 83–4) cites Steriade (1995) as observing that 'the assumptions of this model have often been disregarded in practice' because phonologists have not been consistent in which redundant features they choose to remove. We have already discussed this objection to underspecification theory in the previous chapter. As we have seen, the objection applies to theories based on logical redundancy, but not to the contrastive hierarchy approach to contrastive specification.

Like Itô, Mester and Padgett (1995), Kirchner (1997) considers it an advantage of OT that it can model varying degrees of contrastive specification. If required, the constraint hierarchy can simulate the specification of only contrastive features, or of all features, or of features that are contrastive only in some contexts, and so on.

As has been stressed throughout, the degree of underspecification of phonological representations is an empirical matter. While it is true that various degrees of (under)specification can 'emerge' from different types of constraint rankings, allowing arbitrary orderings is tantamount to adopting the null hypothesis about contrastive specification: namely, that the degree of specification is arbitrary, and free to vary without limit from language to language.

Nothing compels OT to adopt this, or any other, position on contrastive specification. In the following section I will consider how the contrastive hierarchy can be represented in OT.

the hypothesis has to be weakened to allow something like (9), a question I consider further in section 7.9.

6.4 The contrastive hierarchy in OT

Since OT requires a set of interacting constraints that mediate between an input and an output, we must make some decisions concerning what the input and output are when it comes to modelling the contrastive hierarchy.[3]

6.4.1 Outputs

Let us first consider the output. I will continue to assume that the contrastive hierarchy specifies only contrastive features, and does not specify redundant features. For example, given a three-vowel inventory /i, a, u/ and the ordered binary features [high] > [round], the SDA first divides the inventory into [–high] and [+high] sets. The [–high] set consists only of /a/, so the assignment of features to this phoneme stops at this point. The specification of further phonetic features that characterize /a/, such as [–round], [+back], [+low], and so on, is left to a later component.[4] The [+high] set, consisting of /i, u/, requires one further division, effected by [round] in this hierarchy, and other specifications are left for later. That is, the output of the SDA for the inventory /i, a, u/ consists of the contrastive specifications in (10).

(10) Output specifications of the SDA for /i, a, u/, [high] > [round]

$$/\text{i}/ = \begin{bmatrix} +\text{high} \\ -\text{round} \end{bmatrix} \qquad /\text{a}/ = \begin{bmatrix} -\text{high} \end{bmatrix} \qquad /\text{u}/ = \begin{bmatrix} +\text{high} \\ +\text{round} \end{bmatrix}$$

The standard theory of OT posits a single parallel mapping from lexical representation to surface phonetic form. Such a mapping would not mirror the effect of the SDA, but would include also redundant features, as well as the effects of phonological processes. For example, we will see in chapter 7 that the Classical Manchu vowel /u/ is phonetically [round] but does not bear a [round] feature phonologically. Moreover, in many contexts surface [u] represents two different underlying vowels: [ATR] (= Advanced Tongue Root) /u/, and non-ATR /ʊ/. The same is true of Nez Perce surface /i/, which represents the merger of two different underlying phonemes.

Thus, there are a number of different sources that lead to a differentiation between the output of the SDA and the surface phonetic form. In standard OT

3 This section has been influenced by discussions with Daniel Currie Hall and Sara Mackenzie, and by studying Mackenzie (2002) and D. C. Hall (2007).

4 In this implementation of the contrastive hierarchy in OT, I will assume that redundant features are not present at all in the phonology. Other interpretations of the status of redundant features are possible; see, further, section 7.9.

there is no level that corresponds to the representations that are the output of the SDA. Rather, the SDA fits best into a multi-stratal or serial version of OT (cf. Kiparsky 2000, 2002, forthcoming; Rubach 2000, 2003; and Bermúdez-Otero 2003).[5]

6.4.2 Inputs

What should be the input to the OT constraint system? The SDA specifies the set of well-formed contrastive representations for a given inventory. Therefore, an OT implementation of the algorithm should be able, at a minimum, to take fully specified representations as input, and output the corresponding contrastive specifications. Thus, the constraint system acts as a *filter,* sifting fully specified representations and allowing only the contrastive specifications to pass through.

OT theorists have proposed that the constraint evaluation system should be able to give an output for *any* input; this is Prince and Smolensky's (2004) notion of 'richness of the base'. In our case, there are two types of potential additional inputs that could be distinguished. First, there are *illicit* combinations of features, corresponding to phonemes that are not in the inventory. Presenting the grammar with such an input is tantamount to asking what a speaker would do when presented with a segment from another language, or how learners would interpret unfamiliar sounds in terms of their grammar. For example, how does the grammar corresponding to (10) treat an input /o/, that is, a set of features that includes the specifications [–high, +round]? We will consider this issue in section 6.6.

A second type of additional input consists of inputs that are *underspecified* relative to the representations licensed by the contrastive hierarchy. For example, in (10), any segment specified for [round] must also be specified [+high]. What should the OT translation of the contrastive hierarchy do with an input that consists only of the feature [–round]? I will consider this type of case in section 6.7.

In the first instance, however, we will limit inputs to fully specified but legal phonemes; that is, we will first limit the OT constraint evaluation system to the filtering function. It will take fully specified phonemes as input and produce contrastive specifications as output.

5 The claim that OT posits a single mapping from lexical to phonetic representations is often disregarded in practice. We have seen, for example, that Itô, Mester and Padgett (1995) posit outputs that may have a certain amount of underspecification. Presumably, these are intermediate representations that are subject to a further round of specification.

6.4.3 An OT version of the Successive Division Algorithm

For the sake of this discussion, let us take as an example the Classical Manchu vowel system, shown in (11). The line divides vowels into two height classes.

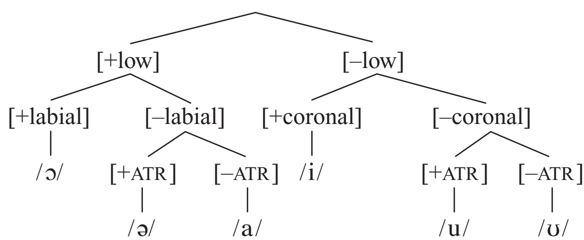
(11) Classical Manchu (Zhang 1996)

i		u
		ʊ
	ə	
		ɔ
	a	

Zhang (1996) proposes that the contrastive feature hierarchy for Classical Manchu consists of the features [low] > [coronal] > [labial] > [ATR]. The SDA, using this hierarchy, creates the tree shown in (12a), and the resulting specifications are listed in (12b).[6]

(12) Contrastive hierarchy for Classical Manchu (Zhang 1996)

a. SDA: [low] > [coronal] > [labial] > [ATR]



b. Output specifications

$$/i/ = \begin{bmatrix} -\text{low} \\ +\text{coronal} \end{bmatrix} \quad /u/ = \begin{bmatrix} -\text{low} \\ -\text{coronal} \\ +\text{ATR} \end{bmatrix} \quad /ʊ/ = \begin{bmatrix} -\text{low} \\ -\text{coronal} \\ -\text{ATR} \end{bmatrix}$$

$$/a/ = \begin{bmatrix} +\text{low} \\ -\text{labial} \\ -\text{ATR} \end{bmatrix} \quad /ə/ = \begin{bmatrix} +\text{low} \\ -\text{labial} \\ +\text{ATR} \end{bmatrix} \quad /ɔ/ = \begin{bmatrix} +\text{low} \\ +\text{labial} \end{bmatrix}$$

Starting at the top of the hierarchy, the feature [low] makes the first dichotomy in the inventory: all segments must be specified for this feature. I will use the constraint MAX [F] to require that any specification of F (+ or –) must be preserved. On the assumption that only fully specified inputs are provided to

6 For expositional reasons I use binary features for this example, in place of the privative features used by Zhang (1996). This example and the rationale for the analysis are discussed in depth in section 7.4.

the grammar, ranking MAX [low] highest will ensure that any underlying value of [low] is preserved in the output.

The next feature in the Manchu hierarchy is [coronal]. If we rank MAX [coronal] next, then any underlying value of this feature will be preserved in the output. However, a segment that is [+low] may not be specified for [coronal]; that is, we must exclude the feature combination [coronal, +low] for any value of [coronal]. Therefore, ranked higher than MAX [coronal], but lower than MAX [low], we require a co-occurrence constraint ruling out this combination of features.

The next feature is [labial]. We repeat the above procedure, specifying a co-occurrence constraint *[labial, –low,] ranked above the constraint MAX [labial].

The fourth and final feature is [ATR]. Here there are two co-occurrence restrictions, one for each branch of the tree descending from [low]. One is *[ATR, +labial, +low], the other is *[ATR, +coronal, –low]. Since [labial] has already been excluded in the domain of [–low], and [coronal] is excluded with [+low], the specifications of [low] in the [ATR] exclusions are redundant here. Following these restrictions we add, as before, MAX [ATR], which requires the preservation of [ATR] wherever it is not excluded.

As the contrastive hierarchy ends at this point, every other feature specification is simply excluded: *[F].

Summing up, we require the constraint types in (13).

(13) Two basic constraint types

a. MAX [F]: ‘Preserve the feature value of F (either + or –).’
b. *[F, Φ], where F is the feature to be excluded, and Φ is the set of feature values (of features with wider scope than F) forming the context of F.

The constraint hierarchy for Classical Manchu is given in (14), and the general procedure for converting any contrastive hierarchy into an OT constraint hierarchy is summarized in (15).

(14) Classical Manchu constraint hierarchy

MAX [low] ≫ *[coronal, +low] ≫ MAX [coronal] ≫ *[labial, –low] ≫ MAX [labial] ≫ *[ATR, +coronal], *[ATR, +labial] ≫ MAX [ATR] ≫ *[F]

a. MAX [low]	Ordered highest.
b. *[coronal, +low]	If [+low], cannot specify a value for [coronal].
c. MAX [coronal]	Must be ordered after (b).
d. *[labial, –low]	After (c); not crucially here, because [coronal] > [labial].
e. MAX [labial]	Must be ordered after (d).

f. *[ATR, +coronal] After (e) because [labial] > [ATR].
g. *[ATR, +labial] After (e), in the same stratum as (f).
h. MAX [ATR] After (f) and (g).
i. *[F] For F = any feature.[7] After (h).

(15) Converting a contrastive hierarchy to a constraint hierarchy
Given an ordering of features:

a. Go to the next contrastive feature in the list, F_i. If there are no more contrastive features, go to (e).
b. In the next stratum of constraints, place any co-occurrence constraints of the form *[F_i, Φ], where Φ consists of feature values of features ordered higher than F_i.
c. In the next stratum, place the constraint MAX [F_i].
d. Go to (a).
e. In the next constraint stratum, place the constraint *[F], and end.

A sample constraint tableau is given in (16).

(16) Constraint hierarchy corresponding to Classical Manchu contrastive hierarchy
Input: [–low, +coronal, –labial, +ATR] (overspecified /i/)

Candidates	MAX [low]	*[cor, +low]	MAX [cor]	*[lab, –low]	MAX [lab]	*[ATR, +cor]	*[ATR, +lab]	MAX [ATR]	*[F]
a. [–low, +cor, –lab, +ATR]				*!		*			****
b. [–low, +cor, –lab]				*!				*	***
c. ☞ [–low, +cor]					*			*	**
d. [–low, –lab]			*!	*				*	**
e. [+low, –lab, +ATR]	*!		*						***
f. [+cor]	*!				*			*	*
g. [–low, +cor, –back]					*			*	***!

7 There is no need to specify that [F] here is different from the contrastive features. If [F] is any feature, then every candidate will incur a violation for each feature it specifies. In the case of contrastive features, this violation will be harmless, because any candidates that lack a required feature will already have been ruled out by a higher constraint.

In this example the input is overspecified with respect to the contrastive specifications allowed by this contrastive hierarchy. The feature specifications of the input vowel correctly characterize the phonetics of Manchu /i/, but in the analysis in (12) the only contrastive features are [–low, +coronal]; the feature values [–labial] and [+ATR] are present phonetically but not phonologically. Candidates (a) and (b) maintain both or one of these noncontrastive features and both are losers because they violate the co-occurrence constraint against having [labial] together with [–low]. Candidate (c) is the winning output: it dispenses with the features [labial] and [ATR], but retains [low] and [coronal], as required by the contrastive hierarchy. Candidate (d) also retains two features of the input, [low] and [labial], but loses because Max [coronal] is ranked higher than Max [labial]. Candidate (e) retains three features in a licit configuration by changing the input [–low] to [+low], but is immediately ruled out because this change violates Max [low]. Candidate (f) retains only [+coronal]. Though through logical redundancy this would uniquely identify the phoneme /i/, it is ruled out as a representation because it, too, violates Max [low] (i.e., no underspecification 'from below' is permitted). Finally, candidate (g) retains the correct specifications from the input but adds a redundant feature, [–back]. It thus incurs an extra violation of *[F] and so loses to (c).

6.5 Constraint hierarchies not corresponding to a contrastive hierarchy

The procedure in (15) can convert any legal contrastive hierarchy into a constraint hierarchy. The converse is not the case, however. We have seen that a number of commentators have observed that OT is capable of simulating a wide spectrum of grammars, ranging from those with highly underspecified representations to those with full specification, as well as grammars intermediate between these. If all of these grammars are indeed attested, then this degree of expressive power is warranted. If, however, there are constraints on the degree of specification allowed in grammars, then it is undesirable to allow constraint hierarchies that produce impossible grammars.

The assumption here, following from the Contrastivist Hypothesis, is that the contrastive hierarchy sets the limits on specification in phonology, at least to a first approximation. If that is the correct position, then we wish to constrain OT to be limited to hierarchies that simulate legal contrastive hierarchies. That is, we could require that OT grammars contain only constraint hierarchies that adhere to (15). In the absence of such a constraint, OT grammars can indeed mimic all manners of specification and underspecification. This can be illustrated with the same Classical Manchu vowel system we considered above.

Let us begin with full specification. Full specification of the features of an inventory can be required simply by promoting all MAX [F] constraints to the top of the constraint hierarchy. A tableau of this type is given in (17).[8] It is clear that only a fully specified output can be a winner, given a fully specified input, since all features must be preserved.[9]

(17) Full specification of all features: all MAX constraints promoted to the top
Input: [–low, +coronal, –labial, +ATR, –back, +high, –nasal] (fully specified /i/)

Candidates	MAX [low]	MAX [cor]	MAX [lab]	MAX [ATR]	MAX [bck]	MAX [hi]	MAX [nas]	*[F, Φ]
a. ☞ [–low, +cor, –lab, +ATR, –bck, +hi, –nas]								*...
b. [–low, +cor, –lab, +ATR, –bck, +hi]							*!	*...
c. [–low, +cor, –lab, +ATR]					*!	*	*	*...
d. [–low, +cor]			*!	*	*	*	*	*...

We can limit the phonology to full specification of 'relevant' features (i.e., features that are contrastive somewhere in the phonology) by promoting only the MAX constraints for contrastive features, as in (18).[10] Features that are redundant everywhere are filtered out, but all segments must be specified for all contrastive features, even if they are not contrastive in any given contrastive hierarchy.

8 The various co-occurrence restrictions are all abbreviated as *[F, Φ] in this tableau, since they have no effect on the outcome. *... represents an indefinite number of violations.

9 Recall that we are limiting inputs to being fully specified, for now. MAX constraints apply only to features present in the input. As we shall see when we consider underspecified inputs, this fact allows unconstrained OT constraint systems to create unusual grammars in which underspecified inputs are required to be underspecified, while fully specified inputs must remain so in the output.

10 The various co-occurrence restrictions are again abbreviated as *[F, Φ]. MAX [G] refers to all the redundant features that play no contrastive role anywhere in the inventory.

(18) Full specification of 'relevant' features: some MAX constraints promoted to the top
Input: [–low, +coronal, –labial, +ATR, –back, +high, –nasal] (fully specified /i/)

Candidates	MAX [low]	MAX [cor]	MAX [lab]	MAX [ATR]	*[F, Φ]	*[F]	MAX [G]
a. [–low, +cor, –lab, +ATR, –bck, +hi, –nas]					*...	*****!**	
b. [–low, +cor, –lab, +ATR, –bck, +hi]					*...	*****!*	*
c. ☞ [–low, +cor, –lab, +ATR]					*...	****	*...
d. [–low, +cor]			*!	*		**	*...

We can mimic the effect of the Pairwise Algorithm, giving a type of underspecification. A set of feature specifications for Classical Manchu consistent with those generated by the Pairwise Algorithm is given in (19).[11]

(19) Pairwise Algorithm specifications for Classical Manchu

Vowel	/i/	/ə/	/a/	/ɔ/	/u/	/ʊ/
[low]	–	+		+		–
[labial]	–		–	+	+	
[ATR]		+	–		+	–

These specifications can be obtained by placing a series of pairwise restrictions ordered in a block above a set of MAX constraints, also in a block. Since the Pairwise Algorithm does not incorporate any kind of ordering of features, it follows that the constraints of the same type also do not have to be ordered. Sample tableaux with two different inputs are given in (20) and (21).

(20) Underspecification by means of the Pairwise Algorithm
Input: [–low, –labial, +ATR, +coronal] (overspecified /i/)

Candidates	*[+low, –lab] *[–low, +lab]	*[–low, +ATR] *[+low, –ATR]	*[–lab, +ATR] *[+lab, –ATR]	MAX [low]	MAX [lab]	MAX [ATR]	*[F]
a. [–low, –lab, +ATR, +cor]		*!	*				****
b. [–low, –lab, +ATR]		*!	*				***
c. [–low, +cor]					*!	*	**
d. ☞ [–low, –lab]						*	**
e. [–lab, +ATR]			*!	*			**
f. [+cor]				*!	*	*	*

11 The feature [coronal] is omitted because the Pairwise Algorithm would fail if it were included.

(21) Underspecification by means of the Pairwise Algorithm
Input: [+low, –labial, −ATR] (overspecified /a/)

Candidates	*[+low, –lab] *[–low, +lab]	*[–low, +ATR] *[+low, –ATR]	*[–lab, +ATR] *[+lab, –ATR]	MAX [low]	MAX [lab]	MAX [ATR]	*[F]
a. [+low, –lab, –ATR]	*!	*					***
b. [+low, –lab, +ATR]	*!		*			*	***
c. [–low, –lab]				*		*!	**
e. ☞ [–lab, –ATR]				*			**
f. [+low, +ATR]					*	*!	**

An OT constraint hierarchy can be made to simulate the theory of RU (see chapter 5, section 3.2). In this theory, one value of each feature, the marked value, is specified on every segment. RU also assumes that a minimal number of features is used. A possible RU analysis of Classical Manchu is given in (22).[12]

(22) Radical Underspecification for Classical Manchu

Vowel	/i/	/ə/	/a/	/ɔ/	/u/	/ʊ/
[low]		+	+	+		
[labial]				+	+	+
[ATR]	+	+			+	

This type of underspecification is easily modelled in OT by positing a set of constraints of the form MAX [m F], where [m F] is the marked value of F.[13] These constraints can be ordered in a single stratum above *[F]. A sample tableau is shown in (23).

(23) Radical Underspecification
Input: [–low, –labial, +ATR, +coronal, +high] (overspecified /i/)

Candidates	MAX [+low]	MAX [+lab]	MAX [+ATR]	*[F]
a. [–low, –lab, +ATR, +cor, +hi]				**!***
b. [–low, –lab, +ATR]				**!*
c. ☞ [+ATR]				*
d. Ø			*!	
e. [+low, +ATR]				**!

12 The feature [coronal] is again omitted here to keep to a minimal feature set.

13 It is not so straightforward to model the derivational aspect of RU – the fact that feature values can be inserted in the course of a derivation. Also, we continue to restrict inputs to fully specified, or overspecified, forms.

We have seen, then, that OT is capable of simulating the effects of any theory of specification or underspecification. This is to be expected, since OT is primarily a theory of constraint interaction, not a theory of specification. If we are to have a theory of specification, OT must be constrained so as to produce the desired representations and exclude the others. To the extent that phonologies adhere to contrastive specifications of the sort that can be produced by the SDA, OT must be constrained, for example along the lines of (15).

More fundamentally, such constraints on the theory are required if OT is to instantiate some version of the Contrastivist Hypothesis. The reason for limiting specifications to those that are contrastive is to capture the relationship between contrast and phonological activity.

6.6 Inputs containing illicit feature combinations

Up to here we have considered only inputs that are as specified or more specified than the contrastively specified outputs. For example, the Manchu vowel /i/, contrastively specified as [–low, +coronal], can be input as such, or with additional redundant features, for example [–low, +coronal, –labial, –ATR, –back]. The constraint system filters out the redundant features and outputs the contrastive representations we have proposed above.

Many OT theorists require the constraint system to do more than filter redundancies from fully specified representations. According to ‘richness of the base’ (Prince and Smolensky 2004), we might expect the constraint system to be able to convert *any* arbitrary input into a licit output. For purposes of this discussion, it is useful to recognize two additional types of inputs: fully specified inputs that contain combinations of features that are not permitted in the given inventory (i.e., input segments that are not part of the inventory); and underspecified inputs (i.e., inputs lacking certain required features). Let us first consider cases of the former kind.

Continuing with our Manchu example, we can ask what the constraint hierarchy in (16) would do with an input [+low, +coronal, –labial, –ATR, –back, –high, etc.], that is, a front low vowel such as /æ/ or /ɛ/. Such a vowel does not exist in Classical Manchu. A tableau showing the results given by the hierarchy in (16) is given in (24).

(24) Illicit fully specified input (segment outside the inventory)
Input: [+low, +coronal, –labial, –ATR, –back, –high, etc.] (/æ/)

Candidates	MAX [low]	*[cor, +low]	MAX [cor]	*[lab, –low]	MAX [lab]	*[ATR, +cor]	*[ATR, +lab]	MAX [ATR]	*[F]
a. [+low, +cor, –lab, –ATR, –bk, –hi]		*!				*			******
b. [+low, +cor, –lab, –ATR]		*!				*			****
c. [+low, +cor, –lab]		*!						*	***
d. [+low]			*		*!			*	*
e. [+low, –lab]			*					*!	**
f. ☞ [+low, –lab, –ATR]			*						***
g. [–low, +cor]	*!				*			*	**
h. [–low, +cor, –lab, –ATR]	*!			*		*			****

The constraint system gives a unique optimal output, [+low, –labial, –ATR], namely /a/. The high ranking of MAX [low] ensures that the underlying specification [+low] is maintained, and the next highest-ranking constraint, *[coronal, +low], rules out any output that contains a coronal feature specification. All the other features must be preserved insofar as they are not excluded. Thus, the competition between the conflicting [+low] and [+coronal] input specifications is resolved in favour of the feature that is higher in the hierarchy.

This situation is tantamount to asking the question: what would speakers of Classical Manchu do when presented with a vowel that is not part of their inventory? This, of course, is part of a larger question: how are foreign (L2) sounds integrated into one's native (L1) phonology? This is an empirical question, and a considerable literature has arisen to address it. Here I will briefly look at the role of the contrastive hierarchy in the treatment of L2 sounds that fall outside the native L1 inventory. I will return to this topic in section 7.7, where I will present an empirical study that makes crucial use of the contrastive hierarchy in accounting for patterns of loanword adaptation. Here, I will consider the matter more abstractly.

One possible approach is the one we took above: in any clash of features, the feature higher on the hierarchy takes precedence over the one lower down, and the rest of the features are retained or not to the extent that they fit into the native contrastive system. To take another example involving Classical Manchu, an input vowel /o/, which we will assume is specified [+low, –coronal, +labial, +ATR] (i.e., the [+ATR] counterpart of /ɔ/), will first of all retain its [+low] specification. Since [coronal] is excluded in the [+low] domain, that feature will be filtered out. In this case, the exclusion of [coronal] has no perceptible effect, since the vowel will in any case be phonetically [–coronal]. Continuing down the hierarchy, [+low] vowels are all specified for [labial], so the [+labial] specification is retained. Vowels specified [+low, +labial] admit of no further specification in Classical Manchu, so the [+ATR] specification is filtered out. The prediction of this approach is that an input /o/ will be converted to its (redundantly) non-ATR counterpart, /ɔ/.

This is the result we obtain by following down the contrastive tree in (12). If L2 words are borrowed following this approach, foreign sounds are assigned to native categories by preserving the feature specifications with widest scope at the expense of those with narrower scope. This is a purely hierarchical approach to loan phonology. It is possible that L2 phonology works this way; however, it is not necessary to apply the contrastive hierarchy to foreign sounds in just this way. Other factors may intercede to lead to different results.

Consider again, for example, the three-vowel system in (10), where [high] > [round]. Suppose we were to present the grammar with an input vowel /o/, that is, a vowel specified [–high, +round]. Following a purely hierarchical approach, we would require [–high] to be preserved above any other consideration. Once the vowel is designated [–high], any specification for [round] must be filtered out, and the result would be [–high], that is, /a/.

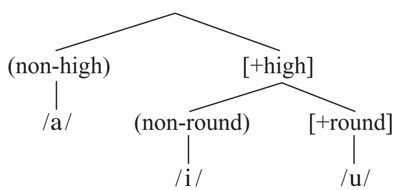
Again, this is perhaps the right outcome. But one can imagine a different outcome, whereby the [+round] specification is perceived by L1 speakers as being more *salient* than the height feature. If the priority is to retain this [+round] specification, then necessarily the height feature will have to be adjusted, giving the output [+high, +round], that is, /u/.

This scenario raises the question of the relation between salience and the contrastive hierarchy. On the approach taken here, saliency is closely tied up with the contrastive hierarchy. Now it could be that salience is to some extent free of language-particular hierarchies. In that case, the contrastive hierarchy and salience would be two coexisting dimensions that each contribute in different degrees to various aspects of the phonology. To the extent that salience would also have to be organized into a hierarchy (to sort out what happens when two features are both salient but conflicting), it would of course be more desirable, from a theoretical point of view, if the contrastive and salience hierarchies could be conflated, or at least related to one another. The matter is, however, empirical, and will ultimately be decided on the evidence.

Before leaving this topic, it is worth pointing to a third position, one which relates salience to the contrastive hierarchy, but which also assigns a role to markedness.

Let us suppose that for each binary feature, one value is unmarked (default) and the other is marked. A more radical implementation of this idea is to adopt privative features, where only the marked value is specified at all, leaving the unmarked value to be literally unmarked. Suppose that in the three-vowel system in (10) the values [+high] and [+round] are marked, and the minus values are unmarked (or even absent). Assigning labels only to marked values, the tree in (10) would appear as in (25).

(25) Contrastive hierarchy for /i, a, u/, [high] > [round]



If marked features were to take priority over unmarked features, then an input /o/ would have the marked feature [+round]. Preserving this feature would result in the output [+high, +round], that is, /u/.

This result can be obtained from a constraint system like the one in (24) if the Max [F] constraints were understood as applying only to marked features. Suppose, for the sake of this discussion, that in Classical Manchu the marked features were [–low], [+coronal], [+labial] and [+ATR]. To keep the marked values identified with plus values, we will for the rest of this example substitute [high] for [low].[14] Then the input /æ/, [–high, +coronal, –labial, –ATR], would have the marked value [coronal]. As shown in (26), the predicted outcome is now /i/, that is, [high, coronal]. In (26) only marked values are shown. A constraint like *[coronal, –high] must be reformulated, assuming there is no value [–high]. We still want to exclude the specification [coronal] in the absence of [high]. We can recast the constraint in terms of licensing (cf. Itô, Mester and Padgett 1995) as [coronal] ⊂ [high] ('[coronal] must be licensed by [high]'), or

14 I adopt the assumption that [–low] or [+high] is the marked feature in Classical Manchu only for purposes of this example, so as to create a mismatch between the highest feature, [high], and the highest marked feature, [coronal]. Elsewhere, however, I continue to assume that [+low] (or [low]) is the marked feature value in Classical Manchu.

as a co-occurence constraint *[coronal, *u* high] ('[coronal] may not be specified if [high] is unspecified').

(26) Marked (privative) features: illicit input (segment outside the inventory)
Input: [coronal] (/æ/)

Candidates	MAX [hi]	*[cor, *u* hi]	MAX [cor]	*[lab, hi]	MAX [lab]	*[ATR, cor]	*[ATR, lab]	MAX [ATR]	*[F]
a. [cor]		*!							*
b. [cor, ATR]		*!				*			**
c. [Ø]			*!						
d. ☞ [hi, cor]									**
e. [hi, cor, ATR]						*!			***

This brief discussion shows that the contrastive hierarchy by itself does not dictate a particular approach to the question of how foreign sounds are incorporated into a phonology. In conjunction with other concepts, such as markedness, it does make available a range of empirical hypotheses.

Privative feature systems necessarily involve underspecification, as seen in (26). Such underspecification is not particularly problematic in this context, since it is permanent. Thus, the presence versus absence of a feature forms a binary opposition, parallel to that between a plus and minus value of a feature. In binary feature systems, however, underspecification takes on a different status, creating in effect a ternary opposition between plus, minus and zero. In the next section we consider the effect of underspecified inputs in a binary feature system.

6.7 Underspecified inputs

The notion of 'richness of the base' can be extended to include not just illicit combinations of fully specified features, but also underspecified inputs; by 'underspecified' I mean here underspecified relative to the representations derived by the contrastive hierarchy. Whereas inputs consisting of illicit combinations of features can be interpreted as foreign sounds being presented to a speaker, it is less obvious in what actual or hypothetical situations a speaker will be presented with an underspecified input. Since it is not possible to isolate individual features in any percept that a speaker can actually be exposed to, it is hard to imagine under what circumstances speakers – let us say of Classical Manchu, to continue with our example – might interpret a vowel that sounds something like [a] as consisting only of the feature [+low], rather than, as required by the contrastive hierarchy, [+low, –labial, –ATR].

Nevertheless, richness of the base has been a basic part of the conceptual system of OT, so let us see what the consequences are of allowing such inputs. In (27), an underspecified input consisting only of [+low] is provided as an input to the constraint system we have been using for Classical Manchu.

(27) Underspecified input
Input: [+low]

Candidates	MAX [low]	*[cor, +low]	MAX [cor]	*[lab, –low]	MAX [lab]	*[ATR, +cor]	*[ATR, +lab]	MAX [ATR]	*[F]
a. *☞ [+low]									*
b. [+low, –lab]									**!
c. [+low, +lab]									**!
d. [+low, –lab, +ATR]									**!*
e. [+low, –lab, –ATR]									**!*
f. [+low, –cor]		*!							**
g. [–low]	*!								*
h. [Ø]	*!								

The constraint system developed so far does not adequately deal with underspecified representations. The high ranking of MAX [low] ensures that the underlying specification [+low] is maintained, and the next highest-ranking constraint, *[coronal, +low], rules out any output that contains a coronal feature specification. Beyond these requirements, however, the only constraint that plays a role in the outcome is *[F], which incorrectly selects (a), the least specified candidate that preserves the input [+low].

The problem is that constraints of the type MAX [F] require only that an input specification of [F] must be preserved. They have no jurisdiction over a feature that is absent from the input. If we consider a contrastive tree such as in (12), we understand that it not only prohibits certain feature combinations, it also *requires* some features to be present. MAX constraints do not do this, nor do the co-occurrence constraints. For while a constraint of the form *[coronal, +low] prohibits certain combinations of features, it does not require any combinations. Candidates that lack some or all of the features mentioned in such a constraint simply evade its effects.

To enforce the presence of required contrastive features we need to add a third type of constraint, requiring the presence of a feature in the context of other features. One way is by an implicational constraint of the form in (28).

(28) Implicational constraint
$\Phi \supset$ F: 'The set of feature specifications Φ requires the presence of feature F.'

When a value of [F] is specified in the input, MAX [F] will ensure it is maintained. When no value is specified, the implicational rule will penalize any candidate that lacks [F].

A set of implicational constraints for Classical Manchu that mirrors the contrastive tree in (12) is given in (29).

(29) Implicational constraints for Classical Manchu
a. [low] The feature [low] must always be specified.
b. [+low] ⊃ [labial]
c. [–low] ⊃ [coronal]
d. [–labial] ⊃ [ATR]
e. [–coronal] ⊃ [ATR]

We now observe a redundancy between the constraints that penalize a feature in a context and those requiring it: the contexts in which these apply are bound to be complementary. Therefore, we can simplify the implicational constraints to simply require the presence of a feature, as long as these are ordered after the constraints forbidding the feature. This type of constraint can take the form shown in (30).

(30) Specification Constraint
SPEC [F]: '[F] must be specified.'

If we add such constraints en bloc after the constraints we have been using till this point, they will filter out candidates that lack required features. For example, in (31) the underspecified candidates [+low], [+low, –labial], and [+low, +labial] are eliminated because they lack features that are required in the context of [+low]. On the other hand, the overspecified [+low, –coronal, –labial, –ATR] does not benefit from its lack of violations of these constraints because it is already excluded by the higher-ranking constraint excluding [coronal] in the context of [+low].[15]

15 The constraint column MAX, *[F, αG] is simply shorthand for all the MAX and *[F, αG] constraints as they are ranked in (27).

(31) Underspecified input, adding SPEC [F], where [F] is one of the contrastive features
Input: [+low]

	Candidates	MAX, *[F, αG]	SPEC [low]	SPEC [cor]	SPEC [lab]	SPEC [ATR]	*[*m* F]
a.	[+low]			*	*!	*	*
b.	[+low, –lab]			*		*!	*
c.	[+low, +lab]			*		*!	**
d.	[+low, –lab, +ATR]			*			**!
e. ☞	[+low, –lab, –ATR]			*			*
f.	[+low, –cor, –lab, –ATR]	*!					*
g.	[–low]	*!		*	*	*	
h.	[Ø]	*!	*	*	*	*	

The addition of the SPEC [F] constraints narrows the field down, but does not suffice to decide which value of [ATR] should be supplied.[16] Presumably, we want to supply the least marked values. Thus, we have to add a set of markedness constraints that penalize the marked value of each feature.

(32) Markedness constraints (contextual)
*[*m* F, Φ]: 'Value m of F is penalized in the context of Φ.'

If markedness does not vary by context, then this type of constraint can be reduced simply to (33).

(33) Markedness constraints (noncontextual)
*[*m* F]: 'Value m of F is penalized.'

In (31), these constraints are added at the bottom of the constraint hierarchy. As observed in the previous note, different rankings of the markedness constraints yield somewhat different results.

16 The same might be said of the feature [labial]. In tableau (31), the candidate [+low, +labial] loses because it lacks a feature, [ATR], even though the combination [ATR, +labial] is excluded in Classical Manchu. Thus, ordering all the SPEC [F] constraints before the markedness constraints favours candidates that maximize the number of specified features. This may not be the correct answer. If, for example, [+labial] were unmarked, one might expect that the optimal way to fill out an input consisting only of [+low] would be [+low, +labial], rather than filling in [–labial] and a value of [ATR]. This result can be obtained by interleaving the markedness constraints with the SPEC [F] constraints.

6.8 Summary

In this chapter I have argued that OT is not itself a theory of contrast in competition with the Contrastivist Hypothesis or the contrastive hierarchy. Rather, OT can incorporate various theories of contrast. Though it is true that some approach to contrast will inevitably 'emerge' from any OT constraint hierarchy, the empirical question is what theories of contrast are allowed to occur in phonology. If, for example, the Contrastivist Hypothesis as implemented by the SDA operating on a contrastive feature hierarchy is the correct theory, then OT should be constrained to implement that theory.

I have shown how OT with different strata (serial, or derivational OT) can incorporate the contrastive hierarchy by modelling the SDA. In OT with multiple strata, the constraints corresponding to the SDA can be segregated from the rest of the phonology; the output of these constraints consists of well-formed contrastive specifications that serve as the input to the phonology proper.

Whether the Contrastivist Hypothesis can be incorporated into a monostratal version of OT is a more difficult question. Such a grammar does not make available a level at which only contrastive feature specifications are visible, and constraints implementing the SDA cannot be kept separate from the other constraints. Moreover, there are problems with requiring that phonemes be underspecified in various ways. I cannot pursue these issues here, but they are considered in some detail by D. C. Hall (2007: chapter 5). Hall concludes that there are real difficulties in incorporating the Contrastivist Hypothesis into a standard monostratal OT.

We will return to some of the issues discussed here in the following chapter, where I take up a series of case studies that provide evidence for the contrastive hierarchy and the Contrastivist Hypothesis.

7 *Evidence for the contrastive hierarchy in phonology*

7.1 Introduction

In section 3.7 I sketched the outlines of a theory of phonology that was distilled from the leading ideas discussed in chapter 3. This theory adopts the Contrastivist Hypothesis, which holds that phonology computes only contrastive features. It determines what the contrastive features in a language are by applying the SDA to a contrastive feature hierarchy for that language. In keeping with the Contrastivist Hypothesis, phonological activity serves as the chief heuristic for determining what the feature hierarchy is for a given language.

Though the ingredients for such a theory were in place by the 1930s, phonological theory did not develop in this direction; why it did not was the subject of chapters 4–6. These chapters show that the theory of section 3.7 has never properly been put to the test. In this chapter I argue that these ideas remain viable and indispensable to an explanatory theory of phonology.

Of course, any contemporary effort to implement such a theory must take account of advances in phonology since the 1930s. For example, the diagnostic given in (38d) of chapter 3, that a contrastive feature must be present in all the allophones of a phoneme, is not consistent with the generative phonological conception that phonology is relatively *abstract* with respect to phonetics. In keeping with Chomsky and Halle's arguments against taxonomic phonemics, it is unlikely that we can put limits on the degree to which a segment may be modified in the course of a derivation. But the principle in (38d) may still have some heuristic value: the fact that all the allophones of a phoneme share a certain feature could lead us to suspect that this feature is contrastive, in the default case, in the absence of stronger conflicting evidence.

An example is the Russian phoneme /i/, which has allophones that vary in backness but share the property of being [−round]. In the absence of conflicting evidence, we would suppose, with Jakobson (1962b [1931]), that [−round] is contrastive for /i/ and [back] is redundant (see chapter 1, section 1).

However, we will see in section 8.3 that there is more compelling evidence from phonological activity that /i/ is contrastively [−back]; the [+back] allophone is the result of spreading of [+back] from a preceding consonant.

In this chapter I will provide further evidence for the Contrastivist Hypothesis and the contrastive hierarchy in the framework of a contemporary theory that has the main properties of the one sketched in chapter 3, section 7. The cases surveyed below draw mainly on research carried out at the University of Toronto since the mid-1990s in the context of the project on ‘Markedness and the Contrastive Hierarchy in Phonology’ (Dresher and Rice 2007). This approach to phonology has come to be known as Modified Contrastive Specification (MCS).

In the following section I will sketch some of the key characteristics of MCS, pointing to similarities with and differences from other contemporary approaches. Particular topics are the relationship between contrast and markedness (section 7.2.1), variation in the feature hierarchy (section 7.2.2) and the relationship between phonology and phonetics (section 7.2.3).

Section 7.3 shows that feature hierarchies, considered apart from any theory of contrast, are widespread in phonological theory and practice; therefore, the incorporation of contrastive hierarchies into MCS does not create a descriptive or explanatory burden not shared by other theories of phonology. Subsequent sections take up, in turn, vowel harmony (section 7.4), metaphony (section 7.5), consonant co-occurrence restrictions (section 7.6), loanword adaptation (section 7.7) and language acquisition (section 7.8). I will argue that evidence in all these domains supports the MCS approach. Finally, section 7.9 considers whether the Contrastivist Hypothesis is too strong, and discusses some possible refinements.

7.2 Modified Contrastive Specification

MCS began with a focus on complexity in phonology (Avery and Rice 1989; Dresher and Rice 1993; Dresher, Piggott and Rice 1994; Dresher and van der Hulst 1998), and evolved to concentrate on markedness and contrast.[1] In this model, complexity in representations is driven by both contrast and markedness. Assuming that each feature has a marked and unmarked value, MCS posits that

1 Some of the material in this section is adapted from a description of MCS written jointly with Keren Rice.

only marked features count toward complexity; thus, segments with fewer marked features are less complex than those with more marked features.[2]

7.2.1 Contrast and markedness

MCS proposes that contrasts are determined by the SDA operating on a hierarchy of features. Since a more marked representation is permitted only if needed to establish a contrast with a less marked one, the theory of MCS leads us to expect a relation between the amount of segmental markedness a system allows and the number and nature of contrasts it has.

The assumption that markedness is related to contrast is inconsistent with the view that markedness scales are universally fixed (Chomsky and Halle *SPE*; Cairns 1969; Kean 1980; Beckman 1997; Lombardi 2002; Prince and Smolensky 2004; de Lacy 2006; see Rice 2007 for discussion). For example, in a vowel inventory with a front and back vowel, say /i, a, u/, either /i/ or /u/ may pattern as marked with respect to phonological activity, because only one contrast is required to distinguish front unrounded from back rounded vowels (say, [labial] or [coronal], but not both).[3] However, if a central vowel such as /ɨ/ or /ə/ is added, the prediction is that both the front and back vowels will pattern as marked with respect to the central vowel. This follows from the assumption that there is no feature [central], with the consequence that now both [labial] and [coronal] are required to distinguish the vowels from each other. It thus follows that the central vowel must be unmarked, a prediction that is empirically supported (see Rose 1993; Walker 1993; and Rice 2007).[4]

An illustration of this principle can be found in developments in the Yupik and Inuit/Inupiaq dialects of Eskimo-Aleut.[5] These dialects descend from a proto-language that had four vowels, as shown in (1).

2 There are many conceptions of markedness in the contemporary literature. As with contrast, the MCS notion of markedness also has roots in the work of Trubetzkoy, particularly in his distinction between *logical* (structural) and *natural* (phonetic) markedness (cf. Rice 2007). In MCS the emphasis is on logical markedness, which is relative to a particular inventory.

3 I will continue the practice of the previous chapters of using the feature names used in my sources. In MCS we have been using vowel features [coronal] and [labial], and sometimes [peripheral] (see Rice 1994, 2002). However, I take no stand here on whether vowel features are identical to consonant features or distinct from them (see Clements and Hume 1995 and Halle, Vaux and Wolfe 2000 for different views). For the purposes of this book, [coronal] is interchangeable with [front], and [labial] with [round].

4 Dispersion Theory (Flemming 2002, Padgett 2003a, b) also takes a systemic view of markedness.

5 ‘Inuit’ is the name used in Canada for this language family, ‘Inupiaq’ is the name used in Alaska. I would like to thank Richard Compton for discussion of this example.

(1) Proto-Eskimo vowels (Fortescue, Jacobson and Kaplan 1994)

i u
ə
a

Phonological patterning in Inuit dialects suggests that the contrastive features for this inventory are as given in (2) (Compton and Dresher 2008). Because of the symmetry of the vowel system, the ordering of the features is not crucial in this inventory: the division lines in (2) depict an ordering in which [low] is the first feature. We also suppose that [labial] is ordered above [coronal], for reasons that will become clear shortly. Only marked feature values are shown.

(2) Contrastive features for dialects distinguishing four vowels

[coronal] | | [labial]
i | | u
| ə |
———————————
a [low]

Yupik dialects and the Diomede subdialect of Bering Strait Inupiaq retain this four-vowel system. However, schwa does not have the same status as the other vowels: according to Kaplan (1990: 147), it 'cannot occur long or in a cluster with another vowel'. The latter phenomenon is characteristic of unmarked elements: they tend to be targets of phonological processes, and they are not triggers (Rice 2007). In this case, schwa assimilates to neighbouring vowels, and does not cause assimilation in other vowels.

The influence of contrast and markedness can be seen in Inuit dialects that have palatalization of consonants. On the assumption that palatalization of a consonant by a vowel is triggered by a contrastive [coronal] feature, /i/ in (1) could trigger palatalization, but /ə/ could not.[6] In most Inuit dialects the vowel represented as */ə/ has merged at the surface with */i/. Some contemporary dialects, however, distinguish between two kinds of *i*: 'strong *i*', which descends from */i/, and 'weak *i*', which comes from */ə/. In North Alaskan Inupiaq, strong *i* triggers palatalization of alveolar consonants, but weak *i* does not. Some examples from the Barrow dialect are given in (3). The suffixes in (a) have initial alveolar consonants following a stem ending in *u*; the suffixes in (b) show palatalization of the suffix-initial consonant following strong *i*; and the forms in (c) show that palatalization does not occur after weak *i*. Note that the palatalization of /t/ results in *s* here and in most Inuit dialects.

6 How palatalization works is the subject of some debate; see Kenstowicz (1994) and T. A. Hall (2007) for overviews and references. All that is important here is that /i/ bears some contrastive feature that triggers palatalization.

(3) Barrow Inupiaq palatalization after strong *i* (Kaplan 1981: 82)

	Stem	*Gloss*	*'and a N'*	*'N plural'*	*'like a N'*
a.	iglu	'house'	iglulu	iglunik	iglutun
b.	iki	'wound'	ikiʎu	ikiɲik	ikisun
c.	ini	'place'	inilu	ininik	initun

Further, weak *i* undergoes a variety of assimilation and deletion processes that do not affect strong *i* or the other vowels *u* and a. For example, weak *i* changes to a before another vowel (4a), but strong *i* does not (4b). According to Kaplan (1981), weak *i* also alternates with *u* in some restricted contexts and with zero (i.e., it syncopates) in other contexts.

(4) Barrow Inupiaq weak and strong *i* before a vowel (Kaplan 1981: 119)

a.	aŋuti	+	-u-	+	-tuq	→	aŋutauruq
	'man'		COPULA		3SG.		'it is a man'
b.	iki	+	-u-	+	-tuq	→	ikiuruq
	'wound'		COPULA		3SG.		'it is a wound'

Following Underhill (1976) and Kaplan (1981), I suppose that dialects that distinguish between strong and weak *i* retain four underlying vowels, as in the proto-language, with the same contrastive features as in (2). Though this analysis is 'abstract' with respect to the surface phonetics, the analysis is committed to specifying the fourth vowel phoneme only as not low, not labial and not coronal (i.e., some non-low unrounded central vowel).[7]

These contrastive marked values account for the fact that /i/ can trigger palatalization, as it has a relevant contrastive feature. The fourth vowel is the least marked, literally, and therefore cannot trigger palatalization, and is more susceptible to receive features from the context.

In over half of the Inuit dialects from Alaska to Labrador there is no longer any distinction between the two kinds of *i*: in all these dialects, etymological */i/ and etymological */ə/ have merged completely as *i*.[8] It is a striking fact that none of these dialects has consonant palatalization triggered by /i/ (Compton and Dresher 2008). Compton and Dresher (2008) posit that [low] and [labial] are ordered ahead of [coronal] in the contrastive hierarchy of the Inuit language family; with only three vowels in the inventory, only the former two features can be contrastive, as shown in (5). Lacking a contrastive feature, /i/ can no longer trigger palatalization. This analysis thus explains a conspicuous gap in the typology of Inuit dialects: there are palatalizing dialects with four underlying

7 See Archangeli and Pulleyblank (1994: 73–84) for an analysis that is similar in spirit, but proceeding from different theoretical assumptions.

8 See Dorais (2003) for a survey.

vowels, and non-palatalizing dialects with three underlying vowels, but no palatalizing dialects with three underlying vowels.

(5) Contrastive features for dialects distinguishing three vowels

i	[labial] u
a [low]	

Many other examples of contrast-driven markedness asymmetries have been adduced. Studies in the general MCS framework include, among others: Avery and Rice (1989), Rice and Avery (1993), Avery (1996), Rice (1996) and Radišić (2007) on asymmetries in the markedness of consonants; and Rice and Avery (1993), Rose (1993), Causley (1999), Frigeni (2003, 2009), Rice (2002, 2003), D'Arcy (2003a) and Rohany Rahbar (2006, 2008) on vowel systems.[9]

7.2.2 Variation in the feature hierarchy

Analyses of many languages in MCS suggest a certain amount of variation in the feature hierarchy. We observed such variation in chapters 3 and 4, and more examples will be presented in the course of this and the following chapter. To cite but one example here, we have seen immediately above that Inuit vowel systems have the order [labial] $>$ [coronal]. Many languages show a similar pattern of /i/ being unmarked in relation to /u/. But this pattern is not universal; we will see in section 7.4 that Manchu-Tungus and Mongolian vowel systems have [coronal] $>$ [labial].

Studies in MCS that demonstrate variability in the feature hierarchy include work on vowel systems (Aka: Balcaen 1998; Japanese: Hirayama 2003; Manchu-Tungus: Zhang 1996; Mandarin: Zhou 1999; Miogliola: Ghini 2001a, b; Persian: Rohany Rahbar 2006, 2008; Québécois French: Mercado 2002; Sumerian: E. J. M. Smith 2007; Yokuts: D'Arcy 2003a) and consonant systems (Czech: D. C. Hall 1998, 2007, 2008; Inuktitut: Compton 2008; Japanese: Rice 1997, 2005; Latvian: Vilks 2002; Mandarin: Wu 1994; Old English: Moulton 2003; Vietnamese: Pham 1997, 1998). Variations are observed in the feature hierarchy for place of articulation (Rice 1996), stop vs continuant consonants, and height vs place for vowels. Avery (1996) found evidence for

9 For a concise review of this issue, see Rice (2007). Not all activity-based diagnostics that have been proposed in the literature are equally reliable. See Rice (2003, 2007, forthcoming) for further discussion of markedness. Rice argues that asymmetries in assimilation provide the most revealing test of markedness. If we found, for example, that /t/ assimilates to /k/ but /k/ does not assimilate to /t/, that would be evidence that /k/ is marked relative to /t/.

variation in voicing systems. In tone systems, high tones are unmarked in some languages and low tones in others (Rice 2003; Rice and Hargus 2005).

The limits of this variation remain to be determined. It may be, for example, that some of the cross-linguistic variation is related to differences in the structure of inventories (Béjar 1998; Casali 2003; Herd 2005). That there are limits is suggested by the fact that certain feature orders produce unnatural-looking inventories, as shown in section 7.3.

7.2.3 Phonology and phonetics

A consequence of the MCS approach is that phonology is necessarily underspecified with respect to phonetics: the number and nature of contrasts that a segment enters into influence, but do not determine, its phonetic realization. Therefore, the contrastive specifications assigned by the phonological component must be supplemented by further principles to derive the detailed phonetic specification of a speech sound.

In the four-vowel system in (1), for example, the vowel /a/ is phonologically specified as being contrastively [low]. The fact that it is realized phonetically as [a] and not [æ] or [ɑ] or [ɒ] is due to other principles. The vowel designated as /i/ in (2) is fundamentally different from the /i/ in (5). The former is contrastively [coronal]; this [coronal] feature is also part of its phonetic realization as [i]. The /i/ in (5) does not have a contrastive [coronal] feature: its contrastive characterization is purely negative, as Trubetzkoy would put it. It is not [low] and it is not [labial]. Phonologically, then, the vowel in (2) that most closely corresponds to /i/ in (5) is /ə/, not /i/. Why, then, does this vowel surface as [i] and not as [ə] or [ɨ]?

7.2.3.1 Phonetic enhancement

Stevens, Keyser and Kawasaki (1986) and Stevens and Keyser (1989) propose that phonological contrasts can be *enhanced* by phonetic specification of non-contrastive features. These enhancements serve either to increase the perceptual salience of the contrastive feature, or to increase the perceptual salience of a contrast. The notion of enhancement was adopted by MCS (Avery and Rice 1989; Rice 1993, 1996; Wu 1994; Dyck 1995; Causley 1999). In the three-vowel system in (5), the contrastively non-low vowels are enhanced by the feature [high], the contrastively [labial] vowel is enhanced by [back], and the non-labial non-low vowel is enhanced by the place feature [coronal].

Enhancement thus also partly accounts for why certain inventories are more common than others; why, for example, /i, a, u/ is more common than /ə, a, u/ or /ɨ, a, u/ (on this, see, further, section 8.3).

7.2.3.2 Phonetics in phonology versus phonological minimalism

MCS was developed in the context of works that argue that the phonology is underspecified with respect to phonetics; in addition to the papers on enhancement mentioned above, these include Keating (1988), Cohn (1993) and Lahiri and Reetz (2002). Kingston and Diehl (1994) also argue that an elaborate phonetic component is required to complement the phonology.

These were followed by proposals that aim to diminish or eliminate the distance between phonetics and phonology by arguing that noncontrastive phonetic features play a role in phonology (e.g., Kirchner 1997, 1998; Steriade 1997; Boersma 1998; Flemming 2001, 2002), and that much that goes on in phonology is sensitive to detailed phonetic information (Pierrehumbert, Beckman and Ladd 2000; Hayes, Kirchner and Steriade 2004). It should be emphasized that the Contrastivist Hypothesis does not require lexical representations to be free of redundancy. As argued in chapter 2, the aim of the SDA is not to eliminate logical redundancy, but to identify contrastive specifications. Nor is anything in the theory based on an assumption that the brain has limited storage capacity. The claim that the phonological component assigns a special role to contrastive specifications is an empirical hypothesis formulated in order to account for phonological patterning. It is not a question of what memories may be stored in the brain, but of how the phonology is organized.

Reacting to the proliferation of phonetic detail in recent approaches to phonology, some phonologists have been exploring minimalist theories of phonological representation (Hyman 2001, 2002, 2003; Morén 2003, 2006). Clements (2001, 2003, 2009) has argued that feature economy plays a role in accounting for phonological inventories (see D. C. Hall 2007 for discussion). Avery and Idsardi (2001) argue for representational economy and underspecification drawing on evidence from laryngeal systems. Versions of phonological minimalism can be found in other phonological traditions as well, such as Dependency Phonology (J. M. Anderson and Ewen 1987; J. M. Anderson 2005; Carr, Durand and Ewen 2005) and Radical CV Phonology (van der Hulst 1995, 1996, 2005).

7.3 Ubiquitous feature hierarchies

The theory of the contrastive feature hierarchy makes two empirical claims. The first claim is that distinctive features in each language are organized into a hierarchy. The second claim is that this hierarchy determines which feature values are contrastive in a given language. In this section I will focus on the first claim and argue that feature hierarchies are widespread in the practice, and in many cases also the theory, of phonology. So ubiquitous are they that it is impossible to avoid considerations of feature ordering in almost any phonological analysis. This should not be surprising; as Halle (1970) remarks, ‘Any structure of any kind of complexity presupposes some form of hierarchy.’

7.3.1 Feature hierarchies in phonological theory

The ordering of features into hierarchies is remarkably pervasive in phonology, even where it is not acknowledged explicitly, and even where one might not be aware of it. I have shown that hierarchies, often implicit, were central to structuralist phonology. In generative phonology, we have seen that feature hierarchies are embedded into markedness theory, underspecification theory and feature geometry. Even theories that have not adopted some version of the Contrastivist Hypothesis have required feature hierarchies.

Feature hierarchies are pervasive in Optimality Theory, in the ranking of faithfulness constraints. For example, the tableau in (6) represents a portion of the OT grammar proposed by Baković (2000) for Nez Perce, which will be discussed in section 7.4.4. In this grammar it is more important to preserve underlying values of [low] than of [back]; similarly, [back] is ranked above [ATR].[10]

(6) OT grammar with a feature hierarchy: [low] > [back] > [ATR]

/–lo, –bk, –ATR/ = ɛ	IDENT [lo]	*[–bk, –ATR]	IDENT [bk]	*[–lo, –ATR]	IDENT [ATR]
a. [+lo, –bk, +ATR] = æ	*!				*
b. [–lo, –bk, –ATR] = ɛ		*!		*	
c. [–lo, +bk, +ATR] = u			*!		*
d. ☞ [–lo, –bk, +ATR] = i					*

10 Baković uses the faithfulness constraint IDENT where I have used MAX. Though there are technical differences between the two that have empirical consequences in certain situations, for our purposes we can regard them as interchangeable: in the type of case discussed here, both require that an underlying specification be preserved.

In this example, an input segment /−low, −back, −ATR/ (say, the vowel /ɛ/) surfaces as [i]. Because of the highly ranked constraint IDENT [low], an input vowel must retain its feature [−low] (hence candidate (a) is excluded); the 'faithful' output [ɛ] is excluded because there is a constraint against being [−back] and [−ATR] at the same time; and candidate (c) does not preserve the underlying [−back] feature, violating IDENT [back]; hence, all these candidates lose to (d).[11] The point is that any ranking of faithfulness constraints implies a feature ordering. Thus, the question of whether and how features are ordered in a grammar is independent of the Contrastivist Hypothesis.

7.3.2 Implicit feature hierarchies in practice

Feature hierarchies are often implicit in at least a partial way in the descriptive practice of phonologists. Consider, for example, the way segment inventories are presented in charts in descriptive grammars. Compare the inventory tables of Siglitun (Dorais 2003: 62),[12] an Inuit (Eskimo-Aleut) language spoken in the Canadian Arctic, and Kolokuma Ị̀jọ̀ (Williamson 1965),[13] an Ijoid (Niger-Congo) language spoken in Nigeria, given in (7) and (8), respectively. I present them as they are given in the sources (with some changes to the phonetic symbols but not to the arrangement).

(7) Siglitun consonants (Dorais 2003: 62)

	Bilabial	Apical		Velar	Uvular
Stops	p	t		k	q
Voiced fricatives	v	l	j	ɣ	ʀ
Voiceless fricatives		ɬ	s		
Nasals	m	n		ŋ	

11 For the sake of concision I have omitted reference to the feature [high] from this example, and the constraints that rule out [e] in Nez Perce.

12 I have simplified Dorais's *j/dj* and *s/ch* to *j* and *s*, respectively. As he makes clear, these are variants of single phonemes.

13 Williamson notes that Back = palatal, velar or glottal, Vl. = voiceless, and Vd. = voiced. She mentions that some speakers have a marginal phoneme /ɣ/, but she omits it from the table. I have added it because it appears to be no less marginal than /h/, which is included.

(8) Consonant phonemes of Kolokuma Ịjọ (Williamson 1965)

	Plosive		Continuant				
			Fricative		Sonorant		
					Non-lateral		Lateral
	Vl.	Vd.	Vl.	Vd.	Oral	Nasal	
Labial	p	b	f	v	w	m	
Alveolar	t	d	s	z	r	n	l
Back	k	g	(h)	(ɣ)	j	ŋ	
Labio-velar	kp	gb					

Note in particular the different placements of /l/ and /j/ in these charts. The Siglitun chart is not as overtly hierarchical as the one for Ịjọ, but it is clear that the feature which characterizes /l/ and /ɬ/ (presumably [lateral]) has very narrow scope, confined to making distinctions among apicals, whereas [nasal] is higher in the hierarchy. Thus, in the Siglitun chart /l/ and /j/ are 'partners', as are /ɬ/ and /s/. The non-nasal sonorants are not set apart in Siglitun, suggesting that the feature [sonorant] is lower in the hierarchy than in Ịjọ.

The chart for Ịjọ expresses a hierarchy in which the feature [continuant] has wider scope than [sonorant] and [voiced], and [lateral] has wider scope than [nasal]. Now /j/ and /ŋ/ are 'partners', and /l/ stands apart. The Ịjọ chart groups all post-alveolar consonants, including /j/, together under the general place 'back', whereas the Siglitun chart distinguishes between velar and uvular places, and groups /j/ with the apicals.

These sorts of examples can be multiplied indefinitely. Descriptive phonologists display phoneme inventories in ways that illuminate their phonological patterning, and these patterns attest to different hierarchical relations among features. This is not to say that any feature hierarchy is equally likely, or even permitted. It is possible to create rather unnatural-looking phoneme inventories by ordering the features in unusual ways. For example, the chart in (9) is wrong as a description of the Siglitun inventory, and probably wrong for any language with a comparable set of consonant phonemes. What is unusual are the relative scopes of place features and manner features. Such examples show that there are limits to the extent that the feature hierarchy can vary cross-linguistically, though it is not clear what these limits are.

(9) Siglitun consonants: unusual feature hierarchy

	Labial	Coronal	Velar	Uvular
Oral stops	p	t	k	q
Nasal stops	m	n	ŋ	
Fricatives	v	ɣ		ʀ
Stridents		s		
Voiced laterals		l		
Voiceless laterals		ɬ		
Glides		j		

7.3.3 Feature scope ambiguities

Finally, it is impossible to escape having to make decisions about the scope of a feature. Such decisions are not always thought of as involving feature ordering, but they do, because the relative scope of features can be expressed in terms of ordering. Deciding on the scope of a feature is particularly important when there are asymmetries in a phoneme inventory.

Eastern and Valencian Catalan, for example, have seven stressed vowels /i, e, ɛ, a, ɔ, o, u/. Analysts agree that /i, u/ are high and /a/ is low; assuming that the main contrast between the mid vowels is [ATR], it must be decided whether this feature is confined to the mid vowels, or if it extends to include the high and low vowels as well.[14] One can find both kinds of analysis: Crosswhite (2001) assumes that [ATR] is confined to the mid vowels (10a), whereas Walker (2005) and Lloret (2008) assign values of [ATR] to all the vowels (10b).

(10) Two analyses of Catalan vowels

a. Eastern Catalan (Crosswhite 2001)

		[+front]		[–front]
[+high]		i		u
	[+ATR]	e		o
	[–ATR]	ɛ		ɔ
[+low]			a	

14 Not all authors use a binary [ATR] feature to characterize this system. Jiménez (1998) uses [RTR] instead of [−ATR]. Wheeler (2005: 56) characterizes /e, o/ as [+close] and /ɛ, ɔ/ as [−close]. Whatever feature is chosen, questions of scope arise (unless it is limited by definition to certain types of phonemes: if, for example, [close] is defined to be applicable only to mid vowels).

b. Valencian Catalan (Walker 2005; Lloret 2008)

		[front]		[back]
[+ATR]	[high]	i		u
		e		o
[–ATR]		ɛ		ɔ
	[low]		a	

In the Catalan case we may simply be dealing with one system whose analysis is in dispute; but there are cases where there is evidence that different dialects with similar-looking inventories actually differ in their feature ordering, and hence in the relative scopes of contrastive features. We saw a number of such examples in chapter 3; here I will present one more.

Both Anywa (Reh 1996) and Luo (Tucker 1994), related Nilotic languages, have a dental/alveolar contrast in the coronal stops; in both languages, the alveolar nasal /n/ has no dental nasal partner. Should /n/ be considered contrastively alveolar (in contrast to the coronal dental stops in general), or is it outside the dental/alveolar contrast, being only redundantly alveolar? Mackenzie (2005, 2009) argues that the two languages adopt different solutions to this question: in Anywa /n/ acts as if it is contrastively alveolar (11a), in Luo it acts neutrally with respect to the contrast (11b).[15]

(11) Nilotic dental/alveolar contrast (Mackenzie 2005)

a. Anywa

Dental	Alveolar	
t̪	t	voiceless stops
d̪	d	voiced stops
	n	nasals

b. Luo

Dental	Alveolar
t̪	t
d̪	d
n	

7.3.4 Feature hierarchies and phonological patterning

Given that feature hierarchies have always been a part of phonology and are here to stay, it is worth reflecting on their significance. Notice that in all the

15 See Mackenzie (2005, 2009), for further details and the reasons for adopting this analysis. Mackenzie observes that the two consonantal inventories are not in fact identical: Luo has a set of prenasalized stops that also participate in the dental/alveolar contrast. It is possible that the slightly different compositions of these segmental inventories contribute to an explanation of why the contrastive hierarchies in these languages are different (cf. Béjar 1998 and D. C. Hall 2007: 152–3 for discussion).

examples reviewed in this section, the issue is not the *phonetic* description of the phonemes. I assume that Siglitun and Ịjọ *l*, *j*, etc. are phonetically similar enough for them to be depicted with the same symbols /l/, /j/, and so on. Any further phonetic details that may distinguish them are not in any case provided in the phoneme charts, and it is unlikely that phonetic details are what account for their different placements in the charts. The same is true of the Catalan vowels and Nilotic dental and alveolar consonants. No one disputes that Catalan *i* and *u* are phonetically ATR, or that Nilotic *n* is phonetically alveolar; the question in each case is whether they function *phonologically* as though they are specified for these features. It follows from the Contrastivist Hypothesis that this amounts to asking whether they are *contrastively* specified for the features in question.

7.3.5 Feature hierarchies: summary

I mentioned at the outset of this section that contrastive hierarchy theory claims that features are ordered, and that this ordering determines the contrastive specifications for a language. I argued in chapter 2 and throughout that feature ordering is the best way to determine contrastive specifications. One might think that having to order the features just for purposes of determining contrast is an unacceptable cost, for it imposes on learners and analysts the burden of arriving at the correct feature ordering. What I have tried to show in this section is that this burden exists independently of any particular theory of contrast. Feature ordering in one form or another is a characteristic of all theories of phonology.

In the rest of this chapter I provide further evidence for the second claim of contrastive hierarchy theory and for the Contrastivist Hypothesis: the feature hierarchy is what determines contrastive features, and only contrastive features are active in the phonology.

7.4 Vowel harmony

An important diagnostic of phonological activity is the spreading of a feature from a segment bearing that feature to neighbouring segments. In this sense, vowel harmony is a fairly reliable indicator of the presence of an active feature or features. The Contrastivist Hypothesis states that phonologically active features are contrastive; the corollary of this principle in the domain of vowel harmony is that harmony triggers should be contrastive features.[16]

16 The situation is less clear with targets and other non-triggers, such as transparent and opaque segments. There are various reasons why segments may block harmony, not all derived from their contrastive status. Similarly, targets may be restricted for reasons beyond their contrastive

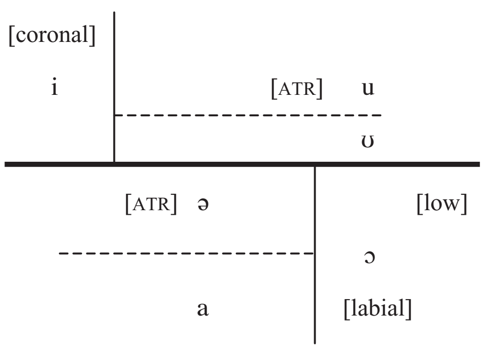
In this section I will review the analysis of Classical Manchu by Zhang (1996), and show that this principle applies in the case of both ATR harmony and labial harmony. The analysis is strikingly supported by diachronic developments in Manchu, showing how changes in the contrastive status of vowels leads to new patterns of phonological activity, as predicted. The analysis of labial harmony is further supported by typological surveys of labial harmony: in all cases, the vowel triggering the harmony can be shown to be contrastive.

To complete this section, I will review an OT analysis of Nez Perce harmony by Baković (2000). Like Classical Manchu, Nez Perce has ATR harmony; however, Baković's analysis appears to suggest that the spreading [ATR] feature is not contrastive. If this analysis were correct it would be a problem for the Contrastivist Hypothesis. Mackenzie (2002) shows, however, that this aspect of the analysis has no empirical support; in her analysis, [ATR] is clearly contrastive.

7.4.1 Classical Manchu (Zhang 1996)

The vowel system of Classical Manchu as analysed by Zhang (1996) was presented in section 6.4.3, where it was used to demonstrate how to convert a feature hierarchy into an OT hierarchy.[17] The vowel system and the proposed contrastive features are shown again in (12). Recall that Zhang (1996) proposes the feature hierarchy: [low] > [coronal] > [labial] > [ATR]. The resulting specifications are shown in (12) in chapter 6, and a corresponding OT feature hierarchy is given in (16) in chapter 6. Here I will focus on the motivation for, and empirical consequences of, this analysis.

(12) Classical Manchu vowel system (Zhang 1996)



status; see van der Hulst and van de Weijer (1995) and Archangeli and Pulleyblank (2007) for surveys. Here I will focus on harmony triggers.

17 The data and analysis of Manchu and related languages in this and the following sections are based on Zhang (1996); see Zhang (1996: 32) for discussion of the transcriptions and phonetic values. See also Zhang and Dresher (1996), Dresher and Zhang (2004), Zhang and Dresher (2004) and Dresher and Zhang (2005) for more details.

The first contrast applies to all the vowels and divides them into a [low] (/a, ə, ɔ/) and non-low (/i, u, ʊ/) set. Splitting the inventory in this manner has the effect of allowing for different contrasts in each set.

The next feature, [coronal] (= [−back]), applies only to /i/. This has a number of important empirical consequences. First, it means that /i/ will receive no further contrastive specifications; notably, it does not receive a specification for [ATR], despite the fact that it is phonetically ATR. Second, because the non-low vowels have already been split from the low vowels, no further features are required to draw a contrast between /i/ on one side and /u/ and /ʊ/ on the other. In particular, these latter vowels have no contrastive specification for [labial] (= [round]), despite the fact that they are phonetically rounded vowels.

No low vowels are eligible to receive a [coronal] specification, so [coronal] has no contrastive effect on the low vowels. Hence, a feature is needed to distinguish between /ɔ/ on one side and /a/ and /ə/ on the other. This contrast is made by the next feature on the list, which is [labial]. The consequence of this is that /ɔ/ is the only contrastively [labial] vowel in the inventory.

The next feature is [ATR]. It distinguishes /u/ from /ʊ/ and /ə/ from /a/. This feature also accounts for the height differences in these pairs of vowels, as [ATR] vowels tend to be higher than their non-ATR counterparts (Archangeli and Pulleyblank 1994). The vowels /i/ and /ɔ/ are not contrastive for this feature, in this ordering.

The evidence for these contrastive specifications is summed up in (13).

(13) Summary of evidence for contrastive specifications of Classical Manchu vowels

a. /i/ lacks contrastive [ATR]: /u/ and /ə/ trigger ATR harmony, but /i/ does not, though /i/ is phonetically [ATR].
b. /u/ and /ʊ/ lack contrastive [labial]: /ɔ/ triggers labial harmony, but /u/ and /ʊ/ do not, though they are phonetically [labial].
c. /i/ is contrastively [coronal]: /i/ triggers palatalization of consonants, suggesting it has some relevant contrastive feature.
d. All vowels are contrastively assigned to one of two height classes: alternations /ə/ ∼ /a/ ∼ /ɔ/ and /u/ ∼ /ʊ/ are limited to a height class. We need one height feature, which we call [low].

A more detailed discussion of the harmony facts follows.

7.4.1.1 ATR harmony

All vowels in a word, apart from /i/, must agree with respect to [ATR], as shown in (14).

(14) ATR harmony in Classical Manchu

a. /ə/ alternates with /a/

xəxə	'woman'	xəxə-ŋgə	'female'
aɢa	'rain'	aɢa-ŋɢa	'of rain'

b. /u/ alternates with /ʊ/

xərə-	'ladle out'	xərə-ku	'ladle'
paqt'a-	'contain'	paqt'a-qʊ	'internal organs'

The alternation between /u/ and /ʊ/ is apparent only after back (velar and uvular) consonants (which also alternate, depending on the following vowel). In other contexts, /u/ and /ʊ/ merge at the surface as [u], except in a few sporadic examples. Zhang (1996) assumes that this is a late phonetic rule, since it does not affect the behaviour of /ʊ/ with respect to ATR harmony, as shown in (15).

(15) Merger of /ʊ/ to [u] except after back consonants

a. Underlying /u/: ATR harmony

susə	'coarse'	susə-tə-	'make coarsely'
xət'u	'stocky'	xət'u-kən	'somewhat stocky'

b. Underlying /ʊ/ not after velar/uvular consonants

tulpa	'careless'	tulpa-ta-	'act carelessly'
tat'ʂ̺un	'sharp'	tat'ʂ̺u-qan	'somewhat sharp'

In each word in (15b) the vowel that surfaces as [u] patterns with non-ATR vowels; compare the forms in (15a). I suppose, following Zhang (1996), that [u] in (15b) derives from /ʊ/, which merges with /u/ as [u] in these environments.

The vowel /i/ is neutral, as shown in (16). It can co-occur in roots with both ATR and non-ATR vowels and with both ATR and non-ATR suffixes (16a, b), and it can itself appear in a suffix following either ATR or non-ATR vowels (16c).

(16) ATR harmony in Classical Manchu: /i/ is neutral

a. /ə/ ∼ /a/ suffix

pəki	'firm'	pəki-lə	'make firm'
paqtʂ̺'in	'opponent'	paqtʂ̺'i-la-	'oppose'

b. /u/ ∼ /ʊ/ suffix

sitərə-	'hobble'	sitərə-sxun	'hobbled/lame'
panjin	'appearance'	panji-sχʊn	'having money'

c. /i/ suffix

əmt'ə	'one each'	əmt'ə-li	'alone; sole'
taχa-	'follow'	taχa-li	'the second'

Surprisingly, when /i/ is in a position to trigger harmony, it occurs only with non-ATR vowels, as in (17).

(17) Stems with only /i/: suffixes with non-ATR vowels

a. /a/ in suffix, not /ə/

fili	'solid'	fili-qan	'somewhat solid'
ili-	'stand'	ili-χa	'stood'

b. /ʊ/ in suffix, not /u/

sifi-	'stick in the hair'	sifi-qʊ	'hairpin'
tʂ'ili-	'to choke'	tʂ'ili-qʊ	'choking'

Despite the fact that it is phonetically an ATR vowel, /i/ does not trigger ATR harmony. This fact is explained if we posit the contrastive specifications in (12), together with the hypothesis that only contrastive values of [ATR] trigger harmony.

The failure of /i/ to trigger ATR harmony is particularly striking given the observation that front high vowels tend to be associated with [ATR], because the gestures required to make a high front vowel are compatible with an advanced tongue root and antagonistic to a retracted tongue root (see Archangeli and Pulleyblank 1994 for discussion and references). While this tendency can account for why /i/ lacks an [RTR] partner, we would expect that /i/, as the ATR vowel par excellence, should trigger ATR harmony if any vowel does. The fact that it does not strengthens the argument that its contrastive status is the key to its neutrality.

7.4.1.2 Labial harmony

Another vowel harmony process in Classical Manchu is labial harmony. A suffix vowel /a/ becomes /ɔ/ if preceded by two successive /ɔ/ vowels (18a), but labial harmony is not triggered by a single short or long /ɔ/ (18b).[18] Nor is labial harmony triggered by the high round vowels /u, ʊ/ (18c, d). As with ATR harmony, only a contrastive feature can serve as a harmony trigger. In this case, only /ɔ/, but not /u/ or /ʊ/, has a contrastive [labial] feature.

(18) Labial harmony in Classical Manchu

a.	pɔtʂ'ɔ	'colour'	pɔtʂ'ɔ-ŋɢɔ	'coloured'
	fɔχɔlɔn	'short'	fɔχɔlɔ-qɔn	'somewhat short'
b.	tɔ-	'alight (birds)'	tɔ-na-	'alight in swarm'
	tɔɔ-	'cross (river)'	tɔɔ-na-	'go to cross'
c.	gulu	'plain'	gulu-kən	'somewhat plain'
	kumun	'music'	kumu-ŋgə	'noisy'
d.	χʊtun	'fast'	χʊtu-qan	'somewhat fast'
	tursun	'form'	tursu-ŋɢa	'having form'

18 On this condition, see Zhang and Dresher (1996) and Walker (2001).

7.4.2 Evolution of Spoken Manchu and Xibe

This analysis of Classical Manchu receives additional support from subsequent developments in the modern Manchu languages. The vowels /ə/ and /u/ undergo changes in their contrastive status, leading to new patterns of phonological activity.

We observed that in Classical Manchu the contrast between /u/ and /ʊ/ is neutralized phonetically to [u] in most contexts, with surface [ʊ] surviving only after uvular consonants and sporadically in other contexts in a few words. It is no surprise, therefore, to see this neutralization continue to completion in Spoken Manchu, a modern Manchu language descended from an ancestor similar to Classical Manchu. In Spoken Manchu, /u/ and /ʊ/ have merged completely to [u], and the phoneme /ʊ/ has been completely lost.

In a contrast-driven approach to vowel systems, the loss of a contrast in one part of the system could have wider effects. In the Classical Manchu system, the contrast between /u/ and /ʊ/ involves the feature [ATR], just like the contrast between /ə/ and /a/. As long as the [ATR] contrast between /ə/ and /a/ is paralleled by a similar contrast between /u/ and /ʊ/, it cannot be mistakenly regarded as a height contrast. But with the loss of /ʊ/, the position of [ATR] in the system becomes much more tenuous. For now the entire burden of the [ATR] contrast would fall on the contrast between /ə/ and /a/. That this contrast is based on [ATR], however, is not clear; it could more straightforwardly be attributed to a difference in height. Indeed, the feature [low], which is required independently, can serve to distinguish /ə/ from /a/.

Therefore, without assuming that the phoneme /ə/ changed phonetically, the loss of /ʊ/ could have indirectly led to a change in the phonological status of /ə/, from [low] to non-low. This reclassification, in turn, could have influenced the phonetic realizations of /ə/, because in Spoken Manchu it is definitely a non-low vowel. Zhao (1989) characterizes it as a mid-high back unrounded vowel, with an allophone [ɤ]; according to Ji, Zhao and Bai (1989), [ə] is in free variation with a high back unrounded vowel [ɯ]. It is reasonable to suppose that there is a mutual influence between phonology and phonetics in such cases. The phonetic properties of a vowel obviously influence its phonological representation; but this influence is not simply one way, and the phonological representation can in turn affect the phonetics, by delimiting the space within which the vowel can range (short of neutralization).

The change in status of /ə/ in turn has consequences for the specification of /u/. Recall that in Classical Manchu there is evidence that /i/ is actively [coronal], but there is no evidence that /u/ and /ʊ/ are actively [labial], though they clearly are phonetically round. The elevation of /ə/ to a non-low vowel, joining /i/ and /u/, changes the situation. Assuming, as before, that [coronal] takes precedence, /i/ is again specified [coronal], distinguishing it from /ə/ and /u/. But now we must still distinguish the latter two vowels from each other. The most straightforward distinction is to extend the feature [labial], already in the system for /ɔ/, to /u/, as diagrammed in (19).

(19) Spoken Manchu after loss of /ʊ/

[coronal]		[labial]	
i	ə	u	
	a	ɔ	[low]

This analysis thus predicts that the reclassification of /ə/ as a non-low vowel should cause /u/ to become contrastively [labial]. This prediction is borne out in Spoken Manchu, as evidenced by the development of a new phoneme /y/, a front rounded vowel that originated as a positional allophone of /i/ followed by /u/, as well as /u/ followed by /i/ (Zhang 1996). The front feature [coronal] is contributed by /i/, but the round feature [labial] must be contributed by /u/.

Further evidence can be found in the related modern Manchu language Xibe. Unlike Spoken Manchu, Xibe retains a labial harmony rule in which /ə/ alternates with /u/ in suffixes: /u/ occurs if the stem-final vowel is round (20b, c), /ə/ occurs otherwise (20a).

(20) Labial harmony in Xibe (Li and Zhong 1986)

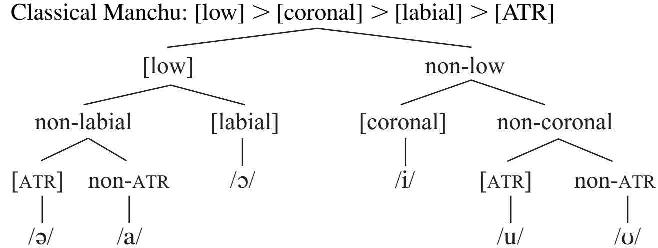
	Classical Manchu	*Xibe*	*Gloss*
a.	gət’ə-xə	gət’ə-xə	‘awoke’
	ɢɔtʂ̡’i-χa	ɢɔçi-χə-	‘cherished’
	ərtə-kən	ərtə-kən	‘somewhat early’
	χantʂ̡i-qan	χantçi-qən	‘somewhat near’
b.	pɔtɔ-χɔ	pɔtu-χu	‘thought’
	fɔχɔlɔ-qɔn	fœχulu-qun	‘somewhat short’
c.	pu-xə	pu-xu	‘gave’
	ḋuʂ̡uxu-kən	dʑyçxu-kun	‘somewhat sour’
	xət’u-kən	xət’u-kun	‘somewhat stocky’
	farχʊ-qan	farχu-qun	‘somewhat dark’

Recall that in Classical Manchu, labial harmony is restricted to the low vowels, and creates an alternation between /a/ and /ɔ/. In Xibe, noninitial vowels tended to be raised – almost always in suffixes, frequently in stem vowels – so an original sequence of the form /a/ - /a/ would become /a/ - /ə/ or /ə/ - /ə/, and a sequence of the form /ɔ/ - /ɔ/ would become /ɔ/ - /u/ or /u/ - /u/. The labial harmony observed in Xibe is not merely a holdover of Classical Manchu labial harmony, however, for in Xibe harmony is triggered not only by /u/ derived from older /ɔ/ (20b), but also by original /u/ (20c). The fact that /u/ triggers and undergoes labial harmony further supports the hypothesis that it has a [labial] specification in Xibe.

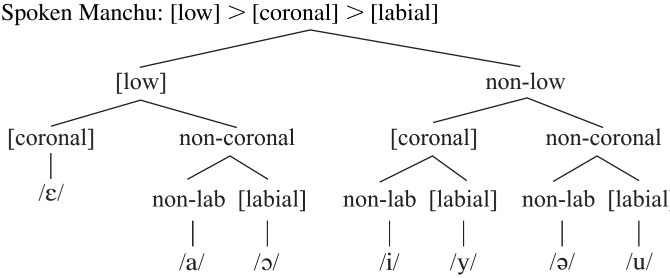
The contrastive hierarchies of the three Manchu languages discussed above are given in (21). Spoken Manchu and Xibe make do with only three contrastive features. They have more vowel phonemes than Classical Manchu because they exploit the possibilities of the three features more fully. For more on the modern Manchu vowel systems and the development of new coronal vowels, see Zhang (1996) and Dresher and Zhang (2005).

(21) Contrastive hierarchies of Manchu languages

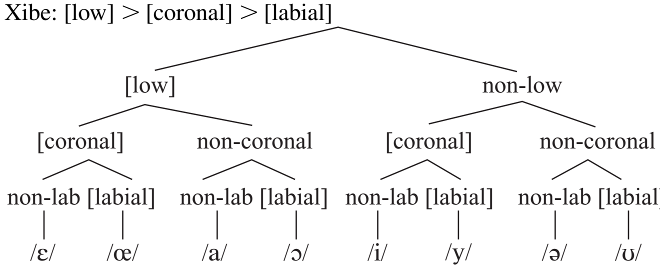
a. Classical Manchu: [low] > [coronal] > [labial] > [ATR]



b. Spoken Manchu: [low] > [coronal] > [labial]



c. Xibe: [low] > [coronal] > [labial]



7.4.3 Typological surveys of labial harmony

Typological surveys of labial harmony in Manchu-Tungus, Mongolian and Turkic languages support the hypothesis that only contrastive features trigger harmony. Zhang (1996: chapter 6) surveys a number of Manchu and Tungusic languages in China and Russia. We have seen that labial harmony in Classical Manchu is limited to the low vowels. On this account, only the low vowel /ɔ/ is contrastively [labial] in this inventory. The same holds for most Manchu-Tungus languages, which have similar vowel inventories. A Tungusic example is Oroqen (Zhang 1996), whose inventory is given in (22); again, only low vowels are triggers and targets of harmony.

(22) Oroqen vowel system (Zhang 1996)

[coronal]			
i ii		u uu ʊ ʊʊ	
e ɛ	ə əə a aa	o oo ɔ ɔɔ [labial]	[low]

Eastern Mongolian languages have a similar type of labial harmony triggered by and affecting low vowels. An example is Khalkha Mongolian (Svantesson 1985; Kaun 1995), shown in (23).[19] Assuming a similar contrastive hierarchy to that of Manchu-Tungus, again [labial] is contrastive only in the low vowels.

(23) Khalkha Mongolian vowel system (Svantesson 1985; Kaun 1995)

[coronal]			
i		u ʊ	
	ə a	o ɔ [labial]	[low]

Turkic languages tend to have symmetrical vowel inventories. They are typically analysed with three features: one height feature and two place features. A typical example is Turkish, shown in (24). Assuming three features, [high], [coronal] and [labial] (or their equivalents), the Turkish vowels exhaust the space of possible values. Therefore, all feature values are contrastive; in particular, [labial] is necessarily contrastive in all vowels that are rounded on the surface.

19 See Dresher and Zhang (2005) for further discussion of the phonemic values of the Khalkha Mongolian vowels.

(24) Turkish vowel system

	[coronal]		non-coronal	
	non-labial	[labial]	non-labial	[labial]
[high]	i	y	ɨ	u
[low]	e	ø	a	o

The theory predicts, therefore, that all round vowels could potentially be triggers of labial harmony in such languages. This prediction is correct, though harmony observes limitations that are not due to contrast, but to other factors. That is, having a contrastive feature is a necessary but not sufficient condition for triggering harmony. We find a variety of labial harmony patterns, where high vowels are favoured as triggers and targets, for reasons unrelated to contrast (Korn 1969; Kaun 1995).

In Turkish, for example, harmony triggers can be high or low, but targets are typically limited to high vowels. In Kachin Khakass (Korn 1969), with the same vowel inventory, both triggers and targets of labial harmony must be high, the opposite of the Manchu-Tungus-Eastern Mongolian pattern.

7.4.4 Nez Perce

Nez Perce, a Penutian language of the Pacific Northwest in the United States, is another language that displays ATR harmony, though of a different character than the Manchu type.[20] The differences (directional harmony in Manchu versus a dominant harmony in Nez Perce) do not change our expectation that [ATR] should be a contrastive feature. An analysis by Baković (2000) seems to put this assumption in doubt. It is interesting, therefore, that this analysis does not appear to be as empirically adequate as an alternative in which [ATR] is a contrastive feature.

7.4.4.1 The Nez Perce vowel system

The surface vowels of Nez Perce (Aoki 1966, 1970) are shown in (25).

(25) Nez Perce surface vowels (Aoki 1966, 1970)

i			u
			ɔ
	æ	ɑ	

Nez Perce has dominant-recessive ATR harmony (B. L. Hall and Hall 1980). All vowels in a word, apart from /i/, must agree with respect to [ATR], and the

20 This section is based on Mackenzie (2002) and Mackenzie and Dresher (2004).

value [−ATR] is dominant. That is, a [−ATR] specification anywhere in the word causes all [+ATR] vowels in the word to become [−ATR]. The vowel /æ/ alternates with /ɑ/ (26) and /u/ alternates with /ɔ/ (27).

(26) ATR harmony: /æ/ alternates with /ɑ/

	Underlying	*Surface*	*Gloss*
a.	/næʔ-mæq/	næʔmæχ	'my paternal uncle'
b.	/næʔ-tɔt/	nɑʔtɔt	'my father'
c.	/mæq-æʔ/	mæqæʔ	'uncle VOC'
d.	/tɔt-æʔ/	tɔːtɑʔ	'father VOC'
e.	/cæqæt/	cæːqæt	'raspberry'
f.	/cæqæt-ɑyn/	cɑqɑːtɑyn	'for a raspberry'

(27) ATR harmony: /u/ alternates with /ɔ/

	Underlying	*Surface*	*Gloss*
a.	/tæwæː-puː/	tæwæːpuː	'the people of Orofino, Idaho'
b.	/sɔːyɑː-puː/	sɔːyɑːpɔː	'the white people'
c.	/tuʔuynu/	tuʔuynu	'tail'
d.	/tuʔuynu-ʔɑyn/	tɔʔɔynɔʔɑyn	'for the tail, crupper'

As illustrated in (28), the vowel /i/ sometimes patterns with [−ATR] vowels (28a, b), and other times with [+ATR] vowels (28c, d), though it is phonetically [+ATR].

(28) Dual patterning of /i/

	Underlying	*Surface*	*Gloss*
a.	/næʔ-ciːc/	nɑʔciːc	'my paternal aunt'
b.	/ciːc-æʔ/	ciːcɑ	'paternal aunt VOC'
c.	/næʔ-iːc/	næʔiːc	'my mother'
d.	/ʔiːc-æʔ/	ʔiːcæʔ	'mother VOC'

Following Jacobsen (1968), Rigsby and Silverstein (1969), Zwicky (1971) and B. L. Hall and Hall (1980), Mackenzie (2002) assumes that surface [i] represents a merger of /i/ and a [−ATR] vowel that can be represented as /ɛ/. In (28a, b) the underlying stem vowel is [−ATR] /ɛ/ and in (28c, d) it is [+ATR] /i/.[21] Thus, every vowel has a counterpart that contrasts with it in the feature [ATR].

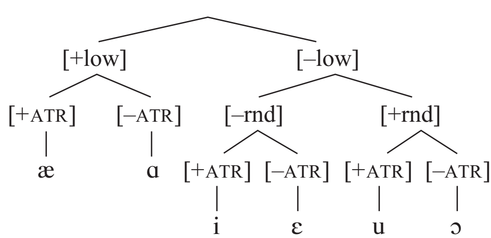
21 In its dual behaviour, Nez Perce /i/ is quite different from Classical Manchu /i/. It is more like Classical Manchu /u/, which represents the surface merger of two underlying vowels when not following a velar or uvular consonant.

(29) Nez Perce underlying vowels

i	[+ATR]	u
ɛ	[−ATR]	ɔ
[+ATR]		[−ATR]
æ		ɑ

By any definition, [ATR] would appear to be a contrastive feature in the underlying vowel system of Nez Perce. What are the other contrastive features? Abstracting away from [ATR], we have a classic three-vowel system, which we can designate /I, A, U/. In such systems it is usual to have a height feature, either [low] or [high], and a place feature, either [back] or [round].[22] The feature [low] is a better choice than [high] because the non-low [ATR] pairs are not strictly [+high], whereas the low pair are both [+low]. Following Jakobson and Halle's (1956) assumption that a contrast between high and low sonority is, preferably, ordered before one based on place (but see Ghini 2001b for a different view), let us order [low] as the first feature. For the second feature, either [round] or [back] are possible; for concreteness, we will pick [round]. This contrast is relevant only among the non-low vowels.[23] Because of the symmetry of the system, it does not matter very much where [ATR] is ordered. For concreteness, we will assume it is ordered third. We thus arrive at the contrastive hierarchy illustrated in (30).

(30) Nez Perce: [low] > [round] > [ATR] (Mackenzie 2002)



22 The feature names are chosen to facilitate comparison with Baković's analysis.

23 If one knew nothing about the phonological patterning of Nez Perce and looked only at the underlying vowel system as pictured in (29), one might think it could be analysed the way Jakobson (1962b [1931]) analysed Standard Slovak (see section 1.1), as three pairs of vowels arrayed into three height classes where each pair is distinguished on the front/back dimension. Such an analysis for Nez Perce totally fails as a basis for capturing the facts of vowel harmony.

7.4.4.2 The analysis of Baković (2000)

An OT analysis of the Nez Perce vowel system is given by Baković (2000). His analysis has some properties in common with Mackenzie's, namely, a hierarchy of featural faithfulness constraints, and constraints to exclude certain combinations of features. However, he arrives at quite different results. Baković arrives at the ranking shown in (31). He proposes that these faithfulness constraints and co-occurrence restrictions are sufficient to exclude non-existent vowels and to ensure that vowels present in the inventory will surface faithfully.[24]

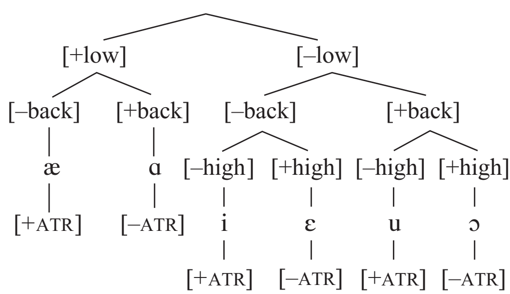
(31) Constraint ranking for Nez Perce (Baković 2000)
*[+back, +ATR] & IO-IDENT [ATR], IO-IDENT [low], *[−back, −ATR] ≫ IO-IDENT [back] ≫ *[−high, +ATR], *[+high, −ATR] ≫ IO-IDENT [high] ≫ *[+back, +ATR] ≫ *[−low, −ATR] ≫ IO-IDENT [ATR]

Looking only at the faithfulness constraints, we find the hierarchy in (32).

(32) Ranking of faithfulness constraints (Baković 2000)
IO-IDENT [low] ≫ IO-IDENT [back] ≫ IO-IDENT [high] ≫ IO-IDENT [ATR]

This constraint hierarchy translates into an ill-formed contrastive hierarchy (33). Features written under the phonemes are redundant.

(33) Contrastive hierarchy based on (32)



The feature [ATR] is phonologically redundant in this hierarchy, though it is the active feature in vowel harmony. It is redundant in the [−low] vowels because of the presence of [high], which does not appear in the contrastive

24 *[+back, +ATR] & IO-IDENT [ATR] is a conjoined constraint. A form violates it only if it contravenes both constituent constraints. Thus, the constraint against a vowel that is [+back, +ATR] is ranked fairly low to allow [u] to surface. The individual constraint IO-IDENT [ATR] is also ranked low, allowing input vowels to surface with a different value for [ATR], if so compelled by higher-ranking constraints. But if an input vowel is [−ATR], a candidate that is [+back, +ATR] will violate the high-ranking conjoined constraint. Thus, a hypothetical [−ATR] input vowel /ʊ/ may not surface as [u], in this analysis.

hierarchy we arrived at earlier; in the [+low] vowels, it is redundant because it is ordered below [back]. How did Baković (2000) arrive at this feature ranking?

Starting from the assumption that inputs are not restricted to language-specific inventories (richness of the base, Prince and Smolensky 2004), Baković introduces constraints to derive the surface inventory. For example, Nez Perce has no vowel [o], that is, a vowel with the features [−low, −high, +round, +back, +ATR]. Baković (2000: 245) proposes that an input vowel with these features will surface as [ɔ]. To ensure this result, a high ranking of faithfulness to [high] is required, as shown in (34).

(34) Role of IO-Ident [high] (Baković 2000)

Input /o/	*[–high, +ATR]	IO-Ident [high]	IO-Ident [ATR]
a. o	*!		
b. u		*!	
c. ☞ ɔ			*

However, Baković adduces no evidence that an input /o/ does in fact surface as [ɔ] and not as [u], or, for that matter, as [i] or [æ]. Therefore, the relatively high ranking of this constraint has no real motivation. Mackenzie's analysis also excludes illicit vowels, though with different results from those proposed by Baković. The contrastive feature hierarchy in (30) translates into the constraint ranking in (35). This ranking also prevents an input /o/ from surfacing as [o], as shown in (36).

(35) Nez Perce constraint hierarchy based on (30)
IO-Ident [low] ≫ *[round, +low] ≫ IO-Ident [round] ≫ IO-Ident [ATR] ≫ *[F]

(36) Evaluation of /o/ with low-ranking IO-Ident [high]

Input /o/ = [–low, –hi, +rnd, +ATR]	Ident [low]	*[round, +low]	Ident [round]	Ident [ATR]	*[F]
a. o = [–low, –hi, +rnd, +ATR]					****!
b. ☞ u = [–low, +rnd, +ATR]					***
c. ɔ = [–low, +rnd, –ATR]				*!	***
d. i = [–low, –rnd, +ATR]			*!		***

Recall that, unlike Baković, I am assuming that the output of the constraints corresponding to the contrastive hierarchy is a segment specified only for contrastive features. Thus, an input /o/, that is, a vowel specified as [−low, −high, +round, +ATR], yields the winning output [−low, +round, +ATR], which surfaces as [u] in Nez Perce.[25] Put differently, an input /o/ is not contrastively distinct from /u/. In this case, there are two sources that both favour a high vowel as the phonetic specification of [−low, +round, +ATR]: in addition to the usual preference for /I, A, U/ over /E, A, O/, [+ATR] favours [+high], as discussed in connection with Manchu /i/.

Similarly, Baković (2000: 246) wishes to ensure that an input /e/ surfaces as [i]. In his analysis, faithfulness to [back] plays a prominent role in preventing /e/ from surfacing as *[ɔ] (37). Again, there are many other ways of excluding this vowel, and we have no empirical evidence to favour one over another. Another way is shown in (38).

(37) Role of IO-IDENT [back] (adapted from Baković 2000)

Input /e/	IDENT [low]	*[–bk, –ATR]	IDENT [back]	*[–hi, +ATR]	IDENT [high]	IDENT [ATR]
a. e				*!		
b. ☞ i					*	
c. ɛ		*!				*
d. ɔ			*!			*
e. æ	*!			*		

(38) Evaluation of /e/ without IO-IDENT [back]

Input /e/ = [–low, –hi, –bk, –rnd, +ATR]	IDENT [low]	*[round, +low]	IDENT [round]	IDENT [ATR]	*[F]
a. e = [–low, –hi, –bk, –rnd, +ATR]					****!*
b. ☞ i = [–low, –rnd, +ATR]					***
c. ɛ = [–low, –rnd, –ATR]				*!	***
d. ɔ = [–low, +rnd, –ATR]			*!	*	***
e. æ = [+low, +ATR]	*!		*		**

25 I omit the feature [back], because it would only incur an extra violation of *[F], and thus would not alter the outcome.

Of course, it is ultimately an empirical question what would happen to hypothetical input vowels /o/ or /e/ in Nez Perce. One could seek to find various kinds of evidence that would bear on this question. For example, we could investigate the fate of loanwords with such vowels when adapted into Nez Perce, or devise perception tests to see how speakers label such vowels, and so on. But in the absence of any such evidence, it is impossible to favour any particular result. Therefore, in making predictions about the fate of illicit vowels, we must be guided by our analysis to the extent that it is based on actual empirical evidence.

To conclude, the analysis presented by Baković (2000) appears to require a ranking of faithfulness constraints that is incompatible with any contrastive hierarchy for Nez Perce. Moreover, this analysis does not draw any connection between contrast and phonological activity. Given its low ranking, the feature [ATR] appears to be redundant, though it is the active feature in vowel harmony. If such an analysis were supported by evidence, it would be a counterexample to the Contrastivist Hypothesis. It is significant, therefore, that this ranking is unmotivated by any empirical facts and relies primarily on unsupported assumptions about what non-existent vowels should map to. Moreover, an alternative analysis exists that conforms to the theory being advocated here. In this analysis, the active feature [ATR], together with [low] and [round], is contrastive in the Nez Perce vowel system; there is no evidence that any other vowel features are active in this language.

7.5 Metaphony and contrast limited by domain

In chapter 2, section 7.1, I considered the possibility that contrastive features are not assigned globally to phonemes over the whole language, but may be limited by position. In such a procedure, contrasts would be evaluated separately for each distinguished position, or domain. We have not seen this kind of domain-limited contrastive evaluation yet, and in principle a number of conditions must be fulfilled for a language to allow separate contrastive domains.

First, the phonemes that occur in one domain must not have alternants in the other domain. For example, in many languages the same underlying consonant may appear in both coda and onset position. An example is the stem-final consonant in English *write*, which appears in word-final position in the uninflected form, and between vowels when a vowel-initial suffix is added, as in *writing* or *writer*. Presumably, there is a single underlying representation of the morpheme *write*, so it would be contradictory, in English, to assign different contrastive features to the stem-final /t/ when it is word final and when it is word medial.[26]

A second condition that could plausibly be put on a contrastive domain is that it should correspond to a category that has independent existence in the grammar. This condition would rule out arbitrary domains such as, for English, the set of consonants that could precede the sequence _*et* (*pet, vet, debt, set, net, yet, get*, etc.) within a word.

The conditions for having separate contrastive domains for evaluating vowels are met in Romance languages that distinguish between *stem* vowels and *desinential* vowels. Desinential vowels occur in a closed class of suffixes and do not alternate with stem vowels. Moreover, stems and desinences constitute important grammatical categories in such languages. Dyck (1995) and Frigeni (2003) argue that contrastive specifications must be assigned separately to desinential vowels in Romance dialects of Spain and Italy (Dyck) and in Campidanian Sardinian (Frigeni).

The evidence in both cases comes from *metaphony*, a type of vowel harmony in which some high desinential vowels trigger raising of some stressed vowels. It is argued that the best account of metaphony triggers in these dialects requires that we distinguish between contrastive and redundant feature specifications. More particularly, the contrasts must be assigned separately to desinential vowels. As in the cases of harmony discussed above, a vowel can trigger metaphony only if it has the appropriate contrastive feature.

7.5.1 Metaphony in Iberian and Italian Romance

In (39) are examples of metaphony in Pasiego (Montañes), as given by Dyck (1995), adapted from Penny (1969). Centralization/laxing of unstressed vowels is not shown. Desinential /u/ triggers raising of stressed /é/ to [í] and stressed /ó/ to [ú]. Stressed /í/, /ú/, and /á/ are not affected (40).

26 The assumption that morphemes have a unique lexical form is not universally held. In theories that permit multiple lexical representations, the concept of a phonological inventory would have to be rethought along with what constitutes a permissible contrastive domain.

Keeping now to the principle that morphemes have unique underlying forms, it may still be the case that children at early stages of acquisition may not yet relate different forms of morphemes. For example, they may have different lexical representations for the final consonant in *write* and the medial consonant in *writer* (since *writer* is not yet decomposed into separate morphemes, it is not accurate to label the *t* in *writer* as 'stem final' for these learners). It follows that child grammar may have separate sub-inventories and contrastive domains that would not be permitted in the adult grammar.

(39) Pasiego metaphony of /é/ and /ó/ triggered by /u/

Unmetaphonized	*Gloss*	*Metaphonized*	*Gloss*
afilit[é]ros	'needle-cases'	afilit[í]ru	'needle-case'
g[ó]rdo	'fat (neuter)'	g[ú]rdu	'fat (masculine)'
ab[jé]rtos	'open (plural)'	ab[jí]rtu	'open (plural)'
k[wé]rpos	'bodies'	k[wí]rpu	'body'

(40) Pasiego metaphony does not affect /í/, /ú/ and /á/

Unmetaphonized	*Gloss*	*Neutral*	*Gloss*
luz m[í]yos	'mine (plural)'	il m[í]yu	'mine (singular)'
bj[ú]da	'widow'	bj[ú]du	'widower'
br[á]θos	'arms'	br[á]θu	'arm'

7.5.1.1 Dyck's Generalization

Dyck (1995), modifying an earlier observation by Penny (1970), formulates the generalization in (41) about Romance metaphony (raising) triggered by desinential vowels.

(41) Generalization about metaphony triggers (Dyck 1995)
Desinential high vowels can trigger metaphony only if they contrast with a mid vowel in the same place.

Note that (41), henceforth Dyck's Generalization, refers to contrasts *only among the desinential vowels*. In every Romance dialect high vowels contrast with mid vowels in stressed syllables; but dialects have different inventories of desinential vowels, ranging from three to five. Because the phonetics of these vowels can vary, I will henceforth represent them schematically as /I ~ E, A, U ~ O/ for three-vowel desinential inventories, /I ~ E, A, O, U/ for four-vowel desinential inventories with a contrast between a high and mid back/labial vowel, and so on.

Dyck's Generalization makes correct predictions about which dialects may exhibit metaphony, and what the possible desinential triggers of metaphony may be in these dialects. First, we expect no raising in dialects with only three desinential vowels /I ~ E, A, U ~ O/, because there is no contrast between /I/ and /E/ or /U/ and /O/. For example, no raising is reported in Leonese dialects, where desinences are phonetically [i, a, u] or [e, a, u], depending on the dialect.

In dialects with four desinential vowels /I ~ E, A, O, U/, the prediction is that raising can be triggered by /U/, not by /I/; in dialects with four desinential vowels /I, E, A, O ~ U/, we predict that raising can be triggered by /I/, not by /U/. Examples of the former type are Central Asturian, North Central Asturian and Montañes dialects of Santander, where /u/ contrasts with /o/, but there is only a marginal, archaic contrast between /i/ and /e/. As expected, raising is triggered by [u], not by [i].

In dialects with five desinential vowels, Dyck's Generalization predicts that both /I/ and /U/ can trigger raising. No Spanish dialects are of this type, but there are Italian dialects, such as Servigliano, that have five desinential vowels and raising triggered by both [i] and [u].

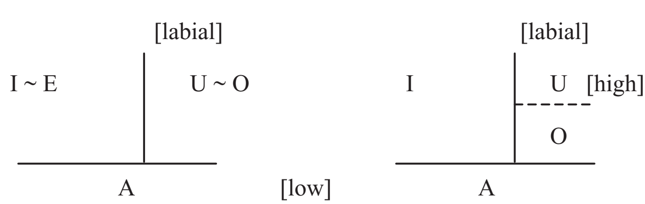
7.5.1.2 Accounting for Dyck's Generalization

In order to account for Dyck's Generalization we must make several assumptions. First, for purposes of evaluating contrasts, vowel inventories must be divided into stem inventories and desinential inventories. Contrasts in each inventory are assessed separately.

Second, we must assume that features in these dialects are ordered [low] > [labial] > [high]. If the first feature is [low], then [high] is not needed in the three-vowel system shown in (42a). The non-low vowels have no contrastive [high] feature to trigger raising, even if they are pronounced as [i] or [u]. In the inventory in (42b), the feature [high] is needed to distinguish between /U/ and /O/, but its scope is limited to the [labial] vowels. Therefore, only /U/ can trigger metaphony, not /I/. These results are strikingly in line with the Contrastivist Hypothesis: only vowels possessing a contrastive feature [high] can trigger metaphony.

(42) Contrastive features in desinential vowels

a. Three desinential vowels b. Four desinential vowels



7.5.1.3 Phonetics of desinential vowels

Dyck's Generalization by itself does not necessarily demonstrate the correctness of the Contrastivist Hypothesis if the data can be explained in other ways. One possible alternative comes to mind immediately. We might reason that in dialects where a high desinential vowel is in contrast with a mid vowel, the high vowel must be phonetically higher than in dialects where only one vowel covers the high and mid space. It is possible, therefore, that metaphony is triggered by contrastively high desinential vowels not because the phonology is concerned with contrastive specifications, but simply because such vowels are the only ones that are phonetically high enough to trigger metaphony. On this account, Dyck's Generalization is explained by phonetics.

Fortunately, the phonetics of Iberian dialects have been richly documented, and it is simply not the case that noncontrastive high desinential vowels are always lower than contrastively high vowels. In a survey of the phonetic descriptions of these dialects, Dyck (1995) shows that while it is true, as a general tendency, that noncontrastive vowels exhibit greater variability than contrastive vowels, there is also a great deal of variation from dialect to dialect. In the Leonese area, for example, there is no phonological contrast between desinential /i/ and /e/. Nominal desinences vary between [i] and [e], depending on dialect and also context (Dyck 1995: 68–71). If the phonetic hypothesis were correct, we would expect metaphony to depend on the phonetic height of the desinential vowel: in dialects where it is always pronounced [i] in certain lexical items we would expect metaphony in those words; where it varies between [i] and [e], we would expect to find variable metaphony. This is not what is found, however: no synchronic metaphony is reported for the Leonese area.

The conclusion is that a purely phonetic account does not suffice here. Whether a vowel is phonetically high at a given time does not predict the possibility of metaphony. Dyck's Generalization remains true only at the level of contrastive features. Thus, the Contrastivist Hypothesis, together with the proposed contrastive hierarchy, remains the best explanation for this generalization.

7.5.2 Metaphony in Campidanian Sardinian

Frigeni (2003) argues for a domain split in the assignment of contrastive features to vowels in Campidanian Sardinian along the lines of Dyck's analysis. She goes one step further, arguing that the contrastive hierarchy for desinential vowels differs from the one that applies to stem vowels, not just in the number of contrastive features, but also in their ordering.

The surface inventories of stem and desinential vowels are shown in (43).

(43) Surface vowels in Campidanian Sardinian

a. Stem vowels

i						u
	e				o	
		ɛ		ɔ		
			a			

b. Desinential vowels

i		u
	a	

Frigeni argues that at the lexical level there are two fewer stem vowels than in (43a), and two more desinential vowels than in (43b) (cf. Bolognesi 1998: 20–1). She argues that the stem vowels [e] and [o] are derived from /ɛ/ and /ɔ/, respectively, by metaphony, as illustrated in (44).

(44) Campidanian Sardinian metaphony

a. No metaphony before /-a/		b. Metaphony before /-i, -u/	
pɔ́tːa	'door F. SG.'	pótːu	'harbour M. SG.'
(a)rːɔ́za	'rose F. SG.'	drómi	'sleep INF.'
sɛ́rːa	'hill F. SG.'	ʒéʁu	'sky M. SG.'
fɛ́ʃta	'party F. SG.'	tésːi	'weave INF.'

Since the surface stem vowels *e* and *o* are derivable by metaphony from the corresponding lax vowels, there is no evidence that they are underlying.

In the preceding section we saw that Spanish and Italian dialects with three desinential vowels do not have metaphony. The reason, according to Dyck's analysis, is that metaphony is caused by a contrastive feature [high] which is only present if there is a contrast between a high and mid vowel at the same place. The Sardinian system in (43b) appears to contradict this generalization, but the contradiction is only apparent. Frigeni (2003) shows that the surface desinential vowels *i* and *u* each represent the merger of two vowels, one which causes metaphony, and one which does not. She argues that the vowels which do not cause metaphony derive from underlying /e/ and /o/, which merge with /i/ and /u/, respectively. Therefore, the underlying stem and desinential vowel systems are as in (45). Capital letters indicate that we are not committing to any particular features beyond a minimal contrastive set.

(45) Underlying vowels in Campidanian Sardinian (Frigeni 2003)

a. Stem vowels			b. Desinential vowels		
I		U	I		U
E		O	E		O
	A			A	

Though the two inventories now look identical, Frigeni argues that their contrastive specifications differ. If we continue to assume that the metaphony trigger is a contrastive feature [high], as in the cases in the previous section, we would expect the mid vowels to raise to high vowels *i* and *u*. Second, metaphony here introduces a new contrast between mid vowels, and it is not clear how spreading [high] to a mid vowel unspecified for any other height feature could lead to this result.

To account for the change undergone by the stem vowels, Frigeni (2003) proposes that the spreading feature in Campidanian Sardinian metaphony is [ATR], not [high]. Hence, [ATR] is the contrastive feature distinguishing desinential /I, U/ from /E, O/. But she demonstrates also that the stem vowels could not be characterized in the same way, or else again we should expect metaphonized vowels to surface as *i* and *u*. She proposes, therefore, that the two domains have different contrastive features, as shown in (46). The stem vowels correspond to a contrastive hierarchy [low] > [labial], [high], and desinential vowels have the contrastive hierarchy [low] > [labial], [ATR].

(46) Underlying contrasts in Campidanian Sardinian (Frigeni 2003)

a. Stem vowels

I			U	[high]
	E	O		
	A			[low]

b. Desinential vowels

I			U	[ATR]
	E	O		
	A			[low]

On this account, metaphony results from the spreading of contrastive [ATR] from a desinential vowel onto a stem vowel. The feature [low] is incompatible with [ATR] in this language, hence /a/ is unaffected by metaphony; the high vowels surface as [ATR] in any case. On the mid vowels, however, the spreading of [ATR] from a desinential vowel has a noticeable effect: /ɛ/ becomes [e] and /ɔ/ becomes [o].

7.5.3 Summary

The generalizations that govern Romance metaphony support the Contrastivist Hypothesis in striking fashion. As predicted, only contrastive features can trigger metaphony. These dialects also exemplify situations where contrast is not evaluated over the entire inventory, but over two different domains; in this case, the desinential vowels are evaluated apart from all the other vowels. These examples also provide evidence that the phonetic realization of a desinential vowel does not predict whether it can be a metaphony trigger. Further, the level at which contrast is required is not at the surface phonetics, but in the underlying lexical representations.

7.6 Consonant co-occurrence restrictions

Mackenzie (2005, 2009) argues that the best analysis of many consonant harmony systems requires specifying certain features as contrastive in terms of a feature hierarchy. For then a simple generalization emerges: consonant harmony applies to segments contrastively specified for the harmonic feature.

In Bumo Izon (an Ijoid language of Nigeria), labial and alveolar implosive and plosive stops may not co-occur in a morpheme (Efere 2001). Thus, implosive /ɓ, ɗ/ may not co-occur with plosive /b, d/, though the plosives may freely occur with each other, as may the implosives (47).

(47) Bumo Izon labial and alveolar plosives and implosives (Efere 2001)

	Plosives		Implosives	
Labials	búbú	'rub (powder in face)'	ɓʊ́ɓaɪ	'yesterday'
Alveolars			ɗɔ́ɗɔ́	'cold'
Mixed	bídɛ́	'cloth'	ɗáɓá	'swamp'

The velar plosive /g/ and the labiovelar implosive /ɠɓ/, however, may freely occur with members of both the plosive and implosive series, as shown in (48).

(48) Bumo Izon velar plosive and labiovelar implosive (Efere 2001)

	Velar plosive /g/		*Labiovelar implosive /ɠɓ/*	
With same	igódó	'padlock'	ɠɓáɓu	'crack (of a stick breaking)'
With different	ɗúgɔ́	'to pursue'	ɠɓódaɠɓóda	'rain (hard)'
	ɓúgí	'to wring (hand)'		

Why are /g/ and /ɠɓ/ exempt from harmony? Consider the inventory of oral stops in this language, shown in (49).

(49) Bumo Izon oral stops (Mackenzie 2005: 174, based on Efere 2001)

		labial	alveolar	velar	labiovelar
plosive	voiceless	p	t	k	kp
	voiced	b	d	g	
implosive		ɓ	ɗ		ɠɓ

Intuitively, the labial and alveolar voiced plosive stops each have an implosive 'partner', whereas the velar and labiovelar voiced stops have no counterparts. Building on Hansson's (2001) observation that contrast seems to play an important role in accounting for these facts, Mackenzie (2005, 2009) presents an analysis in terms of the contrastive hierarchy.

Assuming that the relevant laryngeal feature is [glottalic], Mackenzie (2005) proposes that the contrastive hierarchy for Bumo Izon is: place features > [voiced] > [glottalic]. That is, the consonants are first distinguished by place, in terms of the place categories shown in (49). Within each place, they are then distinguished by [voiced].[27] Now [glottalic] is needed only to distinguish the labials and alveolars. The contrastive features assigned to the voiced stops are

27 Note that this hierarchy differs from what is suggested by the table in (49), where [voiced] appears to take narrower scope than [glottalic]. Such an ordering would assign contrastive [glottalic] features to the labiovelar stops, wrongly predicting that they participate in harmony.

shown in (50). The phonemes that participate in implosive harmony are exactly the ones that are contrastively specified for the harmonizing feature, [glottalic].

(50) Bumo Izon voiced stops: contrastive features (Mackenzie 2005)

	b	ɓ	d	ɗ	g	ɠɓ
place	lab	lab	alv	alv	vel	labvel
voiced	+	+	+	+	+	+
glottalic	−	+	−	+		

7.7 Loanword adaptation as evidence of phonological organization

Loan phonology has been viewed as a source of evidence bearing on the nature of the grammar of the borrowing language at least since Jakobson (1962b [1931]), and, in the framework of generative phonology, at least since Hyman (1970a, b, 1973). Many basic issues in loan phonology remain controversial. Some have argued that the input to loan phonology is analysed in terms of the phonemes of the lending language (Paradis 1988; LaCharité and Paradis 1997; Jacobs and Gussenhoven 2000; Paradis and Prunet 2000). Paradis and her collaborators, for example, assume that speakers responsible for borrowing are bilingual, and familiar with the grammar of the lending language. Others have assumed that the input to loan phonology is phonetic (Silverman 1992; Yip 1993; Kenstowicz 2003; Steriade 2009). It appears that both views may be correct depending on the social context in which borrowing takes place (Kiparsky 1973: section 3.2; Heffernan 2005, 2007).

There is much evidence that the sound structure of one's native language affects one's perception of foreign sounds (Hancin-Bhatt 1994; Best 1995; Flege 1995; Dupoux *et al.* 1999; Brown 2000). In recent years, some studies have argued that noncontrastive features also play an important role in the adaptation of sounds (Brannen 2002; Pater 2003; see Kang 2007 for discussion of this and related issues). However, as with other aspects of phonology, much discussion of contrastive and redundant features in loanword phonology suffers from the lack of a satisfactory account of which features actually are contrastive.

Turning to our main theme, I will focus on the relationship between loan phonology and the contrastive hierarchy of a language. In chapter 6, section 6, we considered abstractly how the contrastive hierarchy might be brought to bear on loan phonology. In section 7.7.1 I review some early proposals by Jakobson, and section 7.7.2 draws on a study by Herd (2005) that argues for the importance of the contrastive hierarchy in accounting for Polynesian loan phonology.

7.7.1 Loanword adaptation and the contrastive hierarchy

We have seen that Jakobson (1962b [1931]) and Jakobson and Lotz (1949) appealed to how native speakers adapt foreign sounds as evidence that the native language uses a particular set of contrastive features. Jakobson (1962b [1931]) cites the alleged relative ease with which native Slovak and Russian speakers adapt front rounded vowels from French or German, in comparison with the greater difficulty Czech speakers have with such sounds, as evidence that the backness and rounding features are dissociated in the former two languages but not in Czech. Presumably, the independence of these features in Slovak and Russian phonology facilitates their combination in novel ways.

Jakobson and Lotz (1949) argue that the difference between velar and palatal place is irrelevant in French; in their analysis, the palatals /ʃ, ʒ/ and velars /k, ɡ/ both have the place feature [+saturation]. In support of this proposal they cite the frequent adaptation of the English velar nasal /ŋ/ as the palatal /ɲ/ in French. On this view, foreign sounds are filtered through the contrastive features of the native language. Though Jakobson and Lotz do not elaborate on how this might work, we can adapt the 'decision tree' model proposed by Jakobson, Fant and Halle (1952) for identifying phonemes in one's native language as one way of instantiating this idea. Thus, a French speaker hearing or attempting to produce English [ŋ] could proceed down through the French contrastive feature hierarchy, chapter 3's (28) in the proposal of Jakobson and Lotz, making a series of binary decisions: going top down, [ŋ] is [−vocalic], [+nasality] and [+saturation]. At this point there are no further contrastive features to be assigned, and the English sound [ŋ] is identified with the French phoneme /ɲ/.[28]

7.7.2 Loanword adaptation in Polynesian languages (Herd 2005)

Herd (2005) studies patterns of adaptation of English words into a number of Polynesian languages. These languages have impoverished consonantal inventories, so many substitutions can be observed. Herd argues that the adaptation patterns provide evidence for the influence of the contrastive hierarchies of the

28 In this case, we could question the strength of the argument by asking what other French phoneme English [ŋ] could be identified with. In the analysis of Martinet (1964), /ɲ/ is assigned a [palatal] place, and the velars /k, g/ are [dorso-velar]. The combination of [dorso-velar, +nasal] does not exist in native Standard French, and Martinet's analysis does not suggest what strategy would be employed by a French speaker in realizing this sound. If the place feature were considered to be contrastively more important, we might expect it to override the nasal feature, incorrectly yielding /g/ as the substitute for [ŋ]. But it is not clear Martinet's analysis would be committed to this approach.

borrowing languages *à la* Jakobson and Lotz. He argues further that phonetic similarity is not sufficient to account for these patterns. Of the many cases he discusses, I will briefly review the adaptation of coronal fricatives into two Eastern Polynesian languages, Hawaiian and New Zealand Māori.

7.7.2.1 Hawaiian

Hawaiian has a famously small consonantal inventory (51).

(51) Hawaiian Consonantal Inventory

p		k	ʔ
			h
m	n		
w	l		

All English coronal obstruents are borrowed into Hawaiian as /k/, including [s], [z] and [ʃ] (52). Note that these segments are not adapted as /h/, which is also a plausible candidate from a phonetic point of view.

(52) Hawaiian adaptation of English coronal fricatives (Herd 2005)

a. [s] → /k/	lettuce → /lekuke/	soap → /kope/
b. [z] → /k/	dozen → /kaakini/	
c. [ʃ] → /k/	brush → /palaki/	machine → /mikini/

7.7.2.2 NZ Māori

NZ Māori has both /k/ and /h/, as well as /t/, though it lacks a phonemic glottal stop (53). In this language, English [s], [z] and [ʃ] are borrowed as /h/, as shown in (54). This is surprising, given that /k/ is available, as in Hawaiian.

(53) NZ Māori Consonantal Inventory

p	t	k	
f			h
m	n	ŋ	
w	r		

(54) NZ Māori adaptation of English coronal fricatives (Herd 2005)

a. [s] → /h/	glass → /karaahe/	sardine → /haarini/
b. [z] → /h/	weasel → /wiihara/	rose → /roohi/
c. [ʃ] → /h/	brush → /paraihe/	sheep → /hipi/

If substitutions are made on the basis of similarity, these facts are hard to explain. As Herd (2005) points out, if coronal fricatives are more similar to /k/ than to /h/ in Hawaiian, why are they more similar to /h/ than to /k/ in NZ Māori? The relevant notion of similarity must be somehow influenced by the different inventories of these languages. Herd proposes that different contrastive specifications are operative in each language.

7.7.2.3 Contrastive specifications of Hawaiian and NZ Māori consonants

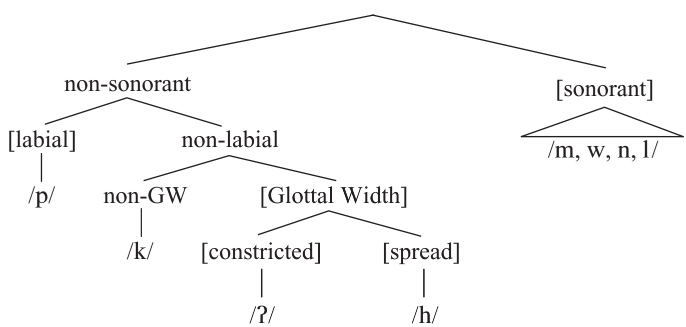
Herd (2005) proposes that the contrastive status of /h/ is different in the two languages. In Hawaiian, /h/ contrasts with /ʔ/. Following Avery and Idsardi (2001), the existence of this contrast activates a laryngeal dimension they call *Glottal Width*. Glottal Width has two values, [constricted] for /ʔ/, and [spread] for /h/.

Herd proposes the feature ordering for Hawaiian shown in (55) (only features relevant to the current discussion are mentioned).

(55) Contrastive hierarchy for Hawaiian (Herd 2005)
[sonorant] > [labial] > Glottal Width ([spread/constricted])

First, [sonorant] distinguishes /m, n, w, l/ from /p, k, ʔ, h/. Next, [labial] splits off /p, m, w/ from the rest. Then laryngeal Glottal Width applies to /ʔ, h/. The result is that /h/ is specified for [spread], /ʔ/ is specified [constricted] and /k/ is the default obstruent (56). Therefore, anything that is not sonorant or labial or laryngeal is adapted to /k/. In particular, [s, z, ʃ] → /k/.

(56) Hawaiian contrastive specifications (Herd 2005)

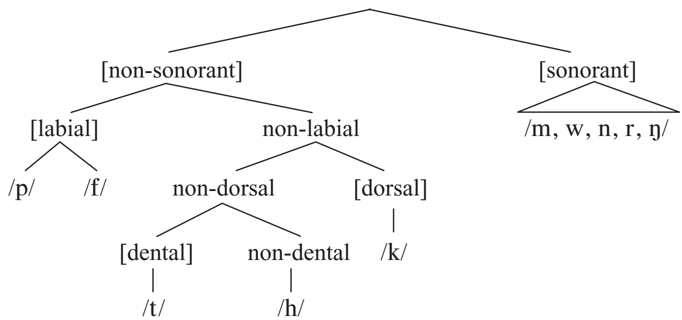


Unlike Hawaiian, NZ Māori has no /ʔ/, so there is no contrast within Glottal Width. Herd (2005) proposes that, lacking such a contrast, [spread] is not accessible as a contrastive feature. This, and the other differences in the inventories of the two languages, result in a different contrastive hierarchy for NZ Māori (57).

(57) Contrastive hierarchy for NZ Māori (Herd 2005)
[sonorant] > [labial] > [dorsal] > [dental]

As in Hawaiian, [sonorant] goes first, splitting off /m, n, ŋ, w, r/, and [labial] follows, applying to /p, f, m, w/. Unlike Hawaiian, [dorsal] is also required, to distinguish /k, ŋ/ from /t, n/. It remains to distinguish /t/ from /h/. Herd proposes to use the feature [dental] to characterize the contrastive property of /t/. This feature accounts for why the interdental fricatives [θ] and [ð] become /t/, not /h/. Thus, in NZ Māori /h/ plays the role of default obstruent, not /k/: /h/ is not sonorant, not labial, not dorsal, and not dental (58). Therefore, [s, z, ʃ] → /h/.

(58) NZ Māori contrastive specifications (Herd 2005)



The different contrastive roles played by /h/ in these languages suggests that they have different ‘pattern alignments’, in Sapir’s terms, despite their very similar phonetic realizations. The differing status of /h/, as well as the presence of /t/ in NZ Māori but not in Hawaiian, also account for the very different contrastive status of /k/ in each language: general default obstruent in Hawaiian, and dorsal obstruent in NZ Māori.

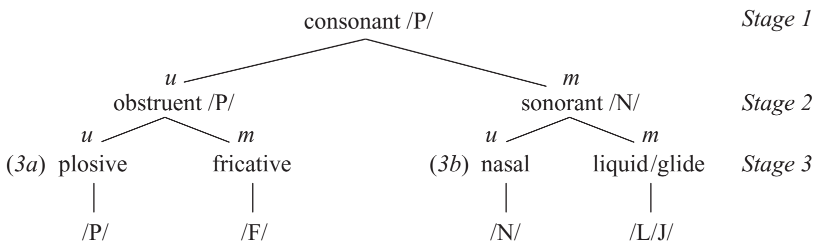
7.7.3 Summary

As mentioned, loan phonology is a diverse phenomenon, and it is unlikely that a single approach can account for all patterns of loanword adaptations. But it suffices for our purposes to show that there exists a class of cases in which loan phonology is sensitive to the contrastive structure of a language; in particular, to the contrastive feature hierarchy. The Polynesian examples discussed above provide a compelling case of this type.

7.8 The acquisition of distinctive features and contrasts

Following the pioneering work of Jakobson (1941) and Jakobson and Halle (1956) discussed in chapter 4, section 3, the notion of a contrastive hierarchy has been fruitfully applied in acquisition studies, where it is a natural way of describing developing phonological inventories (Pye, Ingram and List 1987; Ingram 1988, 1989; Levelt 1989; Dinnsen *et al.* 1990; Dinnsen 1992, 1996; see Dresher 1998a for a review). For example, Fikkert describes the development of segment types in onset position in Dutch as in (59).

(59) Development of Dutch onset consonants (Fikkert 1994)



In Stage 1 there are no contrasts. The value of the consonant defaults to the least marked onset, namely an obstruent plosive, designated here as /P/. The first contrast (Stage 2) is between obstruent and sonorant. The former remains the unmarked option (*u*); the sonorant defaults to nasal, /N/. At this point children differ. Some expand the obstruent branch first (Stage 3a), bringing in marked fricatives, /F/, in contrast with plosives. Others (Stage 3b) expand the sonorant branch, introducing marked sonorants, which may be either liquids, /L/, or glides, /J/. Continuing in this way we will eventually have a tree that gives all and only the contrasting features in the language.

While the contrastive hierarchy has been useful in depicting developing inventories as they appear in children's production, experiments on child and infant perception of phonetic contrasts have appeared to support a different view of phonological acquisition. Beginning with Eimas *et al.* (1971), it has been shown that infants can discriminate fine phonetic distinctions in speech sounds, including sounds that are not discriminated in the ambient language (Trehub 1976; Werker *et al.* 1981; Werker and Tees 1984). Thus, whereas adults have difficulty discriminating certain distinctions not used in their native language, infant perception appears to be 'universal'.[29] A series of studies showed that infants 'tune' their phonetic perceptions in accordance with the distribution of sounds in the language they are acquiring, thus eventually losing the ability to discriminate foreign sounds (Werker and Tees 1984; Kuhl *et al.* 1992).

29 There are also studies showing that certain phonetic contrasts are not as well-discriminated by infants as by adult native speakers (Aslin *et al.* 1981; Polka, Colantonio and Sundara 2001; see also Weiss and Maye 2008).

This tuning occurs in the first year, before the learners have acquired a lexicon. These results have led some to conclude that learners acquire the phonemes of their language before they can produce or understand words. For example, Pinker (1994: 264–5) describes the process as follows:

> By six months, [babies] are beginning to lump together the distinct sounds that their language collapses into a single *phoneme*, while continuing to discriminate equivalently distinct ones that their language keeps separate. By ten months . . . they do not distinguish Czech or Inslekampx *phonemes* unless they are Czech or Inslekampx babies. Babies make this transition before they produce or understand words . . . They must be sorting the sounds directly, somehow tuning their speech analysis module to deliver the *phonemes* used in their language. The module can then serve as the front end of the system that learns words and grammar [emphasis added].

On the face of it, it is hard to see how infants can acquire phonemes without knowing if two utterances are the 'same' or 'different' (Bloomfield 1933). It has been argued that learners are particularly attentive to the distribution of sounds, and can draw certain conclusions about whether a cluster of sounds are to be assigned to one category or to more than one, even in the absence of vocabulary or meaning (Maye 2000; Maye, Werker and Gerken 2002; Weiss and Maye 2008). However, distribution can only take one so far. In fact, there is no evidence that infants have acquired *phonemes* by the age of one. The source for Pinker's claims is the following passage by Kuhl *et al.*:

> Infants demonstrate a capacity to learn simply by being exposed to language during the first half year of life, before the time that they have uttered meaningful words. By 6 months of age, linguistic experience has resulted in language-specific *phonetic prototypes* that assist infants in organizing speech sounds into *categories*. They are in place when infants begin to acquire word meanings toward the end of the first year. *Phonetic prototypes* would thus appear to be fundamental perceptual-cognitive building blocks rather than by-products of language acquisition [emphasis added]. (1992: 608)

What Kuhl *et al.* call 'phonetic prototypes' are not equivalent to phonemes; they are *phones*, phonetic variants of phonemes. Infants become sensitive to the phonetic range and distribution of the sounds of their language, so they can tell, for example, that the pronunciation of a Swedish [i] inserted into an English utterance is somehow anomalous. But this is not the same as learning which phones cluster together to form phonemes.

Nevertheless, the fact that infants are able to make fine phonetic discriminations has sometimes been taken as evidence that children's initial phonological representations are accurate and essentially adult-like (cf. Hale and Reiss 1998).

If that is correct, then it must be the case that the appearance that learners are gradually acquiring phonological contrasts is not a reflection of their linguistic competence, but only of production. This theory is bolstered by anecdotes that children are aware of phonemic contrasts that they are unable to produce themselves; a famous example is Neil Smith's son Amahl protesting when his father said *sip* instead of *ship*, even though Amahl himself pronounced both as *sip* (N. V. Smith 1973).

If learners' phonological representations were adult-like from the beginning, we would no longer have evidence that the system of contrasts is learned gradually, nor would we have evidence for a contrastive hierarchy in acquisition. In fact, we would not even have evidence that contrast is important in acquisition, beyond the distribution of surface allophones. However, there is evidence that we cannot draw these conclusions from the above studies.

In sharp contrast to the excellent performance of young children on phonetic discrimination tasks is their inability to utilize fine phonetic differences in word recognition tasks (Stager and Werker 1997; Werker *et al.* 2002; Pater, Stager and Werker 2004). For example, the 14-month-old children studied by Stager and Werker could not distinguish minimally different nonce words such as *bin* and *din* in a word recognition task (when the 'words' were associated with objects), though they could distinguish them in a pure discrimination task. It follows that purely phonetic perception does not translate immediately into phonological representation. The results are consistent with the view that phonological representations do not contain all the details available to phonetic perception (Werker *et al.* 2002; Pater, Stager and Werker 2004; Pater 2004). Fikkert and Levelt (2008) argue that phonological representations are underspecified to begin with, in support of the 'constructionist' or 'emergentist' view of acquisition inspired by production studies. Fikkert (2007) proposes that there is evidence from perception that supports the constructionist interpretation of the production studies.

Putting everything together, we have a picture of a learner going in two directions simultaneously. At the phonetic perceptual level, child learners begin by attending to many potential sources of contrasts, and are more able than adults to discriminate sounds not used in the ambient language (Eimas *et al.* 1971; Werker *et al.* 1981). Acquisition of the native language requires that they 'tune' their perceptual system to the contrasts used in their language, while learning to disregard contrasts that are not used (Werker and Tees 1984; Kuhl *et al.* 1992). Meanwhile, phonological representations are impoverished to begin with (Fikkert 2007). Infants' rich perception of phonetic contrast does not translate into a system of phonological representations (Stager and

Werker 1997). Phonological representations are built into systems of increasing complexity (Rice and Avery 1995), based on the input from phonetic perception together with evidence from the grammar, which itself becomes more complex and removed from the initial percepts (Dresher 1996, 1999).

An important part of phonological learning is the acquisition of the contrastive feature hierarchy. The evidence presented throughout this book suggests that this hierarchy cannot possibly be present from the beginning, because it depends not just on accurate phonetic perception, but on an understanding of various subtle aspects of phonological patterning.

7.9 Refining the Contrastivist Hypothesis

Throughout this study I have been assuming that the Contrastivist Hypothesis is as stated in (42) in section 3.7: 'The phonological component of a language L operates only on those features which are necessary to distinguish the phonemes of L from one another' (D. C. Hall 2007: 20). This formulation captures the intuitive idea that the 'phonemic content' or 'pattern alignment' of a phoneme is made up of its contrastive feature specifications. I have cited examples that support the Contrastivist Hypothesis to the extent that they suggest that only contrastive features, as identified by the SDA operating on a contrastive hierarchy, are active in the phonology. The existence of such cases does not, however, exclude the possibility that there are other cases in which demonstrably non-contrastive features are also active in the phonology. Such cases would show that the Contrastivist Hypothesis as stated above is too strong.

In this section I will consider the empirical adequacy of the Contrastivist Hypothesis in this sense. I will conclude that the Contrastivist Hypothesis in its strongest form cannot be maintained. This does not mean that the hypothesis must be abandoned; rather, it can be refined so as to allow for an important class of apparent counterexamples while maintaining the essential spirit of the hypothesis.

7.9.1 Is the Contrastivist Hypothesis too weak?

Before considering if the Contrastivist Hypothesis is too strong, I would like to briefly consider whether it might also be too weak. Saying that a theory is too weak means that it is not sufficiently constrained, and thus is hard or even impossible to falsify.

I addressed this question at the end of chapter 3, where I showed that the Contrastivist Hypothesis is easily falsifiable. It is enough to find examples of more features being phonologically active than are permitted to be contrastive. As the review in the next chapter will show, many theorists assume that such cases are common, to the extent that they allow noncontrastive features to freely figure in the phonology. I conclude, then, that the Contrastivist Hypothesis is not too weak, but is stronger in this regard than many competing approaches.

7.9.2 Is the Contrastivist Hypothesis too strong?

The real question to be addressed is whether the Contrastivist Hypothesis is too strong, and if so, in what ways? When considering counterexamples, it is important to distinguish between apparent counterexamples and real ones.

It is easy enough to find apparent counterexamples to the Contrastivist Hypothesis that turn out not to be real counterexamples to the theory presented here. One reason for this is that many studies make different assumptions about which features are contrastive. Many phonologists continue to arrive at contrastive specifications by something like pairwise comparison, as will become clear in chapter 8. This approach, we have seen, takes logical redundancy as its basic criterion for deciding if a feature specification is contrastive or not. Such theories typically designate too few feature specifications as being contrastive. That is, there will typically be features that the SDA in a certain ordering designates as contrastive that pairwise comparison designates as redundant. If such a feature is active, then we have an apparent violation of the Contrastivist Hypothesis, but one which dissolves when we recognize the feature in question to be contrastive.

For example, the famous unpaired fricatives and affricates of Russian are active in voicing assimilation (see chapter 4, section 6, and section 8.7 below). If pairwise comparison is the procedure for assigning contrastive features, the [−voiced] specifications of these phonemes will be designated as redundant, with the result that a redundant feature triggers assimilation, in violation of the Contrastivist Hypothesis. Adopting the contrastive hierarchy as the procedure for assigning contrasts results in a different conclusion: the features in question are contrastive, and there is no violation of the Contrastivist Hypothesis.

To take another example, Yowlumne labial harmony is triggered by both /o/ and /u/. A number of theories to be discussed in chapter 8 find that the [labial] (or [+round]) specification on /u/ is redundant; this could be because [back] must take precedence over [round], or because /u/ is not sufficiently 'crowded', or because [labial] is not the sole feature that distinguishes /u/ from any other phoneme, and so on. In any such theory the crucial harmonic feature is noncontrastive. But in the approach taken here, [labial] is ordered high enough in the feature hierarchy to be contrastive on both /u/ and /o/, and there is no violation of the Contrastivist Hypothesis.[30]

It is only to be expected that there will be many such cases, given the widespread use of pairwise comparison in determinations of which features are contrastive. Therefore, many apparent violations of the Contrastivist Hypothesis can be resolved with a different, arguably more empirically adequate, assignment of contrastive feature specifications.

Another source of uncertainty concerns the dividing point between the phonological component, in which only contrastive features are computed, and the postlexical or phonetic components where this limitation does not obtain. If a post-phonological rule that refers to redundant features is incorrectly assigned to the phonological component, we will create an apparent counterexample to the Contrastivist Hypothesis that will disappear once the rule is reassigned to its correct component.

That said, not all counterexamples to the Contrastivist Hypothesis can be resolved in these ways, and we are left with real counterexamples that have to be accounted for. D. C. Hall (2007) presents one class of cases of this kind, and proposes in response a slight modification of the Contrastivist Hypothesis.

7.9.3 'Prophylactic' features (D. C. Hall 2007)

Yowlumne Yokuts provides a real counterexample to the Contrastivist Hypothesis. We have seen that the underlying vowel system of Yowlumne is specified by two contrastive features, repeated here in more detail as (60).

(60) Yowlumne underlying vowels

[–round]	[+round]	
i iː	u uː	[+high]
a aː	o oː	[–high]

In lexical (underlying) forms, Yowlumne has a symmetrical vowel system where each short vowel has a long counterpart. Underlying long high vowels, however, are not pronounced as such, but are lowered. The vowel /uː/ lowers to /oː/, as expected, but lowered /iː/ comes out as [eː], not as [aː]. Inspection of (60) reveals that the allophone [eː] cannot be accommodated with the two features [high] and [round]. A third feature is required to keep [eː] distinct from

30 Yowlumne does pose a real problem, however, which is discussed in the next section.

[aː], and whatever this feature is, it cannot be contrastive. Since the lowering rule feeds further phonological rules, such as shortening (see Kenstowicz and Kisseberth 1977 for a detailed exposition), it is unlikely that it is simply a late phonetic rule. This, then, is a case where a noncontrastive feature is needed in the phonology.

D. C. Hall (2007, 2008) proposes that the noncontrastive feature [+low] must be attributed to /aː/.[31] When /iː/ lowers, it loses its specification [+high], but does not take on [+low]. Thus, it remains distinct from /aː/. Hall observes that the function of the redundant feature [low] is purely passive: it serves only to distinguish segments that would otherwise be neutralized. He calls such features *prophylactic*, defined as in (61).

(61) Prophylactic features (D. C. Hall 2007: 87–8)
A prophylactic feature is a redundant feature that is crucially present in the representation of a segment before the phonological computation begins, but which is invisible to all phonological rules.

Hall discusses several such examples. To cite one more, Czech /t/ and /ř/ (IPA r̝) are distinguished only by the feature [laryngeal], in the analysis of Hall (2007). Devoicing, on this account, is effected by the addition of [laryngeal]. However, when /ř/ devoices it does not merge with /t/, but appears as voiceless [ř̥], an allophone that does not exist as a distinct phoneme in Czech. To prevent a merger with /t/, Hall proposes that /ř/ bears the prophylactic feature [vibrant]. Like Yowlumne [low], this feature does not figure in the phonological computation: it does not trigger rules, and it is not referred to by rules. However, its presence prevents the merger of two phonemes.

Positing prophylactic features represents a minimal retreat from the Contrastivist Hypothesis. It remains to be specified under what conditions such features typically arise, and whether other types of counterexamples must be recognized. In the meantime, the range of cases where the Contrastivist Hypothesis is upheld and contributes to illuminating analyses suggests that it is well worth maintaining and refining as a basic principle of phonological patterning.

31 Hall (2007, 2008) adopts privative features [high], [low] and [peripheral] (for [round]). For ease of exposition I will continue to use binary features. Hall observes that the presence of [low] on /a/ predicts that it should remain a low vowel when it is rounded to /o/. He argues that this is indeed the case, and that the vowel /o/ is transcribed by Newman (1944: 19) as ɔ, ‘as in German *voll* and English *law*’.

7.10 Summary

In this chapter I have presented a series of cases that support the Contrastivist Hypothesis as a theory of phonology, and the contrastive hierarchy as a theory of phonological contrast. I have proposed that MCS incorporates the leading ideas of chapter 3 in a contemporary context. I also showed that feature ordering is an inescapable part of phonology, and that much phonological theory and practice incorporates, often tacitly, various aspects of the theory defended here. In some sense, then, this theory makes explicit what has been implicit in phonological thinking for a long time.

The cases discussed are drawn from different domains of phonology: vowel and consonant harmony, loanword adaptation, and acquisition. I have argued that the Contrastivist Hypothesis and the contrastive hierarchy are robustly supported in all these domains, and contribute to a comprehensive and unified account of phonological theory and development.

Of course, many contemporary approaches to phonology do not share the principles defended here, to a lesser or greater extent, and it is to these that I turn in the next chapter.

8 *Other approaches to contrast in phonology*

8.1 Introduction

In this chapter I will consider some other approaches to phonological contrast that have been advanced in the recent phonological literature. I will start with theories that are conceptually quite different from the approach I have been taking and then consider those that have more in common with it.

In section 8.2 I consider an alternative explanation of the typology of labial harmony triggers reviewed in section 7.4.3. There, I argued that the observed relation between harmony triggers and inventories supports the Contrastivist Hypothesis. Kaun (1995) advances what appears to be a very different explanation, grounded in perceptual-functional phonetic constraints. I will argue that her account is not, in fact, a real alternative to the explanation I proposed in chapter 7.

Dispersion Theory is concerned with contrast at a perceptual phonetic level, and in section 8.3 I review a dispersion-theoretic account of a phonological change in the history of Russian, proposed by Padgett (2003a). I will argue that this version of Dispersion Theory is enmeshed in descriptive and explanatory complications that may be difficult to sort out. I will present an MCS-style alternative solution that is conceptually much simpler.

Section 8.4 looks at Structured Specification, or natural classes, theory (Broe 1993; Frisch 1996; Frisch, Pierrehumbert and Broe 2004). Rather than designate features as being ‘contrastive’ or ‘redundant’, as the SDA does, this theory is indirectly sensitive to the contrasts in an inventory. It computes the similarity of pairs of phonemes as a function of how many natural classes they share and do not share. It has been proposed that this metric can provide a theory of similarity that plays a role in phonological processes. I will show that this similarity metric simply does not make the right predictions with respect to the phonological processes discussed in the previous chapter.

The theory of Clements (2001, 2009) is the subject of section 8.5. Clements proposes a feature hierarchy that governs segmental specification, as well as a theory of markedness, as in MCS. His theory differs from the one advocated here in assuming that these are universally fixed and independent of phonological activity. The theory of Calabrese (1994, 1995, 2005), discussed in section 8.6, also incorporates contrasts and hierarchy, and pays attention to feature activity, but with a different set of assumptions about how contrasts are achieved and what markedness is. I will argue that phonological activity is an essential source of evidence bearing on the contrastive feature hierarchy, and that the evidence from activity shows that the hierarchy is not universal.

Finally, section 8.7 considers the notion of 'minimal contrast' which plays a prominent role in the work of a number of contemporary phonologists whose theoretical assumptions are otherwise fairly diverse. I will show that minimal contrast arises from the concept of pairwise comparison, which was shown to have fatal flaws in chapter 2.

8.2 A perceptual-functional account of labial harmony triggers?

In section 7.4.3, I showed that typological surveys of labial (or rounding) harmony in Manchu-Tungus, Mongolian and Turkic languages support the hypothesis that only contrastive features trigger harmony. In particular, I showed that labial harmony triggers could be shown to be contrastive, given a contrastive hierarchy for vowels that orders [coronal] ahead of [labial]. This ordering appears to be quite stable in these language families, though it is not universal. Moreover, the differences between typical labial harmony triggers in Manchu-Tungus and Mongolian on one hand, and in Turkic languages on the other, could be related to their different inventories, and different contrastive representations.

Kaun (1995) proposes what appears to be an alternative to a contrastive account of labial harmony systems, one that is grounded in a perceptual-functional analysis. Closer inspection reveals, however, that her account presupposes a contrastive analysis such as the one presented here.

According to Kaun (1995), labial harmony is governed by a number of constraints. The main ones relevant to the present discussion are given in (1).

(1) Constraints responsible for labial harmony (Kaun 1995)

a. EXTEND[RD]: 'The autosegment [+round] must be associated with all available vocalic positions within a word.'

b. EXTEND[RD]IF[−HI]: 'The autosegment [+round] must be associated with all available vocalic positions within a word when simultaneously associated with [−high].'

Constraint (1a) provokes labial harmony triggered by both high and low round vowels, as occurs (potentially) in Turkic languages; (1b) is meant to account for the Manchu-Tungus–Mongolian type of labial harmony, which is triggered only by low vowels.

If grammars are permitted to freely rank these constraints, we would have no explanation for the correlation between inventories and type of labial harmony. It remains to be explained why Manchu-Tungus–Mongolian languages typically use Extend[Rd]If[−Hi], whereas Turkic languages use Extend[Rd]. To account for this correlation, Kaun proposes that Extend[Rd]If[−Hi] is dominant only if there is greater *perceptual crowding* in the non-high vowels than in the high vowels. To implement this notion in her formal theory, she adopts the convention that Extend constraints may operate only on *contrastive* feature values.

Kaun (1995) proposes no theory for identifying which values are contrastive. However, since her constraints do not refer directly to perceptual crowding but only to contrast, it follows that in her theory a segment has a contrastive feature if and only if it is crowded.

Deriving contrast from crowding is problematic when we consider languages like Yowlumne Yokuts. As we saw in section 5.3, Yowlumne Yokuts has height-bounded labial harmony triggered by both high and non-high vowels. The Yowlumne high vowel space is not crowded, however; on the contrary, it has optimal separation (2).

(2) Yowlumne Yokuts underlying vowel inventory

iː i		uː u
	aː a	oː o

Kaun (1995: 159) cannot explain why both /u(ː)/ and /o(ː)/ trigger labial harmony, since, as she assumes, [labial] is not contrastive in the high vowels. But there is no basis for this assumption. For reasons discussed in section 5.2, the Yowlumne feature hierarchy has [labial] ordered above [coronal]. Yowlumne /i/ is quite different from the /i/ in Manchu-Tungus and many Mongolian languages. Thus, /i/ is the epenthetic vowel in Yowlumne, and does not appear to cause palatalization or other modifications in neighbouring segments. For these reasons, Archangeli (1984) proposes that /i/ is the unspecified vowel in Yowlumne. Only two features can be contrastive in this inventory, and they are [labial] (or [round]) and [high], as shown in chapter 5, (27). Since [labial] is a contrastive feature on both /u(ː)/ and /o(ː)/, it is a potential harmony trigger; crowding is not required.

It follows, then, that if we interpret 'perceptual crowding' literally, the hypothesis that crowding drives contrast which in turn drives harmony is false. The hypothesis can be saved if we derive contrastive values from the contrastive hierarchy and interpret crowding in terms of contrast, so that a feature is crowded if and only if it is contrastive; but then crowding plays no role in the explanation. I conclude that Kaun (1995) does not present a genuine alternative to a Contrastivist account in terms of a hierarchical theory of contrast.

8.3 Dispersion-theoretic approaches to contrast

Another way of incorporating contrast into phonological theory has been explored in Dispersion Theory (Flemming 2002, 2004; Ní Chiosáin and Padgett 2001; Padgett 2003a, b), an approach that derives from the work of Liljencrants and Lindblom (1972) and Lindblom (1986). I will first briefly review the rudiments of Liljencrants and Lindblom's model (section 8.3.1), and then cite a simulation by D. C. Hall (1999) that suggests that their results are also consistent with a model that posits contrastive specifications plus enhancement.

In section 8.3.2 I present Padgett's account of the East Slavic post-velar fronting. Padgett (2003a) argues that the contrastive status of the velar consonants is a key to understanding a sound change in East Slavic whereby velars fronted before [ɨ]. While this insight is correct, I will take issue with another aspect of Padgett's analysis: whereas he sees the contrast between [i] and [u] as being a crucial part of the motivation for the change, I will propose an MCS analysis in which the relation between [i] and [ɨ] is what drives the change (section 8.3.3). I will argue (section 8.3.4) that Padgett's approach raises intractable issues of implementation, and that the MCS analysis is much simpler.

8.3.1 Liljencrants and Lindblom's dispersion model

Liljencrants and Lindblom (1972) hypothesized that functional perceptibility considerations might account for the shape of phonological inventories; thus, vowel inventories in which the vowels are well dispersed over the available space, such as /i, a, u/, tend to be common, whereas vowel inventories in which the vowels are less well distributed, such as /i, a, ɨ/, are rare or non-existent.

In Liljencrants and Lindblom's model, vowels are evenly placed on a circle within a simulated vowel space. The vowels then move away from each other as much as they can until an equilibrium is reached, where no further movement can increase the overall space between vowels. D. C. Hall (1999) attempts to replicate Liljencrants and Lindblom's results with his own program. Hall reports that his program, like theirs, selects /i, a, u/ as the optimal three-vowel inventory, providing that certain starting points are chosen: one vowel, which eventually becomes [u], starts with a lower second formant (F2) than the other vowels, and a second vowel, which ends up as [a], starts with a higher first formant (F1). Other starting points can produce unattested inventories.

D. C. Hall (2007: 146) proposes that the starting points can be viewed as phonological feature specifications; in the above example, /u/ is specified as [peripheral] (equivalent to [back] or [round]) and /a/ is specified as [low]. On this view, the program does the work of phonetic enhancement, assigning phonetic values to the phonological specifications that increase the salience of the contrastive features. Thus it is arguable that the basic results of Dispersion Theory can be captured at least as well in terms of contrastive features interacting with enhancement. As Hall (2007: 169) points out, enhancement has an advantage in that it does not require 'any sort of global comparison or evaluation of phonetic distance applied to the inventory as a whole, unlike the dispersion-theoretic approaches'.[1]

Building on the work of Liljencrants and Lindblom, Flemming (2002) proposes that inventories represent a compromise between two conflicting pressures: a pressure to maximize the number of contrasts in an inventory, and a conflicting pressure to keep members of the inventory maximally distinct. Thus, a nine-vowel inventory such as /i, ɪ, e, ɛ, a, ɔ, o, ʊ, u/ allows for many more lexical contrasts than the three-vowel inventory /i, a, u/; in the latter inventory, however, each vowel is allocated more of the vowel space than in the former.

Padgett (2003a, b) proposes that Dispersion Theory (DT) not only is a theory of inventories, but also plays a role in the workings of the phonology; in particular, he proposes an instantiation of DT which he argues is the way that considerations of contrast are brought to bear on the phonology. He illustrates this approach with an analysis of an East Slavic sound change known as post-velar fronting.

8.3.2 A dispersion-theoretic analysis of the East Slavic post-velar fronting

8.3.2.1 The velars and /i/ in Russian

Modern Russian consonants contrast in palatalization: palatalized consonants are paired with non-palatalized (or perhaps velarized) counterparts (3). The

1 See D. C. Hall (1999, 2007) for an extended discussion and critique of Liljencrants and Lindblom's (1972) dispersion model, and Kingston (2007) for a critical review of the issue.

consonants /j, ts, tʃ^j, ʒ, ʃ, ʃ^{j}ː/ are unpaired (/ʃ^{j}ː/ may actually be a sequence of consonants and not a phoneme; it does not act like the partner of /ʃ/).[2]

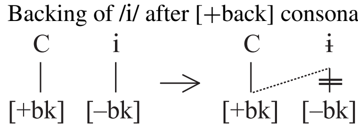
(3) Russian consonant phonemes

	Labial		Dental		Post-alveolar		Palatal	Velar	
Stop	p	p^j	t	t^j				k	k^j
	b	b^j	d	d^j				g	g^j
Fricative	f	f^j	s	s^j	ʃ	ʃ^{j}ː		x	x^j
	v	v^j	z	z^j	ʒ				
Affricate			ts			tʃ^j			
Nasal	m	m^j	n	n^j					
Lateral			l	l^j					
Rhotic			r	r^j					
Glide							j		

Russian has five vowel phonemes, /i, e, a, o, u/. Both palatalized and non-palatalized consonants can occur before back vowels, word finally, and pre-consonantally. Both palatalized and non-palatalized consonants may occur before /i/, which varies allophonically: it is pronounced [i] after palatalized consonants and [ɨ] after non-palatalized consonants.[3]

Padgett (2003a: 45) follows Farina (1991) in supposing that palatalized consonants are specified [−back], and 'plain' (actually velarized) consonants are specified [+back]. The vowels /i/ and /e/ are specified [−back], and /a/, /o/ and /u/ are specified [+back]. The allophony of /i/ is accounted for by the rule in (4).

(4) Backing of /i/ after [+back] consonants



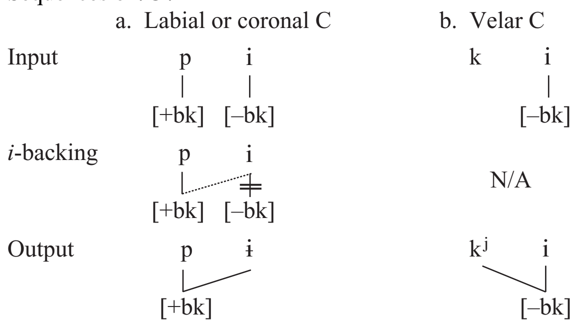
Velars /k, g, x/ behave differently from other consonants: they have front allophones before /i/ and /e/, and back allophones before /a, o, u/. Padgett observes that this patterning follows naturally if velars have no specification for [back] at the point that the backing rule (4) would apply, as shown in

2 Note that the terms 'paired' and 'unpaired' refer in this section to palatalization, not to voicing, as in previous sections.

3 Before /e/, only palatalized consonants may appear across word boundaries; some speakers have non-palatalized consonants before /e/ within roots in some loanwords (Padgett 2003a: 43; Timberlake 2004: 58).

(5).[4] Then *i*-backing fails to apply; instead, the [−back] /i/ causes the velar to palatalize.

(5) Sequences of /Ci/



Padgett (2003a) argues that this solution is not satisfactory, however, because it depends on specifying the underlying representation of /i/ as [−back]; according to the principle of richness of the base (Prince and Smolensky 2004), the grammar should give the correct output for any input, including an input /i/ that is [+back] (i.e., /ɨ/). Clearly, an underlying sequence /kɨ/ will result in surface [kɨ], not [kʲi].

8.3.2.2 Diachronic changes

Padgett proposes further that, whatever one may think of richness of the base, there was a time when /kɨ/ *was* the input to velar fronting. There was a time in the history of Russian when velars occurred before [ɨ] but not before [i] (6c). Between the twelfth and fourteenth centuries East Slavic underwent a change, post-velar fronting, whereby sequences like [kɨ], [ɡɨ] and [xɨ] fronted to [kʲi], [ɡʲi] and [xʲi], respectively (6d). This fronting did not cause a merger with sequences deriving from historical [kʲi], [ɡʲi] and [xʲi] (6a), because these sequences had mutated to palato-alveolars, such as [tʃʲi], prior to post-velar fronting (6b). This change, traditionally known as the First Velar Palatalization (FVP), opened up a gap in the inventory that could be exploited by post-velar fronting.

4 Given that the chart in (3) shows both palatalized and non-palatalized phonemic velars, readers may wonder why velars have no contrastive specification for [back]. We will see below that in earlier Russian there were no phonemic palatalized velars; a contrast was created through borrowings and analogy (Padgett 2003a: 46–7; Timberlake 2004: 59–60). In native words only non-palatalized velars occur before /i, e/ within morphemes, a legacy of the earlier situation. We could suppose that an underlying [–back] feature is deleted from velars in this context.

(6) Slavic sound changes

a. Prior to changes

pi	pɨ	pu
k^{j}i	kɨ	ku

b. First Velar Palatalization (Common Slavic)

pi	pɨ	pu
tʃ^{j}i		
	kɨ	ku

c. Rise of palatalized consonants (post-Common Slavic)

p^{j}i	pɨ	pu
tʃ^{j}i		
	kɨ	ku

d. Post-velar fronting (East Slavic)

p^{j}i	pɨ	pu
tʃ^{j}i		
k^{j}i		ku

8.3.2.3 The dispersion-theoretic analysis of Padgett (2003a)

Since Jakobson (1962a [1929]), there have been attempts to relate post-velar fronting to the contrastive status of the velars following FVP. Analysts differ, however, in the way they incorporate contrast into the analysis. Padgett (2003a) proposes that post-velar fronting occurred because [k^{j}i] makes a better contrast with [ku] than does [kɨ]. The labial and coronal consonants participate in a three-way contrast (represented by [p^{j}i] ~ [pɨ] ~ [pu]); therefore, the contrast between [pɨ] and [pu] cannot be 'improved' by fronting [pɨ] to [p^{j}i] without neutralizing the contrast between [pɨ] and [p^{j}i]. Following FVP, there were no longer sequences [k^{j}i] in contrast with [kɨ] or [ku].

To implement this analysis, Padgett follows Flemming's 1995 dissertation (published in revised form as Flemming 2004) in assuming that possible inputs and candidate forms within OT can include not only individual forms, but sets of forms. In Ní Chiosáin and Padgett's (2001) interpretation, the objects of analysis are taken to be entire languages. Padgett (2003a: 51) writes that 'this daunting prospect is made manageable by means of extreme idealization'. The idealization starts by limiting the set of relevant forms to the ones in (7).

(7) Set of East Slavic 'words' (Padgett 2003a: 53)

pi	pɨ	pu	pau
p^{j}i	p^{j}ɨ	p^{j}u	p^{j}au
tʃi	tʃɨ	tʃu	tʃau
tʃ^{j}i	tʃ^{j}ɨ	tʃ^{j}u	tʃ^{j}au
ki	kɨ	ku	kau
k^{j}i	k^{j}ɨ	k^{j}u	k^{j}au

Padgett posits a family of SPACE constraints that penalize sets of forms that do not allow for sufficient perceptual contrast along designated dimensions.

As part of the extreme idealization of his analysis, Padgett restricts attention to the colour dimension, that is the properties of backness and roundness that are primarily signalled by the second vowel formant. The SPACE constraint employed in his analysis is defined as in (8).

(8) SPACE constraint for East Slavic (Padgett 2003a)
SPACE$_{(\text{COLOUR})}$ $\geq$ 1/2: 'Potential minimal pairs differing in vowel colour differ by at least 1/2 of the full vowel colour range.'

The definition of the SPACE constraint in (8), which is one instantiation of the general definition that would apply for any dimension and any portion of the range, refers to 'potential minimal pairs', which Padgett defines as in (9).

(9) Potential minimal pairs (Padgett 2003a: 54)
A potential minimal pair is a pair of words having the same number of segments, and all but one of whose corresponding segments are identical.

It would take us too far afield to review Padgett's entire analysis, but it will suffice to focus on the crucial step in which post-velar fronting took place, which Padgett represents as in (10).

(10) Post-velar fronting: SPACE ≫ IDENT(COLOUR) (Padgett 2003a: 74)

		*MERGE	SPACE	ID-COL
	pi$_1$ pɨ$_2$ pu$_3$ kɨ$_5$ ku$_6$ tʃʲi$_4$			
a.	pi$_1$ pɨ$_2$ pu$_3$ kɨ$_5$ ku$_6$ tʃʲi$_4$		***!	
b. ☞	pi$_1$ pɨ$_2$ pu$_3$ kʲi$_5$ ku$_6$ tʃʲi$_4$		**	*
c.	pi$_1$ pɨ$_2$ pu$_3$ kʲi$_5$ kɨ$_6$ tʃʲi$_4$		***!	**
d.	pi$_{1,2}$ pu$_3$ kʲi$_5$ ku$_6$ tʃʲi$_4$	*!		**

Padgett (2003a) supposes that the input to (10) is essentially the set of surface forms in (6b), a stage after FVP, but immediately before post-velar fronting (Padgett (2003a: 73 n. 25) notes that he omits the palatalization in [pi$_1$], though this form was undoubtedly palatalized). He proposes that what precipitated post-velar fronting was a reranking of the constraints SPACE and IDENT(COLOUR). Prior to this reranking, IDENT(COLOUR) ensured that an underlying /kɨ/ would surface as such.[5] Following the reranking, the faithful candidate (10a) becomes less optimal than (10b), with [kʲi] as the surface correspondent to input /kɨ/, because the latter has better separation from [ku]. Candidate (10d) has optimal spacing on the colour dimension, but violates *MERGE, which penalizes any surface mergers of forms that are underlyingly distinct.

The intuition behind the DT analysis is that the trigger for post-velar fronting is the possibility of input /kɨ/ gaining better separation from /ku/ by fronting. To implement this idea formally, Padgett must resort to a series of ‘extreme idealizations’ that are rather problematic, for reasons I will discuss below. First, however, I will present an alternative analysis of post-velar fronting that also crucially depends on the contrastive status of the sequence /kɨ/, but which puts the emphasis on the relationship between [i] and [ɨ] rather than on the relationship of these vowels to [u].

8.3.3 A Modified Contrastive Specification analysis of post-velar fronting

The theory of MCS allows for a different solution to post-velar fronting, one that does not require the problematic selection of sets of inputs, and which is actually closer to the spirit of Jakobson’s 1929 analysis. Following standard chronology, I assume the sequence of grammars in (11), starting with post-Common Slavic at a point after FVP but before the East Slavic post-velar fronting. In this stage, which I arbitrarily designate as Stage 1, vowels but not consonants have contrastive values of [back] (11). Palatalization applies to spread [−back] from /i/ to a preceding consonant.

(11) Stage 1: Vowels, not consonants, are contrastively [back]

Underlying	/p i/	/p ɨ/	/tʃ i/	/k ɨ/
	[–bk]	[+bk]	[–bk]	[+bk]
Palatalization	pʲ i	—	tʃʲ i	—
	[–bk]		[–bk]	
Phonetic	[pʲi]	[pɨ]	[tʃʲi]	[kɨ]

A major event in the history of Slavic was the fall of the jers. Jers were short vowels, one front and one back; like other front vowels, the front jer

5 Subscripts in (10) indicate input–output correspondents. Thus, input *kɨ*$_5$ corresponds to output *kɨ* in candidate (a), and to $k^j i$ in candidates (b)–(d). The subscript ‘$_6$’ on the output form *kɨ* in candidate (c) indicates that it corresponds to input *ku*.

caused palatalization. When the jers fell, palatalization that had been triggered by the front jer became opaque as a synchronic process. This situation led to a reanalysis of palatalized consonants as underlying (Jakobson 1962a [1929]: 57; Shevelov 1964: 497). Now vowels as well as most consonants had contrastive values for [back], as shown in (12).

(12) Stage 2: Vowels and paired consonants are contrastively [back]

Underlying	/pʲ i/	/p ɨ/	/tʃʲ i/	/k ɨ/
	[–bk][–bk]	[+bk][+bk]	[–bk][–bk]	[+bk]
Palatalization	—	—	—	—
Phonetic	[pʲi]	[pɨ]	[tʃʲi]	[kɨ]

In the grammar (12), the velars, unlike the labials and dentals, do not have contrastive values for [back]. This is because velars are unpaired, original [kʲ] having become [tʃʲ]. Palatalization does not apply synchronically in the forms in (12), but presumably remains in the grammar to account for palatalization of consonants by front vowels across morpheme boundaries.

By this stage the contrastive status of [ɨ] was becoming unclear: in most cases, [i] follows a front consonant and [ɨ] follows a back consonant. Other developments, such as the disappearance of [ɨ] at the beginning of words, led to [i] and [ɨ] being in complementary distribution: [ɨ] followed back consonants and [i] occurred elsewhere. This regularity led to a reanalysis of the contrastive status of [ɨ], which ceased being an independent phoneme, and became a conditioned allophone of /i/, as shown in (13) (Jakobson 1962a [1929]: 70; Shevelov 1964: 503).

(13) Stage 3: /ɨ/ reanalysed as [+back] allophone of [–back] /i/

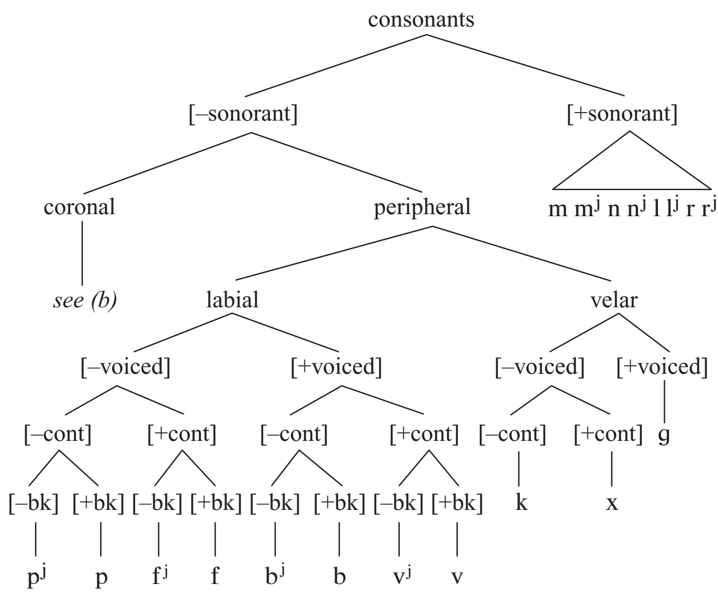
Underlying	/pʲ i/	/p i/	/tʃʲ i/	/k i/
	[–bk][–bk]	[+bk][–bk]	[–bk][–bk]	[–bk]
i-Backing	—	p i [+bk][–bk]	—	—
Palatalization	—	—	—	kʲ i [–bk]
Phonetic	[pʲi]	[pɨ]	[tʃʲi]	[kʲi]

Once [ɨ] was reanalysed as a [+back] allophone of [−back] /i/, it required an adjacent [+back] consonant, or some other donor, to give it a [+back] feature; lacking that, underlying /i/ would surface as [i]. Labial and dental consonants had a contrastive [+back] feature to spread, but velars did not. Hence, an original sequence [kɨ], once reanalysed as deriving from /ki/, would begin to surface as [k^{j}i].

The above analysis requires a contrastive hierarchy for East Slavic along the lines of (14), generating the tree in (15).[6]

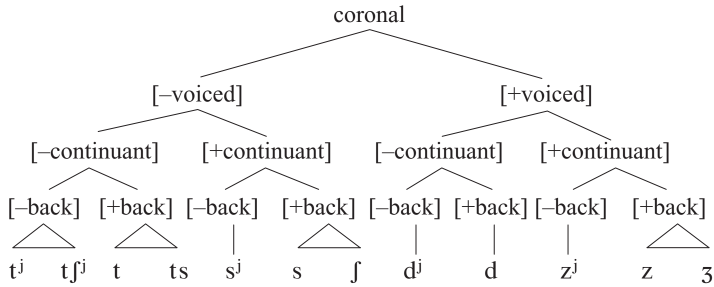
(14) Contrastive feature hierarchy for East Slavic obstruents
[sonorant] > major place features > [voiced] > [continuant] > [back] > other coronal features

(15) Contrastive feature tree for East Slavic obstruents
a. Labials and velars



6 I disregard here the special status of Russian /v/ with respect to voicing and devoicing assimilation (Hayes 1984; Kiparsky 1985; Padgett 2002; D. C. Hall 2007). Hall (2007: 65–6) proposes that Russian has three classes of consonants: sonorants, specified for [Sonorant Voice]; obstruents, specified for [Laryngeal]; and /v/, which has neither specification.

b. Coronals



In the analysis sketched above what drives the fronting of [kɨ] to [k^ji] is the contrastive status of the velar together with the reanalysis of [ɨ] as a conditioned allophone of [−back] /i/.[7] The distance of [ɨ] from [u] plays no direct role in the analysis, hence there is no need for the phonology to compute the relative spacing of vowels from each other.[8]

8.3.4 Comparing the analyses

Padgett's DT analysis and the MCS analysis presented above have in common that they connect the East Slavic post-velar fronting to the contrasts obtaining in the East Slavic inventory. However, the way in which this idea is implemented in each theory is very different. I believe that the version of DT presented in Padgett (2003a) runs into significant conceptual and technical difficulties.

8.3.4.1 Richness of the base

Padgett (2003a: 47) rejects any solution that requires the input to post-velar fronting to be specified as /i/ (and not /ɨ/) on grounds of richness of the base.

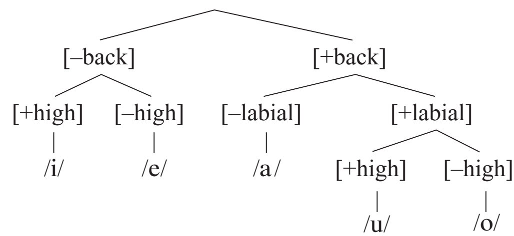
As discussed in section 6.4, in the conception of phonological organization proposed here the appropriate place for richness of the base to be observed is at the phonological level that produces well-formed contrastive specifications as its output. A learner of East Slavic at the stage just prior to post-velar fronting (13), for example, would acquire the contrastive inventory of that

7 This analysis is close in spirit to Jakobson's. According to Jakobson (1962c: 70), once the vowels [*i*] and [*ɨ*] were transformed into allophones of a single phoneme, there was a tendency to unify the phoneme, which manifested itself by generalizing the fundamental allophone after an unpaired (i.e., velar) consonant. In other words, the underlying contrastive features of a phoneme (the 'fundamental allophone', or basic variant) will tend to surface in the absence of a process that would act to change them.

8 The presence of /u/ in the inventory plays an indirect role, for it enters into a determination of the system of contrasts for the language.

stage. The above analysis assumes a partial contrastive hierarchy for the five vowel phonemes /i, e, a, o, u/ as shown in (16).[9]

(16) East Slavic partial contrastive hierarchy: [back] > [labial] > [high]



This hierarchy translates into the OT constraint hierarchy in (17).

(17) Constraint hierarchy corresponding to (16)
Max [back] ≫ *[labial, −back] ≫ Max [labial] ≫ *[high, −labial] ≫ Max [high] ≫ *[F]

Richness of the base holds at the level of the input to the constraint hierarchy in (17). Any vowel that is input to the constraints in (17) will emerge as one of the five contrastive vowel phonemes in (16). These vowels, in turn, serve as the input to the phonology proper. At this point, richness of the base is no longer a factor. Therefore, there can be no input /kɨ/ to the phonology of East Slavic at the stage immediately preceding post-velar fronting.

Padgett (2003a) must assume something similar. In order to be able to describe a sequence of sound changes, he adopts, 'as an expository convenience', the synchronic base hypothesis (Hutton 1996; Holt 1997), which holds that the input at each historical stage is the output of the previous stage.

This hypothesis appears to be too limited. Taken literally, the synchronic base hypothesis is a return to the Neogrammarian conception of sound change as applying to surface forms (see Dresher 1993 for discussion). In the view of classical generative phonology, which I adopt here, sound change must be understood relative to the entire grammar. Some sound changes may cause restructuring of underlying forms, so that the input (underlying form) of one stage approximates the output (surface form) of an earlier stage. More typically, however, sound change does not result in a complete restructuring of underlying forms, but rather remains in the grammar as a synchronic rule.

Post-velar fronting provides a good illustration of this principle. In the MCS analysis, the immediate input to the stage that produced post-velar fronting,

9 This hierarchy is underdetermined by the data we have seen: it could be that [low] was higher in the order than [labial], for example. What is important for our analysis is that /i/ is contrastively [−back].

/ki/, is *not* the output of the previous stage, which was [kɨ]. In this case, the important change occurred at the underlying level: the previously independent phoneme /ɨ/ was reanalysed as an allophone of /i/.

Leaving aside the adequacy of the synchronic base hypothesis, the essential point is that in Padgett's DT analysis, as in the MCS analysis, the input to the relevant part of the phonology does not adhere to richness of the base. Therefore, Padgett's original objection to the analysis in (5) loses its force.

8.3.4.2 'Extreme idealization'

The evaluation of sets of input and candidate forms, required in order to implement the DT analysis, is an obvious departure from standard OT, and from standard generative phonology more generally. While evaluating a set of forms takes more resources than evaluating single forms, what I wish to focus on here is the arbitrary nature of the set of forms in (7). These forms were chosen because they have just the properties that allow the analysis to proceed.

Padgett (2003a: 50–3) defends this approach as being an idealization that is common practice in phonology: 'Limiting the words considered for an analysis actually makes explicit what is implicit in the practice of phonology.' He compares a phonologist demonstrating an analysis of English aspiration by choosing only a small number of words and relevant candidate forms, say the word *pat* and the three candidates [pʰæt], [pæt] and [bæt]. Padgett points out that the analyst is unlikely to choose words or forms that do not effectively illustrate the important aspects of the analysis.

While this is certainly true, this case is entirely different from what Padgett is proposing. In the example of English aspiration, the limitation occurs only in the *demonstration* of the analysis: there is no limitation on what forms may be *considered by the grammar*. A correct analysis of English aspiration ought to give the correct results for any randomly chosen list of English words. Padgett's 'extreme idealization' goes beyond the mere demonstration of his analysis: it is built into the evaluation procedure of the grammar. To compare with the English aspiration example, Padgett would have to show that the same results would be obtained if other sets of forms were presented to the grammar.

It is hard to see how this could be the case, since the DT analysis limits not only the forms that could be considered, but the dimensions of contrast as well. Thus, the only SPACE constraint that is allowed to play a role in the analysis is $\text{SPACE}_{(\text{COLOUR})}$, which evaluates vowel spacing along the second formant. But other dimensions of contrast also exist, some of them quite relevant to this case. For example, the grammar in (10) rates candidate (b) as better than candidate (a) because [kʲi, ku] is a better contrast than [kɨ, ku]. At the same time, though, [kʲi, tʃʲi] is a *worse* contrast than [kɨ, tʃʲi]. It has not been shown that this contrast is less important. Indeed, we could argue the opposite: while there is no real evidence that there is something problematic about keeping [kɨ] distinct from [ku], we know from the history of Slavic that [k^ji] *is* liable to be turned into [$tʃ^j$i]. It was this very change that created the gap in the inventory in the first place.[10] Therefore, it is unclear that the DT analysis can go through if we take into account other dimensions of contrast. This, again, is very different from the example of English aspiration.

8.3.4.3 Minimal pairs

The SPACE constraints in the DT analysis refer crucially to the notion of 'potential minimal pairs', that is, words distinguished by a single segment. Just as phonemes distinguished by a single phonetic property are hard to find at the phonetic level, so surface minimal word pairs are more elusive than one might suppose.

First, the existence of genuine minimal pairs that satisfy the definition in (9) is greatly compromised by phonetic effects. As pointed out by Chomsky (1964: 94), the substitution of a segment [Q] in place of [P] in the phonetic frame [RPS] will not necessarily result in [RQS], since [Q] might affect the neighbouring segments differently from [P]. The result is more likely to be [R′QS′], where R′ and S′ differ from R and S, respectively. An example arises in (10), where [pi, pɨ, pu] are supposed to differ minimally only in the vowel. However, [pi] is in fact [p^ji], which does not form a minimal pair with either [pɨ] or [pu]. Strictly speaking, then, the SPACE constraint would not be able to evaluate the separation between [p^ji] and [pɨ] or [k^ji] and [ku].

Second, although the definition in (9) refers to 'a pair of words', Padgett (2003a: 78–9) makes clear that the SPACE and *MERGE constraints do not operate with actual East Slavic words, but with *possible* words. For if they evaluated actual words, the analysis would predict that post-velar fronting would occur only in cases where a word containing the sequence [kɨ] actually formed a minimal pair with a word containing the sequence [ku]. In words containing [kɨ] for which there was no minimal pair with [ku], SPACE would evaluate the vowel [ɨ] as having 100 per cent of the colour space to itself, in which case [ɨ] is actually the optimal vowel. In fact, post-velar fronting applied to all cases of [kɨ]. As Padgett (2003a: 79) correctly observes, 'the absence of forms such as [k^ji] was a systematic gap, not an accidental one'. But then the existence of minimal pairs as defined in (9) plays no role in this change.

10 Padgett (2003a: 59) cites Guion (1998) as showing that [k^j] is easily confusable with [$tʃ^j$].

8.3.4.4 Why are velars treated separately?

In comparing the merits of the DT and MCS analyses it is necessary to clarify what each analysis explains and leaves unexplained. Up to here I have been pointing out problems for the DT analysis, so fairness requires that we subject the MCS analysis to comparable scrutiny.

Let us look again at the proposed East Slavic contrastive hierarchy in (14) and (15). If this analysis is a descriptively adequate hypothesis about the phonology of East Slavic, then (14) should be part of the grammar internalized by native speakers. As in other cases, however, the MCS analysis does not answer how a learner would know to arrive at this particular feature ordering. To that extent, the analysis does not meet the criterion of explanatory adequacy, as set out in Chomsky (1957): given the data of a language, we cannot explain how learners are led to the (presumably) correct grammar.

Let us consider in particular why the velars are not contrastively specified for the feature [back]. This follows from the ordering of the major place features, as well as [voiced] and [continuant], ahead of [back]. If the feature [back] were ordered higher in the hierarchy, then the velar consonants would also receive contrastive specifications for [back], despite the fact that they are 'unpaired'. Therefore, how the features come to be ordered in the right way is an unexplained step in the analysis. One might suspect, then, that some of the complexity of Padgett's DT account could arise from an effort to make explicit this aspect of the analysis. That is, the learner who in the MCS theory has to arrive at a contrastive hierarchy must be making some sort of comparisons between forms; perhaps we could understand the evaluation of idealized sets like (7) as an attempt to explain how this sort of comparison works.

Examination of the DT analysis shows, however, that this is not at all the case. The DT analysis *also* has to assume, without explanation, that labials and coronals are somehow treated together, but that velars are distinct. Thus, Padgett (2003a: 52–3) writes in connection with the set of forms in (7):

> I am assuming that only the kinds of distinctions evident in this group of forms are relevant to an analysis of the sound changes of interest. For example, it will be important to treat velars on the one hand separately from dentals and labials on the other. In this idealization, non-velars are represented by [p]. Differences among the various labials and dentals of Slavic are not relevant to the analysis.

In other words, the DT analysis does not attempt to explain how it comes about that velars are treated separately from labials and dentals: rather, it presupposes that this is so, and builds the difference into the representations presented to the grammar. Therefore, the DT analysis has no advantage, with respect to explanatory adequacy, to compensate for its descriptive problems.

Given that neither theory (nor any other one, to my knowledge) can explain how learners of East Slavic come to know that velars should be regarded as noncontrastive for [back], the advantage shifts back to the MCS analysis. For in the MCS theory, it is a given that learners must arrive at *some* contrastive hierarchy for their language. Moreover, this same hierarchy must account not just for the behaviour of velar consonants and the feature [back], but also for the other phonemes and features as well. Thus, multiple sources of evidence exist bearing on the construction of the contrastive hierarchy of a language (see, further, section 8.7).

It is not clear whether there are comparable limits on DT theory. The idealizations made to account for post-velar fronting are particular to that problem. Nothing in the theory as presented in Padgett (2003a) would prevent us, for example, from grouping together East Slavic velars and labials against coronals if that suited a different problem.

8.4 Structured Specification theory and contrast via natural classes

Structured Specification, or natural classes, theory (Broe 1993; Frisch 1996; Frisch, Pierrehumbert and Broe 2004) proposes that phonological processes are sensitive to similarity: harmony processes and co-occurrence restrictions affect segments that are highly similar. A similarity metric is derived from the natural classes created by the phonemes of an inventory; it measures similarity by computing the feature classes shared by two phonemes, and dividing by the number of shared plus unshared classes. This method of computing similarity weights features differently depending on how much they contribute to creating distinct natural classes. Features that uniquely distinguish members of the inventory will contribute more to the similarity metric than logically redundant features; in this way, the computation indirectly takes language-particular contrasts into account. But it does not designate features as being contrastive or redundant, and it arrives at the same computation for every similar inventory.

For example, suppose we have a vowel system /i, e, a, o, u/. Let us assume the seven features in (18).[11]

11 For the purposes of this demonstration I adopt monovalent features capable of expressing all natural classes, as required by Frisch (1996). For this example I assume there is no feature [non-round] (i.e., [round] is privative).

(18) Features for a five-vowel system

a. [syllabic]	/i, e, u, o, a/				
b. [high]	/i, u/	c. [low]	/a/	d. [open]	/e, o/
e. [front]	/i, e/	f. [round]	/o, u/	g. [back]	/a, o, u/

The number of *potential* classes (i.e., mathematically possible classes, not necessarily existing or even phonologically possible) consists of the power set of the set of features minus the empty set: that is, all the sets that can be formed from combinations of these features. These sets include one set consisting of all seven features (not possible in this example, since some features are antagonistic), seven sets consisting of six features (leaving out one of the seven features in each one), and so on. The number of such classes for *n* features is $2^n - 1$; for $n = 7$, this comes to 127 classes. Of these we can eliminate phonologically impossible classes (e.g., those with [high] and [low], or [front] and [back], and so on). Of the possible potential classes, only those with distinct members make up a natural class. In our example, there are 11 natural classes (19).

(19) Natural classes derived from (18)

a. [syll, open, round, back]	/o/	b. [syll, open, front]	/e/
c. [syll, high, round, back]	/u/	d. [syll, high, front]	/i/
e. [syll, low, back]	/a/		
f. [syll, round, back]	/o, u/	g. [syll, front]	/i, e/
h. [syll, open]	/e, o/	i. [syll, high]	/i, u/
j. [syll, back]	/a, o, u/		
k. [syll]	/i, e, a, o, u/		

To see how the inventory influences the number of natural classes, consider /a/ (19e), which is characterized as [syllabic, low, back]. As /a/ is the only low vowel in the inventory, it is uniquely characterized as [low], as [syllabic, low], and as [low, back]. But because all these classes have the same membership, they count as only one natural class, represented by the most fully specified description.

The similarity measure for each pair of phonemes is shown in (20). It rates /u, o/ as the most similar pair, more similar to each other than the seemingly parallel /i, e/. The parallelism is disturbed by the fact that /u, o/ are both [back] as well as [round], whereas /i, e/ share only [front]. Therefore, /u, o/ share 3 classes ([syllabic, round, back], [syllabic, back] and [syllabic]), whereas /i, e/ share only 2 ([syllabic, front] and [syllabic]).

(20) Similarity measures (S = shared classes, U = unshared classes)

Pair	*S*	*U*	*S/S+U*	*Pair*	*S*	*U*	*S/S+U*
a. /u, o/	3	4	0.43	b. /i, e/	2	4	0.33
c. /o, a/	2	4	0.33	d. /u, a/	2	4	0.33
e. /e, o/	2	5	0.29	f. /i, u/	2	5	0.29
g. /i, a/	1	5	0.17	h. /e, a/	1	5	0.17
i. /i, o/	1	7	0.13	j. /u, e/	1	7	0.13

The classes [syllabic, round, back] and [syllabic, back] are distinct because /a/ is a member of the latter, but not the former. In the case of unshared classes, the features [round] and [back] function together, because /a/ is neither [high] nor [open], in this system of features, so [syllabic, high/open, round, back] = [syllabic, high/open, round] = [syllabic, high/open, back].

The features [back] and [round] do similar contrastive work in most classes; therefore, the weight they are accorded by the similarity measures is reduced, somewhat capturing the notion that these features make each other logically redundant in most cases, without actually designating one as contrastive and the other as redundant. Though their weight is reduced, they are not entirely disregarded: thus, /u, o/ are computed to be more similar to each other than /i, e/.

This procedure is very sensitive to the choice of features: different choices of features can change the results quite substantially. If we had a feature [non-round], for example, the front and back vowels would appear to be more symmetrical. The choice of a feature [open], defined as applying only to mid vowels, is responsible for the somewhat counterintuitive result that /i/ and /e/ are equally similar to /a/, as are /u/ and /o/. On the other hand, since the feature set is presumably universal, whatever set is chosen will give the same results to every inventory with these phonemes. This poses a problem if, as we have seen, different five-vowel inventories have different contrastive structures.

Frisch, Pierrehumbert and Broe (2004) propose that similarity as computed using natural classes is successful in accounting for the co-occurrence restrictions (what they call 'OCP effects') on consonantal roots in Arabic. This is so, they argue, because all features contribute in some measure to these restrictions. However, the most basic fact about the Arabic (and Semitic, more generally) constraints is inconsistent with Structured Specification theory. As they write (Frisch, Pierrehumbert and Broe 2004: 198), 'Since OCP effects in Arabic only apply to consonants that share major place of articulation features, we stipulate that the natural classes used in the similarity computations are only those natural classes containing a place of articulation feature. Thus, nonhomorganic consonant pairs will have similarity 0...' This type of stipulation contradicts the basic assumption that the restrictions are sensitive to overall similarity as computed by natural classes; rather, it is more naturally accounted for in a theory where place features are hierarchically superior to other features. That is, certain features count more than others with respect to specific phonological processes and constraints.

In the following subsections I apply Structured Specification theory to examples of vowel harmony, consonant co-occurrence and loan phonology discussed above. I will show that, if these processes are supposed to affect segments that are most similar to each other, then the similarity metric makes the wrong predictions in each case.

8.4.1 Vowel harmony

We observed (section 7.4) that the vowels that trigger ATR and labial harmony in Classical Manchu are just those that are contrastively specified for the [ATR] and [labial] features, respectively. Specifically, the vowels /u/ and /ə/ are specified for a contrastive [ATR] feature, and may not co-occur with /ʊ/ and /a/, which are contrastively non-ATR. The vowel /i/ is not contrastively specified for [ATR], and may co-occur with any of the above vowels. The feature [labial] contrastively distinguishes between /ɔ/ and /a/, and these are the only vowels that participate in labial harmony in Classical Manchu.

On the assumption that ATR harmony targets vowels that are highly similar to each other, we expect that /u/, /ə/, /ʊ/ and /a/ should all be more similar to each other than any of them are to /i/. Of course we expect /u/ to be very similar to /ʊ/, and /ə/ to be very similar to /a/, since these vowels are distinguished only by the feature [ATR]; any feature-based theory would derive this result. The test of the theory is whether /u/ and /ʊ/ are also similar to /ə/ and /a/.

To test this prediction, I computed similarity measures for the vowels of Classical Manchu, using the features in (21), which mirror the features used by Zhang (1996). Following the usual practice in Structured Specification theory (Broe 1993; Frisch 1996), I have given each marked feature an unmarked counterpart, so that all classes of vowels can be picked out. The similarity measures are shown in (22).

(21) Vowel features for Classical Manchu

a. [syllabic]	/i, ə, a, ɔ, u, ʊ/		
b. [back]	/ə, a, ɔ, u, ʊ/	c. [front]	/i/
d. [low]	/ə, a, ɔ/	e. [non-low]	/i, u, ʊ/
f. [round]	/ɔ, u, ʊ/	g. [non-round]	/i, ə, a/
h. [ATR]	/i, ə, u/	i. [non-ATR]	/a, ɔ, ʊ/

(22) Similarity measures for Classical Manchu

	Phonemes	*Similarity*	*Harmony*		*Phonemes*	*Similarity*	*Harmony*
a.	/ɔ, ʊ/	0.45	None	b.	/a, ɔ/	0.45	Labial
c.	/u, ʊ/	0.42	ATR	d.	/ə, a/	0.42	ATR
e.	/i, u/	0.33	None	f.	/i, ə/	0.33	None
g.	/ə, u/	0.29	ATR	h.	/a, ʊ/	0.23	ATR
i.	/ɔ, u/	0.21	None	j.	/ə, ɔ/	0.21	Labial
k.	/i, ʊ/	0.15	None	l.	/i, a/	0.15	None
m.	/ə, ʊ/	0.13	ATR	n.	/a, u/	0.13	ATR
o.	/i, ɔ/	0.07	None				

It is evident from the results in (22) that the similarity between two phonemes as computed by this metric has no bearing on whether or not these segments may co-occur in a word. We can set no threshold above which harmony is required. To explain why certain vowels participate in harmony, we require a theory that picks out which features are contrastive.

8.4.2 Consonant co-occurrence

It has been proposed (Hansson 2001; Rose and Walker 2004) that consonant harmony and other restrictions on the co-occurrence of consonants depend on similarity, as computed by Structured Specification theory. Mackenzie (2005, 2009) argues, however, that this theory rather consistently fails to account for common types of consonant harmony and co-occurrence.

Recall Bumo Izon (section 7.6), in which labial and alveolar voiced stops must be all pulmonic or all implosive in a word; the velar pulmonic voiced stop and the labiovelar implosive consonant do not participate in this restriction. We saw that the phonemes that participate in implosive harmony are exactly the ones that are contrastively specified for the harmonizing feature, [glottalic] (Mackenzie 2005). Hansson (2001) observes that the Structured Specification / natural classes approach will not work for this case. The following demonstration of why this is so is based on the discussion in Mackenzie (2005).

For the purposes of computing natural classes, let us assume the features in (23).[12] Similarity computations for this inventory are shown in (24).

12 I observed above that the similarity metric is very sensitive to the choice of features, and different numbers are obtained if we make different decisions about what the features are. For example, the labiovelars can be considered to be combinations of [labial] and [velar] rather than as [labiovelar]. I have not found any combination of features that gives the required results.

(23) Features for Bumo Izon stops

a. [stop]	/p, b, ɓ, t, d, ɗ, k, g, kp, ɠɓ/		
b. [labial]	/p, b, ɓ/	c. [coronal]	/t, d, ɗ/
d. [velar]	/k, g/	e. [labiovelar]	/kp, ɠɓ/
f. [pulmonic]	/p, b, t, d, k, g, kp/	g. [glottalic]	/ɓ, ɗ, ɠɓ/
h. [voiced]	/b, ɓ, d, ɗ, g, ɠɓ/	i. [voiceless]	/p, t, k, kp/

(24) Similarity measures for Bumo Izon

Phonemes	*Similarity*	*Harmony*	*Phonemes*	*Similarity*	*Harmony*
a. /ɗ, d/	0.4	Yes	b. /ɓ, b/	0.4	Yes
c. /ɠɓ, g/	0.22	No			
d. /ɗ, g/	0.2	No	e. /ɓ, g/	0.2	No
f. /ɠɓ, d/	0.18	No	g. /ɠɓ, b/	0.18	No
h. /ɗ, b/	0.17	Yes	i. /ɓ, d/	0.17	Yes

As already pointed out by Hansson (2001), the similarity metric of Structured Specification theory makes the wrong predictions in these cases. For example, it rates /g/ as more similar to implosive /ɓ/ than /d/ is, and more similar to implosive /ɗ/ than /b/ is. It thus incorrectly predicts that /g/ and /ɠɓ/ should be the most likely of all the voiced stops to participate in implosive harmony with stops at different places of articulation.

As with vowel harmony, the simplest account appeals to contrastive features as derived by the contrastive hierarchy. No better alternatives have been proposed, to my knowledge.

8.4.3 Loanword adaptation

We observed in section 7.7.2 above that English /s/ is borrowed into Hawaiian as /k/, but into NZ Māori as /h/. In Herd's (2005) analysis, these adaptations follow from the contrastive hierarchies of these languages. If these adaptations are based on perceived similarity, then /s/ is more similar to /k/ than to /h/ in Hawaiian, but more similar to /h/ than to /k/ in NZ Māori. Do the different inventories of these languages yield these results in Structured Specification theory?

To test these predictions, I assumed the features in (25) for Hawaiian, with /s/ added to the inventory for purposes of comparison. The results are shown in (26). The similarity metric rates /s/ as most similar to /h/ or to /l/; the expected winner, /k/, ranks third.

(25) Features for Hawaiian consonants + /s/

a. [non-syllabic]	/p, m, w, n, l, k, ʔ, h, s/		
b. [consonantal]	/p, m, n, l, k, s/		
c. [sonorant]	/m, w, n, l/	d. [obstruent]	/p, k, ʔ, h, s/
e. [labial]	/p, m, w/	f. [coronal]	/n, l, s/

g.	[dental]	/n/	h.	[non-dental]	/l, s/
i.	[dorsal]	/k/	j.	[laryngeal]	/ʔ, h/
k.	[stop]	/p, m, n, k, ʔ/	l.	[continuant]	/w, l, h, s/
m.	[spread]	/h/	n.	[constricted]	/ʔ/
o.	[lateral]	/l/	p.	[liquid]	/l/
q.	[nasal]	/m, n/			

(26) Similarity measures for Hawaiian consonants and /s/

	Phonemes	*Similarity*		*Phonemes*	*Similarity*
a.	/s, h/	0.36	b.	/s, l/	0.36
c.	/s, k/	0.29	d.	/s, p/	0.25
e.	/s, n/	0.19	f.	/s, ʔ/	0.15
g.	/s, w/	0.14	h.	/s, m/	0.11

Assuming the feature specifications in (27), the results for NZ Māori are even worse (28). As in Hawaiian, the similarity metric rates /s/ as more similar to /h/ than to /k/. In NZ Māori this is a good result, because English words with /s/ are adapted as /h/; the problem is that /h/ is ranked as only the fourth most similar phoneme to /s/, behind /f/, /t/ and /r/.

(27) Features for NZ Māori consonants + /s/

a. [non-syllabic] /p, f, m, w, t, n, r, k, ŋ, h, s/
b. [consonantal] /p, f, m, t, n, r, k, ŋ, s/

c.	[sonorant]	/m, w, n, r, ŋ/	d.	[obstruent]	/p, f, t, k, h, s/
e.	[labial]	/p, f, m, w/	f.	[coronal]	/t, n, r, s/
g.	[dental]	/t, n/	h.	[non-dental]	/r, s/
i.	[dorsal]	/k, ŋ/	j.	[laryngeal]	/h/
k.	[stop]	/p, m, t, n, k, ŋ/	l.	[continuant]	/w, f, r, h, s/
m.	[spread]	/h/	n.	[non-lateral]	/r/
o.	[liquid]	/r/	p.	[nasal]	/m, n, ŋ/

(28) Similarity measures for NZ Māori consonants and /s/

	Phonemes	*Similarity*		*Phonemes*	*Similarity*
a.	/s, f/	0.47	b.	/s, t/	0.38
c.	/s, r/	0.35	d.	/s, h/	0.31
e.	/s, k/	0.25	f.	/s, p/	0.21
g.	/s, n/	0.16	h.	/s, ŋ/	0.11
i.	/s, w/	0.11	j.	/s, m/	0.10

8.4.4 Summary

Although Structured Specification theory and the natural classes method of computing similarity are interesting proposals for taking contrast in inventories into account without designating features as contrastive, the results surveyed above suggest that typical phonological processes are not sensitive to similarity as computed in this way. In every case, the Contrastivist Hypothesis, implemented by the contrastive hierarchy, provides simple successful accounts of the phonological patterns, suggesting that phonology is particularly sensitive to contrastive feature specifications.

8.5 Clements's approach to contrastive specification

Clements (2001, 2003, 2009) has been developing an approach to contrastive feature specification that has some affinities with MCS, as well as some important differences. Like MCS, Clements (2001) proposes that only *active features* are specified in the phonology, where his notion of 'active feature' is essentially the same as the one assumed here. However, Clements does not tie activity to contrast, as supposed in the strong version of the Contrastivist Hypothesis. Clements proposes three conditions for feature specification, corresponding to the lexical, phonological and phonetic levels, respectively (29).

(29) Conditions for feature specification (Clements 2001: 77–8)

a. Lexical level: distinctiveness
A feature or feature value is present in the lexicon if and only if it is distinctive (in a sense to be defined).

b. Phonological level: feature activity
A feature or feature value is present at a given phonological level if it is required for the statement of phonological patterns (phonotactic patterns, alternations) at that level.

c. Phonetic level: pronounceability
Feature values are present in the phonetics if required to account for relevant aspects of phonetic realization.

Condition (a) requires that only contrastive features may be specified in the lexicon; the method of determining what these are is discussed below. Condition (b) allows for the specification of features if required in the phonology, that is, if they are active. This condition represents a severe weakening of the Contrastivist Hypothesis, essentially abandoning it and allowing redundant features to be added freely in the phonology, with no restrictions. Given the success of the Contrastivist Hypothesis in the various types of cases discussed in this book, it is preferable to relax the strong Contrastivist Hypothesis in less drastic ways first, if necessary, for example as proposed in section 7.9.

As in MCS, Clements (2001) assumes that contrastive specifications at the lexical level are introduced by means of a feature hierarchy, called an *accessibility hierarchy*; this is slightly revised and called a *robustness scale* by Clements (2009). This scale is presented in (30).

(30) Robustness Scale for consonant features (Clements 2009: 46–7)
a. [±sonorant]
[labial]
[coronal]
[dorsal]
b. [±continuant]
[±posterior]
c. [±voiced]
[±nasal]
d. [glottal]
e. others

Unlike the assumption I have pursued, that the contrastive hierarchy may vary cross-linguistically, Clements assumes that the robustness scale is universal. The empirical import of this claim is substantially weakened by condition (b) in (29), however: since phonological activity is not evidence bearing on which features are contrastive, it is difficult to prove or disprove any proposed set of contrastive features for a particular language.

Clements argues that the robustness scale is supported in a general sense by the distribution of phonological inventories and the order of appearance of contrasts in the course of acquisition. Cross-linguistic evidence can certainly be useful in establishing general tendencies and perhaps even some default feature orderings. However, in the absence of any language-particular criteria for identifying contrastive features, one cannot tell if a particular language follows the general ordering or not. Second, the cross-linguistic evidence itself could be compromised, as D. C. Hall (2007: 187–8) points out. For example, according to the scale in (30), a contrast between /t/ and /s/ involves the feature [continuant] rather than [strident], because the former feature is higher in the scale. But then, as Hall observes, we cannot take this analysis as evidence that [continuant] is universally preferred over [strident].[13]

To sum up, Clements makes stronger claims about the universality of the feature hierarchy than the approach I have been taking here, but weaker claims about the connection between contrast and phonological activity. The two positions are closely linked: it is easier to argue for a universal feature hierarchy if one does not take into account phonological activity; conversely, it is easier to argue that active features are contrastive if one is not bound to a universal feature hierarchy. The two approaches can be brought closer together if we

13 Hall (2007: 188) also points to valid typological evidence that Clements (2009) draws on to support the robustness scale; see Hall (2007) for an extended discussion of feature economy from a contrastivist point of view.

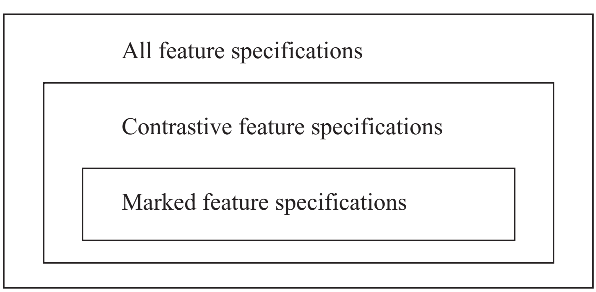
consider the robustness scale in (30) as a hypothesis about the default universal feature hierarchy, but one that is subject to variations under language-particular conditions.

Herd (2005) takes this approach, adopting the hierarchy proposed by Clements (2001) as an initial hypothesis for the contrastive hierarchies of the Polynesian languages in the study discussed above in section 7.7.2. Taking into account the loanword evidence, he proposes that variation may be introduced into the hierarchy under specific conditions. For instance, in the hierarchy of Clements (2001), [spread] is ranked higher than [continuant], with the consequence that the contrast between /h/ and a stop should involve [spread], not [continuant]. As we have seen, [spread] distinguishes /h/ from other phonemes in Hawaiian, but the loanword evidence suggests that [continuant] is the contrastive feature in NZ Māori. Herd proposes to keep the general ordering of [spread] > [continuant], but allows [spread] to be accessible only if it contrasts with another glottal segment.[14]

8.6 Calabrese (2005): Visibility Theory

Calabrese (2005: 68–9) proposes a theory of markedness and of contrast. As in MCS, he proposes that some rules target only *contrastive* feature specifications. Other rules target only *marked* contrastive specifications, a position that is also consistent with MCS practice. However, he also freely allows rules to target *all* feature specifications, a major weakening of the Contrastivist Hypothesis. The three classes of rules are diagrammed in (31). Calabrese calls this Visibility Theory, an approach taken also by Halle (1995), Halle, Vaux and Wolfe (2000), Vaux (2000) and Nevins (2004, 2005).

(31) Accessible features in Visibility Theory (Calabrese 2005: 68)



14 In the revised robustness scale (30), [continuant] has been moved above [spread glottis], so a slightly different modification of the scale would be required to promote the latter feature over [continuant] in the presence of a laryngeal contrast.

In addition to allowing phonological rules to freely access noncontrastive features, Calabrese's approach to markedness and contrast are different from what is assumed here. I will first briefly review his theory of markedness (section 8.6.1), and then his method for determining contrastive specifications (section 8.6.2). In section 8.6.3, I look at the way Calabrese's approach and MCS treat the difference between systematic and accidental gaps in an inventory.

8.6.1 Calabrese: markedness filters

Continuing in the *SPE*–Kean tradition, Calabrese (1988, 1994, 1995) proposes that markedness is encoded in a series of universal markedness statements, or filters, of the form in (32).

(32) Format for marking statements (Calabrese 1988, 1994)
*[αF, <u>βG</u>] / [___, X]
means that [βG] is not permitted in the context of a segment bearing [αF] and X, where F, G are features; α, β are + or −; and X is a set of feature specifications.

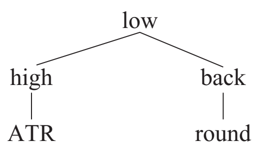
For example, *[+back, <u>–round</u>] / [___, –low] means that [–round] is not permitted on a segment specified [–low, +back]. These statements can be overridden in particular grammars, resulting in marked segments. The markedness statements (implicitly) encode a hierarchy of features against which contrast is assessed, in that feature G has a marked value in the context of F and X. Such a statement presupposes, in our terms, that F and the features that constitute X are ordered ahead of G; in this example, the ordering is [low] > [back] > [round]. Calabrese (2005: 126) proposes that the ordering of most of the features can be derived from the robustness scale proposed by Clements (2009). Calabrese observes that, to the extent that the features are ordered by independent principles, it is not necessary to designate feature G as the dependent feature by means of an underscore; the notation can thus be simplified to *[αF, βG]/. . .

Calabrese's approach to markedness statements not only is hierarchical, but also has a binary aspect, evident in the format adopted for marking statements. According to Calabrese (1994: 42), 'by restricting the focus of the statement to only two features, we capture the intuition that the basic relationships between feature specifications involve pairings of features, as observed by Stevens, Keyser & Kawasaki (1986). For example, there is a special acoustic relationship between [+back] and [–round] which is not shared by [–low]'. In this way, his version of markedness theory has connections with enhancement in MCS.

Calabrese (2005: 439) proposes the marking statements in (33) for vowels. These statements are consistent with a single partial ordering of features, as shown in (34).

(33) Marking statements for vowels (Calabrese 2005)
  a. *[–low, –high]
  b. *[–high, +ATR]
  c. *[+low, –back]
  d. *[–back, +round]
  e. *[+high, –ATR]
  f. *[+back, –round] / [___, –low]
  g. *[+low, +round]
  h. *[+low, +ATR]
  i. Prohibition: *[+high, +low]

(34) Partial feature ordering presupposed by (33)



The partial ordering of the markedness statements predicts that inventories can be expanded in certain ways but not in others. When all the statements in (33) are active, the only vowels permitted are /i, a, u/. These vowels thus have a Degree of Complexity of 0. Deactivating (33a) enables a vowel to be specified [–low, –high], adding the vowels /ɛ, ɔ/ to the basic set. These vowels have Degree of Complexity 1.

The marking statements predict that an [ATR] contrast must be added to inventories in the order shown in (35).

(35) Order of expansion of vowel inventories (Calabrese 2005)

a. Unmark (33a)

i u
ɛ ɔ
a

b. Unmark (33b)

i u
e o
ɛ ɔ
a

c. Unmark (33e)

i u
ɪ ʊ
e o
ɛ ɔ
a

d. Unmark (33h)

i u
ɪ ʊ
e o
ɛ ɔ
A a

An [ATR] contrast must be added in the mid vowels (35b), before it can be added in the high vowels (35c), and both of these must be present before an [ATR] contrast can be added in the low vowels (35e). This generalization may hold of Romance, and of typical African languages with ATR or RTR vowel contrasts, but it does not appear to be true of Manchu-Tungus languages, which typically have two height classes, each of which may have ATR or RTR contrasts. In these languages, there is no phonological class of [–low, –high] vowels, though there are vowels that are [+high, –ATR]; see Zhang (1996) for discussion of such systems.

More generally, a universal markedness scale of the type in (33) is unable to take into account variation in markedness, and does not relate markedness to contrast (see section 7.2.1). Rather than make markedness depend on contrast, as in MCS, Calabrese makes contrast depend on markedness, as I will show in the next section.

8.6.2 Calabrese's approach to contrast

In addition to markedness, Visibility Theory makes reference to contrastive feature specifications. Therefore, a procedure for selecting contrastive specifications is required. Calabrese (2005: 438) proposes a procedure that is, in essence, the pairwise method combined with a feature hierarchy: contrastive specifications are those that are not logically redundant; where two features make each other logically redundant, priority is given to the feature that is higher in the robustness scale. Calabrese's definition is given in (36).

(36) Formal definition of contrast (Calabrese 2005)

i. Given an active constraint *[αF, βG] in a system S, [–βG] is noncontrastive in the context of [αF], and [–αF] is noncontrastive in the context of [βG];

ii. if *[αF, βG] and [–αF, –βG] are active in S, [–βG] is noncontrastive in the context of [αF] and [βG] is noncontrastive in the context of [–αF] where [G] is marked with respect to [F].

Clause (i) defines as noncontrastive those feature specifications that are logically redundant in pairs of features. For example, given a constraint *[+low, +round], it is predictable that a low vowel must be [–round]; by the same token, any vowel that is [+round] must be [–low]. If we compare this with a contrastive hierarchy where [low] > [round], then the first conclusion is consistent with that derived by the contrastive hierarchy: [round] is not contrastive in the domain of [low]. The second conclusion allows the lower feature to make the higher one noncontrastive.

The danger of the procedure in clause (i) is that features can make each other noncontrastive, thereby failing to distinguish between phonemes. This is one of the characteristic shortcomings of identifying system redundancy with logical redundancy (see section 2.5). For such cases, clause (ii) stipulates that one chooses the feature that is higher in the ordering as contrastive.

Consider how this procedure will apply to the high-vowel inventory /i, y, u/. To simplify the discussion, we will assume that [+high] is contrastive and focus on [back], [round] and [ATR].

(37) Assigning contrastive specifications (Calabrese 2005)

a. Begin with full specifications

	i	y	u
high	+	+	+
back	–	–	+
round	–	+	+
ATR	+	+	+

b. Features that make each other logically redundant

[high] and [ATR] [high] > [ATR]

c. Reduced feature set

	i	y	u
high	+	+	+
back	–	–	+
round	–	+	+

d. Contrastive specifications

	i	y	u
high	+	+	+
back		–	+
round	–	+	

Like the pairwise method, the procedure begins with full specification (37a) (some features not shown).[15] The features [high] and [ATR] make each other logically redundant (37b), a problem for the Pairwise Algorithm. In Calabrese's procedure, such feature pairs fall under clause (ii) of (36), whereby the feature higher in the hierarchy, in this case [high], renders the lower one, [ATR], noncontrastive. This effectively reduces the set of potentially contrastive features to those in (37c). Now clause (i) of (36) enacts the Pairwise Algorithm, resulting in the familiar pairwise contrastive pattern of (38d), where /y/ is the middle phoneme that forms a minimal pair with each of the others.

This procedure for arriving at contrastive specifications is open to many of the objections raised in chapter 2 against versions of the pairwise method: it fails the Distinctness Condition, and it uses both feature ordering and pairwise comparison when the former alone is sufficient. The most important objection, however, is empirical: this theory does not allow for variation in the feature

15 Unlike the Pairwise Algorithm, Calabrese begins by designating all feature values as contrastive. Thus, his procedure technically assigns noncontrastive values.

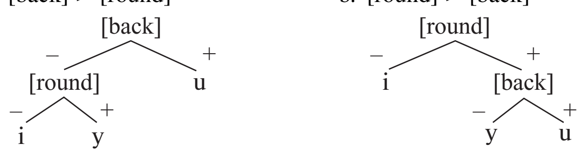
hierarchy, and so will arrive at the same contrastive relations for every similar-looking inventory. We have seen throughout that this is an undesirable result.

Curiously, Calabrese (2005: 430–3) levels the same objection against the theory in Dresher (2002), the same as the one being advocated here. Calabrese first mischaracterizes it as being identical to the approach taken in Clements (2001), which he considers to be a simpler reformulation of Dresher (2002). We have seen that a crucial difference between Clements's approach and the one taken here is that Clements assumes a universal feature ordering (the robustness scale) that does not take phonological activity into account, whereas I assume that the feature order is variable, and that the chief source of evidence for what the hierarchy is in a language derives from the pattern of phonological activity in that language.

Calabrese observes, correctly, that if we suppose that [back] is universally ordered above [round], then the SDA will assign the contrastive feature [–back] to both *i* and *y* in (37), as we can see in (38a). He observes, also correctly, that this is the wrong result for languages like Finnish and Hungarian which have palatal (front/back) harmony, in which the front unrounded vowel /i/ is neutral. The correct analysis is one in which the front unrounded vowels are not contrastively specified for the harmonizing feature, which requires the ordering [round] > [back] (38b). This may be a problem for Clements, but since I do not assume a universal ordering, the ordering [round] > [back] is possible.[16]

(38) Two contrastive hierarchies for /i, y, u/

a. [back] > [round] b. [round] > [back]



Calabrese's argument actually cuts the other way. If Trubetzkoy is to be believed about Polabian (section 3.3.3), in that language the three vowels /i, y, u/ have different contrastive relations from the ones they have in Finnish and Hungarian; in Polabian, /i/ and /y/ are both contrastively [–back] against [+back] /u/. That is, whereas in Finnish and Hungarian /y/ and /u/ are partners with respect to [back] while /i/ is neutral, in Polabian /i/ and /y/ are partners with respect to [round] while /u/ is neutral. These different relations correspond to different contrastive hierarchies. It is Calabrese's theory that cannot capture this variability.

16 See D'Arcy (2003b) for a contrastive analysis using other features.

Similar arguments can be made for other inventories where we have seen variation in the system of contrasts. To take just one more example, consider again the four-vowel /i, a, o, u/ system of Yowlumne, discussed above in several different contexts. We have seen that the phonological patterning of this language is best accounted for by assuming that [labial] (or [round]) is contrastive and [coronal] (or [front]/[back]) is not, because the language is characterized by height-bounded rounding harmony. Yowlumne vowels can be assigned the specifications in (39).[17]

(39) Yowlumne vowels: full specifications

	i	u	a	ɔ
low	−	−	+	+
high	+	+	−	−
back	−	+	+	+
round	−	+	−	+
ATR	+	+	−	−

Following Calabrese's procedure for marking contrastive features, we observe that [low] and [high] make each other mutually logically redundant, so [high] is designated noncontrastive. [ATR] is also fully predictable given [low], and vice versa; since [ATR] ranks lower, it is also made noncontrastive. Since /a/ and /ɔ/ are a minimal pair based on [round], those feature values must be marked as contrastive. But even though /i/ and /u/ make a parallel pair distinguished by [round], [round] in this case is trumped by [back]: in the non-low vowels they predict each other, but [back] ranks higher in the universal scale. Therefore, we arrive at the contrastive representations in (40). This set of contrastive representations does not at all reflect the patterning in the language, and contradicts the intuition, shared by every analysis of this language, that Yowlumne vowels are divided on the basis of rounding, not backness.[18]

(40) Yowlumne contrastive specifications by (36)

	i	u	a	ɔ
low	−	−	+	+
back	−	+		
round			−	+

17 The vowel usually transcribed as *o* is here given as a low vowel ɔ, following Newman (1944: 19); see also D. C. Hall (2007: 115) and section 7.9.

18 Calabrese could still capture rounding harmony, but he would have to characterize it as targeting all features, not just contrastive features.

8.6.3 Systematic and accidental gaps

In Calabrese's approach, designation of a feature as noncontrastive depends on the markedness statements. Calabrese (2005: 443–9) argues that this dependence allows for an account of the famous problem in Russian voicing assimilation, whereby the obstruents /x, ts, tʃ/ pattern with the other contrastively voiceless obstruents, though they have no voiced (minimal) counterparts. Calabrese (2005: 445–6) observes that this would be a problem if contrast were based only on 'the structural arrangement of the system' (that is, in terms of minimal pairs), because these segments should be designated as noncontrastive for voicing, though they pattern as though they are contrastively [–voiced].

However, Calabrese argues that his procedure for assigning noncontrastive features will not designate these phonemes as noncontrastive. This is because the procedure in (36) crucially refers to marking statements. There is a universal marking statement that says sonorants should not be voiced, and the procedure designates them as noncontrastive for [voiced]. But there is no universal marking statement making fricatives and affricates voiceless. Thus, the absence of their voiced counterparts is an 'accidental' gap, not a systematic one. Calabrese proposes that the accidental absence of segments from a phonological inventory does not determine the underlying specifications of the segments of that inventory. Therefore, the unpaired fricatives and affricates can be treated as contrastively voiceless.

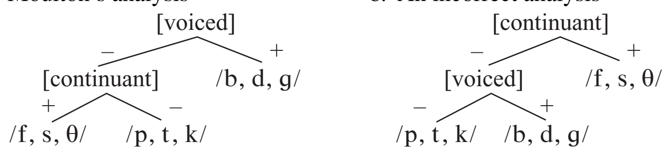
While it gives the correct result in this case, this procedure seems to contradict the intuitive notion of what it means to be contrastive in a system. The SDA is a much simpler way of arriving at the same result while limiting the evaluation of contrast to contrasts actually observed within the system in question. In the contrastive hierarchy proposed in (14), the feature [sonorant] is high in the hierarchy and divides the sonorants from the obstruents. Since the sonorants are not contrastive for [voiced], they do not participate in voicing assimilation. All non-sonorant consonants receive a contrastive value for [voiced]. All that is required for the unpaired obstruents to be contrastively voiceless is for the voicing feature to be sufficiently high in the order.

Calabrese's distinction between systematic and accidental gaps in an inventory can be modelled in terms of the contrastive hierarchy. The difference between gaps that are perceived as systematic as opposed to those that are perceived as accidental is a function of where in the hierarchy the gap occurs. Thus, the lack of voiced sonorants looks systematic when the [sonorant] feature is ordered above [voiced], for then no sonorant consonant has a specification for [voiced]. The same gap would look more accidental if [voiced] were ordered ahead of [sonorant]. This conclusion is presented more formally in (41).

(41) ‘Gaps’ in the light of the Successive Division Algorithm
If τ is a segment with feature [αG], and if Φ is the minimal set of ordered features that distinguishes τ from all the segments that have [–αG], then the absence of a segment σ that differs minimally from τ by the feature G is ‘accidental’ iff G $>$ at least one of Φ (τ is contrastively [αG]), and is ‘systematic’ if all of $\Phi > G$ (τ is redundantly [αG]).

An interesting example is provided by Moulton (2003), in an analysis of Old English fricatives. The Old English fricatives /f, s, θ/ do not contrast lexically with respect to voicing, but, like the Russian unpaired fricatives and affricates, they enter into voicing alternations, and have voiced allophones. Moulton proposes that they are contrastively [–voiced], a specification that follows from the contrastive hierarchy in (42a).

(42) Old English fricatives (Moulton 2003: 161)
a. Moulton’s analysis b. An incorrect analysis



Divided up as in (42a), the lack of voiced continuants could be called ‘accidental’: the feature [continuant] occurs in the domain of [–voiced], but there happen not to be any continuants under [+voiced]. Another way of dividing up the inventory (42b) makes the gap appear to be more ‘systematic’: voicing contrasts are permitted in the domain of [–continuant], but not in the domain of [+continuant]. On this analysis the fricatives have no contrastive specification for [voiced], a result that is contradicted by their behaviour with respect to voicing alternations. It follows that whether a gap is considered systematic or accidental is to some extent a matter of point of view, influenced by the scope relations assigned to the relevant features.

8.7 ‘Minimal contrast’

Though Calabrese’s procedure for assigning (non)contrastive specifications is different in a number of ways from the Pairwise Algorithm, it shares the assumption that contrast is determined (only in part, in Calabrese’s case) by considerations of logical redundancy and minimal pairs. I have argued throughout this book that contrastive specifications must be derived hierarchically, and not by inspection of minimal pairs. This was the main theme of chapter 2, and I have shown in subsequent chapters that contrast based on minimal pairs has been a recurring approach in the history of phonology. It is fitting, then, to conclude our survey of contemporary approaches to contrast by emphasizing that the minimal pairs approach is not just a historical artefact, but remains a popular method for determining contrastive specifications, even uniting phonologists who do not otherwise share the same frameworks. We have seen an appeal to minimal differences in the work of Padgett and Calabrese discussed above. Campos Astorkiza (2007) proposes a special role for *minimal contrast*, two sounds that differ in just one property. Nevins defines contrast as in (43).

(43) Contrastiveness (Nevins 2004: 142)
A segment S with specification αF is *contrastive* for F if there is another segment S′ in the inventory that is featurally identical to S, except that it is –αF.

I have raised various logical and empirical arguments against basing contrast on minimal pairs, and I will not rehearse them all here. It is worth emphasizing, however, that many supposed minimal pairs appear to be so only because the analyst is willing to overlook phonetic differences that are known not to be relevant to the phonological patterning of the sounds. That is, many analyses that appear to be based on minimal pairs actually tacitly impose a partial feature ordering on the data, silently removing irrelevant phonetic differences.

The surface phonetics is a particularly inhospitable place to look for minimal pairs; recall D. C. Hall's observation (2007: 165), cited in section 2.5.4, that phonetic enhancement and other modifications tend to obscure the phonologically contrastive dimensions, augmenting them with further phonetic distinctions. A typical example is provided by Campos Astorkiza, who diagrams the Lithuanian vowel inventory as in (44).

(44) Lithuanian vowel inventory (Campos Astorkiza 2007: 32)

a. Phonemic inventory / b. Phonetic realization

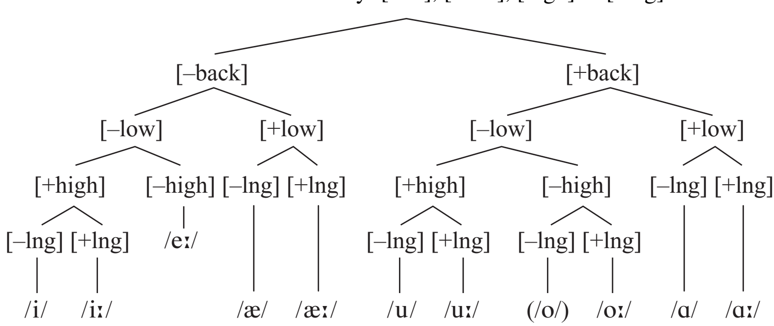
Front		Back			Front		Back	
/i/	/iː/	/u/	/uː/	High	[ɪ]	[iː]	[ʊ]	[uː]
	/eː/	(/o/)	/oː/	Mid		[eː]	([ɔ])	[oː]
/æ/	/æː/	/ɑ/	/ɑː/	Low	[ɛ]	[æː]	[a]	[ɑː]

Campos Astorkiza (2007) presents a DT analysis of contrast in Lithuanian vowels along the lines of Ní Chiosáin and Padgett (2001) and Padgett (2003a, b). In her phonemic analysis, represented by (44a), the vowels of Lithuanian are distinguished along three contrastive dimensions: height, backness and length. Campos Astorkiza wishes to designate as minimally different any two vowels that differ only along one of these dimensions; for example, /i/ and /iː/ differ only in duration, /i/ and /u/ differ only in backness, and so on.

While this may be the correct analysis at the phonemic, underlying, level of representation, a glance at (44b) makes clear that the phonetic reality is different. From a purely phonetic point of view, Lithuanian long vowels do not differ from their short counterparts only with respect to duration. Each pair displays a further phonetic difference that is potentially contrastive in some other language: compared to their short vowel counterparts, the long high vowels are either higher, or tenser, or have more advanced tongue root, or any two or three of these. The front short low vowel is higher than its long counterpart, and the short back (or central) low vowel is less retracted than its long counterpart.[19] Further, the non-low vowels contrast in roundness as well as in backness. Nor is it clear at the phonetic level that the number of front/back contrasts should be limited to two categories: representing the vowel [a] as [central] rather than [back] appears to be closer to phonetic reality.

Thus, the phonemic analysis in (44a) cannot be read off the surface phonetics, nor is the surface the appropriate level at which to seek these contrastive dimensions. Rather, the analysis in (44a) reflects a contrastive hierarchy in which [low], [back] and [high] are ordered ahead of [long], which is ordered ahead of all other features, which are noncontrastive, as shown in (45).[20]

(45) Lithuanian contrastive hierarchy: [low], [back], [high] > [long]



19 It is quite common for vowel length (quantity) contrasts to become obscured and ultimately reanalysed as quality contrasts; see Kabak (2004) on Turkic and Korean, and Rohany Rahbar (2008) on Persian.

20 Since the unpaired vowel /e:/ functions as a long vowel, it is necessary to stipulate that Lithuanian vowels with no contrastive value for length are [+long] by default.

The Russian consonant system provides a final illustration of the limitations of basing contrastive specification on minimal contrasts. With respect to the dichotomy between palatalized and non-palatalized plain (or velarized) consonants, the traditional terminology of 'paired' and 'unpaired' consonants might at first lead one to think that minimal contrast successfully predicts which consonants are contrastive for the feature [back]: 'paired' consonants are contrastive for [back], since this is the main feature that distinguishes a paired palatalized consonant from its non-palatalized counterpart. By the same token the 'unpaired' velar consonants do not have a contrastive [back] feature, as predicted.[21]

The velars, however, are not the only 'unpaired' consonants with respect to palatalization. There are unpaired coronal consonants /j, ts, tʃʲ, ʒ, ʃ/, and they do not act like the velars: /ts, ʒ, ʃ/ pattern as hard consonants, specified as [+back], and /j, tʃʲ/ pattern as soft consonants, specified [–back]. Minimal contrast would lead us to expect these consonants to be noncontrastive for the feature [back], because there are no corresponding consonants that differ only with respect to this feature: for example, there is no underlying phoneme /ʒʲ/ to provide a minimal contrast with /ʒ/, or /tʃ/ to serve as the hard counterpart of /tʃʲ/. Thus, minimal contrast cannot account for why the unpaired coronals do not behave like the unpaired velars.

In a hierarchical conception of contrast, however, such differences between velars and coronals can be expected, if not predicted. Russian coronal consonants participate in more contrasts than the velars, so there are more chances that the feature [back] will be above other features in the coronals than in the velars. In the contrastive hierarchy in (14), illustrated by the trees in (15), the feature hierarchy for the velars runs out before [back] can play a contrastive role. Even though the same hierarchy obtains among the coronals, there are more coronals than velars, and [back] can play a contrastive role in every coronal consonant. In the case of /dʲ, d/ in (15), the contrast in [back] is the lowest feature in the contrastive hierarchy. In the other cases, it is not, and the notion of 'partnership' plays out at a higher level. Thus, given the hierarchy in (14), the palatalized partner of /ʒ/ is /zʲ/, a partner it shares with /z/.

In a hierarchical conception of contrast assignment, not all features can have lowest scope, so we expect that minimal pairs may represent the limiting cases, where contrast is required, but not the only cases of contrastive features. A famous example is the feature [voiced] in Russian, which has played an

21 This is the case for earlier versions of Russian, such as the East Slavic situation discussed in section 8.3. Modern Russian presents a more complex situation.

important role in the development of phonological theory, and which has been discussed here in a number of contexts (see especially section 4.6). With respect to this feature, too, there are ‘paired’ and ‘unpaired’ consonants, but here *all* the unpaired consonants have a specification for [voiced], against the prediction of minimal contrast: this is true of coronal /ts, tʃʲ/, velar /x/ and labial /f/ (assuming /v/ does not contrast minimally with it).

Such consonants could pattern as noncontrastive for [voiced] if [voiced] were further down in the Russian contrastive hierarchy. But we have seen that [back] must be quite low in the hierarchy, and there is reason to suppose that [continuant] is also lower than [voiced] in many Slavic languages (Radišić 2007). Thus, it is a prediction of the hierarchical approach to contrast that minimal pairs could not be an adequate basis for predicting contrastive specifications for *all* of these features.

The results are as we expect. One of these features, [voiced], is relatively high in the hierarchy, and takes as its domain all the obstruent consonants, whether they occur in minimal pairs or not. Another feature, [back], is relatively low in the hierarchy, and its contrastive behaviour begins to approach what would be predicted by minimal contrast. Even here, though, minimal contrast is not observed everywhere, but is violated in the coronals, which act as contrastive for [back] in the absence of minimal contrast.

9 *Conclusion*

This book has been about the contrastive hierarchy in phonology. In one sense, 'phonology' refers to phonological theory as it has been understood and practised from the 1920s until today. I have tried to trace the history of the notion of a contrastive feature hierarchy, and its various guises as it has appeared and disappeared in different versions of phonological theory. In another sense, 'phonology' can be understood as a component of the language faculty whose properties phonologists are trying to discover. In this sense, I have looked for evidence of the role that the contrastive hierarchy plays in contributing to an understanding of this cognitive domain.

In both aspects of this study the fate of the contrastive hierarchy has been closely connected to another central concept, the Contrastivist Hypothesis. The Contrastivist Hypothesis states that only contrastive features are active in the phonology. To properly test whether this hypothesis is correct, we need to have reliable ways of identifying the 'contrastive features'. We also need to know what it means for a feature to be 'active', and what place 'in the phonology' refers to; but my main concern here has been contrast. I have argued that the contrastive hierarchy is crucial to the Contrastivist Hypothesis, because it provides the proper way to identify contrastive features. Conversely, the Contrastivist Hypothesis is crucial to the contrastive hierarchy, because the former supplies the main empirical reasons for being interested in the latter, at least as far as phonology is concerned.

These two concepts have not always been linked in phonological theory. On the contrary, they have more usually been dissociated, with sometimes one, sometimes the other coming to the fore. The story that emerges from the preceding chapters can be summed up as follows. Something like the Contrastivist Hypothesis was adopted by the principal phonological theorists in the first half of the twentieth century, but the notion of a contrastive hierarchy remained inchoate. Though it played an implicit role in the most interesting structuralist contrastivist analyses, it remained unformulated, and competed with a pairwise approach that saw contrast in terms of logical redundancy and minimal pairs.

The 1950s were a turning point. For the first time the contrastive hierarchy was explicitly proposed as governing feature specifications. At the same time, however, the Contrastivist Hypothesis began to be deemphasized. The contrastive hierarchy came to be viewed mainly as a way of minimizing redundancies in phonological specifications, not as a way of capturing phonological generalizations. The Contrastivist Hypothesis was abandoned in early generative phonology, and when redundant feature values were shown to cause technical and empirical problems, the contrastive hierarchy also disappeared from phonological theory.

But not entirely. In chapters 5 and 6 I traced the influence of these concepts in generative phonology and in Optimality Theory. As in earlier incarnations of phonological theory, they went their separate ways. When something approximating the Contrastivist Hypothesis was proposed, the minimal pairs method was adopted for identifying contrastive features, with unfortunate results. When feature hierarchies were proposed, they did not have a clear connection to contrastive specification.

In the next two chapters the focus was on empirical evidence for the Contrastivist Hypothesis and the contrastive hierarchy in the context of contemporary conceptions of phonological theory. I outlined the main tenets of the theory of Modified Contrastive Specification (MCS), in which the Contrastivist Hypothesis is united with the contrastive hierarchy, and reviewed a number of case studies that adopt this framework. I also looked at some other contemporary approaches to contrast, arguing that the MCS approach accounts better for the data.

This study has only touched on a few aspects of a multifaceted topic. I hope to have shown that the contrastive hierarchy merits its former designation as a 'pivotal principle' of linguistic structure, and that the Contrastivist Hypothesis is a fruitful approach to phonological theory, capable of providing illuminating analyses of phonological patterns, and posing interesting questions that remain to be answered.

References

Abaglo, Poovi, and Diana Archangeli (1989). Language-particular underspecification: Gengbe /e/ and Yoruba /i/. *Linguistic Inquiry* 20: 457–80.

Anderson, John M. (2005). Old English *i*-umlaut (for the umpteenth time). *English Language and Linguistics* 9: 195–227.

Anderson, John M., and Colin J. Ewen (1987). *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. (1985). *Phonology in the twentieth century: theories of rules and theories of representations.* Chicago: University of Chicago Press.

Aoki, Haruo (1966). Nez Perce vowel harmony and Proto-Sahaptian vowels. *Language* 42: 759–67.

(1970). *Nez Perce grammar* (University of California Publications in Linguistics 62). Berkeley: University of California Press.

Archangeli, Diana (1984). Underspecification in Yawelmani phonology and morphology. Doctoral dissertation, MIT, Cambridge, Mass. Revised version published 1988 by Garland Press, New York.

(1988). Aspects of underspecification theory. *Phonology* 5: 183–207.

Archangeli, Diana, and Douglas Pulleyblank (1989). Yoruba vowel harmony. *Linguistic Inquiry* 20: 173–217.

(1994). *Grounded phonology.* Cambridge, Mass.: MIT Press.

(2007). Harmony. In de Lacy (2007), 353–78.

Aslin, Richard N., David B. Pisoni, Beth L. Hennessy and Alan J. Perey (1981). Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. *Child Development* 52: 1135–45.

Avery, Peter (1996). The representation of voicing contrasts. Doctoral dissertation, University of Toronto.

Avery, Peter, B. Elan Dresher and Keren Rice (eds.) (2008). *Contrast in phonology: theory, perception, acquisition*. Berlin: Mouton de Gruyter.

Avery, Peter, and William J. Idsardi (2001). Laryngeal dimensions, completion and enhancement. In T. A. Hall (2001), 41–70.

Avery, Peter, and Keren Rice (1989). Segment structure and coronal underspecification. *Phonology* 6: 179–200.

(2004). The representational residue: the role of contrast in phonology. Paper presented at the 12th Manchester Phonology Meeting, 20–22 May 2004, University of Manchester. Ms., Department of Languages, Literatures, and Linguistics, York University, Toronto, and Department of Linguistics, University of Toronto.

Baković, Eric (2000). Harmony, dominance and control. Doctoral dissertation, Rutgers University, New Brunswick, NJ.

Balcaen, M. Jean (1998). Neutralization and enhancement: vowel harmony in Aka. Ms., Department of Linguistics, University of Toronto.

Beckman, Jill N. (1997). Positional faithfulness. Doctoral dissertation, University of Massachusetts, Amherst.

Béjar, Susana (1998). Segmental inventory and the representation of voicing contrasts. Presented at the Montreal–Ottawa–Toronto Phonology Workshop, University of Ottawa, February 1998. Ms., Department of Linguistics, University of Toronto.

Bermúdez-Otero, Ricardo (2003). The acquisition of phonological opacity. In Jennifer Spenader, Anders Eriksson and Östen Dahl (eds.), *Variation within Optimality Theory: proceedings of the Stockholm Workshop on 'Variation within Optimality Theory'*, 25–36. Stockholm: Department of Linguistics, Stockholm University.

Best, Catherine T. (1995). A direct realist perspective on cross-language speech perception. In Winifred Strange (ed.), *Speech perception and linguistic experience: issues in cross-language research*, 171–204. Timonium, Md.: York Press.

Bloomfield, Leonard (1933). *Language*. New York: Holt.

Boersma, Paul (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives* (LOT International Series 11). The Hague: Holland Academic Graphics.

Bolognesi, Roberto (1998). *The phonology of Campidanian Sardinian: a unitary account of a self-organizing structure* (HIL Dissertations 38). The Hague: Holland Academic Graphics.

Brannen, Kathleen (2002). The role of perception in differential substitution. *Canadian Journal of Linguistics / Revue canadienne de linguistique* 47: 1–46.

Broe, Michael B. (1993). Specification theory: the treatment of redundancy in generative phonology. Doctoral dissertation, University of Edinburgh.

Broe, Michael B., and Janet B. Pierrehumbert (2000). Introduction. In Michael Broe and Janet B. Pierrehumbert (eds.), *Papers in laboratory phonology V: Acquisition and the lexicon*, 1–7. Cambridge: Cambridge University Press.

Broecke, M. P. R. van den (1976). *Hierarchies and rank orders in distinctive features*. Assen: Van Gorcum.

Brown, Cynthia A. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In John Archibald (ed.), *Second language acquisition and linguistic theory*, 4–65. Oxford: Blackwell.

Cairns, Charles E. (1969). Markedness, neutralization and universal redundancy rules. *Language* 45: 863–85.

(1988). Phonotactics, markedness and lexical representation. *Phonology* 5, 209–36.

Calabrese, Andrea (1988). Towards a theory of phonological alphabets. Doctoral dissertation, MIT, Cambridge, Mass.

(1994). A constraint-based theory of phonological inventories. In Wolfgang U. Dressler, Martin Prinzhorn and John R. Rennison (eds.), *Phonologica 1992: proceedings of the 7th International Phonology Meeting*, 35–54. Turin: Rosenberg & Sellier.

(1995). A constraint-based theory of phonological markedness and simplification procedures. *Linguistic Inquiry* 26: 373–463.

(2005). *Markedness and economy in a derivational model of phonology*. Berlin: Mouton de Gruyter.

Campos Astorkiza, Judit Rebeka (2007). Minimal contrast and the phonology – phonetics interaction. Doctoral dissertation, University of Southern California, Los Angeles.

Carr, Philip, Jacques Durand and Colin J. Ewen (eds.) (2005). *Headhood, elements, specification and contrastivity: phonological papers in honour of John Anderson*. Amsterdam: John Benjamins.

Casali, Roderic F. (2003). [ATR] value asymmetries and underlying vowel inventory structure in Niger-Congo and Nilo-Saharan. *Linguistic Typology* 7: 307–82.

Causley, Trisha (1999). Complexity and markedness in Optimality Theory. Doctoral dissertation, University of Toronto.

Chao, Yuen Ren (1954). Review of Jakobson, Fant, Halle, *Preliminaries* ... *Romance Philology* 8: 40–6.

Chen, Moshe (1972). Mivta'a shel ha'ivrit hayisra'elit [The pronunciation of Israeli Hebrew]. *Leshonenu* 36: 212–300.

Cherry, E. Colin, Morris Halle and Roman Jakobson (1953). Toward the logical description of languages in their phonemic aspect. *Language* 29: 34–46. Reprinted in Jakobson (1962c), 449–63. Also in Makkai (1972), 323–32.

Chomsky, Noam (1951). Morphophonemics of Modern Hebrew. Master's thesis, University of Pennsylvania.

(1957). *Syntactic structures*. The Hague: Mouton.

(1964). Current issues in linguistic theory. In Jerry A. Fodor and Jerrold J. Katz (eds.), *The structure of language,* 50–118. Englewood Cliffs, NJ: Prentice-Hall.

Chomsky, Noam, and Morris Halle (1956). On the logic of phonemic description. Paper presented at the MIT Conference on Speech Communication, June 16 1956.

(1968) [*SPE*]. *The sound pattern of English.* New York: Harper & Row.

Christdas, Prath (1988). The phonology and morphology of Tamil. Doctoral dissertation, Cornell University, Ithaca, NY.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook* 2: 225–52.

(1987). Toward a substantive theory of feature specification. In James Blevins and Juli Carter (eds.), *Proceedings of NELS 18*, 79–93. Amherst, Mass.: GLSA.

(2001). Representational economy in constraint-based phonology. In T. A. Hall (2001), 71–146.

(2003). Feature economy in sound systems. *Phonology* 20: 287–333.

(2009). The role of features in phonological inventories. In Eric Raimy and Charles Cairns (eds.), *Contemporary views on architecture and representations in phonological theory*, 19–68. Cambridge, Mass.: MIT Press.

Clements, G. N., and Elizabeth V. Hume (1995). The internal organization of speech sounds. In Goldsmith (1995), 245–306.

Cohn, Abigail (1993). Nasalization in English: phonology or phonetics? *Phonology* 10: 43–81.

Compton, Richard (2008). Contrast in Inuit consonant inventories. In Susie Jones (ed.), *Proceedings of the 2008 annual conference of the Canadian Linguistic Association*, www.chass.utoronto.ca/~cla-acl/actes2008/CLA2008_Compton.pdf.

Compton, Richard, and B. Elan Dresher (2008). Palatalization and 'strong' /i/ across Inuit dialects. In Susie Jones (ed.), *Proceedings of the 2008 annual conference of the Canadian Linguistic Association*, www.chass.utoronto.ca/∼cla-acl/actes2008/CLA2008_Compton_Dresher.pdf.

Crosswhite, Katherine M. (2001). *Vowel reduction in Optimality Theory*. New York: Routledge.

D'Arcy, Alex (2003a). Yowlumne reexamined: a challenge for Contrastive Specification. In Daniel Currie Hall (ed.), *Toronto Working Papers in Linguistics (Special Issue on Contrast in Phonology)* 20: 21–46. Toronto: Department of Linguistics, University of Toronto.

(2003b). Unconditional neutrality: vowel harmony in a two-place model. Ms., Department of Linguistics, University of Toronto.

de Lacy, Paul (2006). *Markedness, reduction and preservation in phonology*. Cambridge: Cambridge University Press.

(ed.) (2007). *The Cambridge handbook of phonology*. Cambridge: Cambridge University Press.

Dinnsen, Daniel A. (1992). Variation in developing and fully developed phonetic inventories. In Charles A. Ferguson, Lisa Menn and Carol Stoel-Gammon (eds.), *Phonological development: models, research, implications*, 191–210. Timonium, Md.: York Press.

(1996). Context-sensitive underspecification and the acquisition of phonetic contrasts. *Journal of Child Language* 23: 31–55.

Dinnsen, Daniel A., Steven B. Chin, Mary Elbert and Thomas W. Powell (1990). Some constraints on functionally disordered phonologies: phonetic inventories and phonotactics. *Journal of Speech and Hearing Research* 33: 28–37.

Dorais, Louis-Jacques (2003). *Inuit uqausiqatigiit: Inuit languages and dialects (second, revised edition)*. Iqaluit: Nunavut Arctic College.

Dresher, B. Elan (1983). Postlexical phonology in Tiberian Hebrew. In Michael Barlow, Daniel P. Flickinger and Michael T. Wescoat (eds.), *Proceedings of the West Coast Conference on Formal Linguistics,* Vol. II, 67–78. Stanford, Calif.: Stanford Linguistics Association.

(1985). Constraints on empty positions in tiered phonology. In Dana Mullen and Jan van Voorst (eds.), *Cahiers linguistiques d'Ottawa* 14: 1–51. Ottawa: Department of Linguistics, University of Ottawa.

(1993). Review of Richard M. Hogg, *A grammar of Old English. Volume 1: Phonology*. *Phonology* 10: 146–57.

(1996). Learnability and phonological theory. In Jacques Durand and Bernard Laks (eds.), *Current trends in phonology: models and methods*, Vol. I, 245–66. University of Salford: European Studies Research Institute (ESRI).

(1998a). Child phonology, learnability, and phonological theory. In Tej Bhatia and William C. Ritchie (eds.), *Handbook of language acquisition*, 299–346. New York: Academic Press.

(1998b). On contrast and redundancy. Paper presented at the annual conference of the Canadian Linguistic Association, May, Ottawa. Ms., Department of Linguistics, University of Toronto.

(1999). Charting the learning path: cues to parameter setting. *Linguistic Inquiry* 30: 27–67.

(2002). Determining contrastiveness: a missing chapter in the history of phonology. In Sophie Burelle and Stanca Somesfalean (eds.), *Proceedings of the 2002 annual conference of the Canadian Linguistic Association*, 82–93. Montreal: Département de linguistique et de didactique des langues, Université du Québec à Montréal.

(2003). Contrast and asymmetries in inventories. In Anna-Maria di Sciullo (ed.), *Asymmetry in grammar II: morphology, phonology, acquisition,* 239–257. Amsterdam: John Benjamins.

(2005). Chomsky and Halle's revolution in phonology. In James McGilvray (ed.), *The Cambridge companion to Chomsky*, 102–22. Cambridge: Cambridge University Press.

(2008). The contrastive hierarchy in phonology. In Avery, Dresher and Rice (2008), 11–33.

Dresher, B. Elan, and Harry van der Hulst (1998). Head-dependent asymmetries in phonology: complexity and visibility. *Phonology* 15: 317–52.

Dresher, B. Elan, Glyne Piggott and Keren Rice (1994). Contrast in phonology: overview. In Carrie Dyck (ed.), *Toronto Working Papers in Linguistics* 13,1: iii–xvii. Toronto: Department of Linguistics, University of Toronto.

Dresher, B. Elan, and Keren Rice (1993). Complexity in phonological representations. In Carrie Dyck (ed.), *Toronto Working Papers In Linguistics* 12: i–vi. Toronto: Department of Linguistics, University of Toronto.

(2007). *Markedness and the contrastive hierarchy in phonology*. www.chass.utoronto.ca/∼contrast/.

Dresher, B. Elan, and Xi Zhang (2004). Contrast in Manchu vowel systems. In Carsten Naher (ed.), *Proceedings of the First International Conference on Manchu-Tungus Studies (Bonn, August 28 – September 1, 2000) II: Trends in Tungusic and Siberian Linguistics*, 23–43. Wiesbaden: Harrassowitz.

(2005). Contrast and phonological activity in Manchu vowel systems. *Canadian Journal of Linguistics / Revue canadienne de linguistique* 50: 45–82.

Dupoux, Emmanuel, Christophe Pallier, Kazuhiko Kakehi and Jacques Mehler (1999). Epenthetic vowels in Japanese: a perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25: 1568–78.

Dyck, Carrie (1995). Constraining the phonology–phonetics interface, with exemplification from Spanish and Italian dialects. Doctoral dissertation, University of Toronto.

Efere, Emmanuel Efereala (2001). The pitch system of the Bumo dialect of Izon. *University of British Columbia Working Papers in Linguistics (Current research on African languages and linguistics)* 4: 115–259.

Eimas, Peter D., Einar R. Siqueland, Peter Jusczyk and James Vigorito (1971). Speech perception in infants. *Science* 171: 303–6.

Farina, Donna Marie (1991). Palatalization and jers in Modern Russian phonology: an underspecification approach. Doctoral dissertation, University of Illinois at Urbana-Champaign.

Fikkert, Paula (1994). *On the acquisition of prosodic structure* (HIL Dissertations 6). Dordrecht: ICG Printing.

(2007). Developing representations and the emergence of phonology: evidence from perception and production. Paper presented at PaPI'07, Phonetics and Phonology in Iberia, Braga, Portugal, June 2007. [To appear in *Papers in Laboratory Phonology 10*. Berlin: Mouton.]

Fikkert, Paula, and Clara C. Levelt (2008). How does Place fall into place? The lexicon and emergent constraints in children's developing phonological grammars. In Avery, Dresher and Rice (2008), 231–68.

Flege, James Emil (1995). Second-language speech learning: theory, findings, and problems. In Winifred Strange (ed.), *Speech perception and linguistic experience: issues in cross-linguistic research*, 233–77. Timonium, Md.: York Press.

Flemming, Edward (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18: 7–44.

(2002). *Auditory representations in phonology*. London: Routledge. [Revised version of University of California, Los Angeles doctoral dissertation, 1995.]

(2004). Contrast and perceptual distinctiveness. In Hayes, Kirchner and Steriade (2004), 232–76.

Fortescue, Michael, Steven A. Jacobson and Lawrence D. Kaplan (1994). *Comparative Eskimo dictionary with Aleut cognates*. Fairbanks: Alaska Native Language Center.

Frigeni, Chiara (2003). Metaphony in Campidanian Sardinian: a domain-based analysis. In Daniel Currie Hall (ed.), *Toronto Working Papers in Linguistics* (*Special Issue on Contrast in Phonology)* 20: 63–91. Toronto: Department of Linguistics, University of Toronto.

(2009). Sonorant relationships in two varieties of Sardinian. Doctoral dissertation, University of Toronto.

Frisch, Stefan (1996). Similarity and frequency in phonology. Doctoral dissertation, Northwestern University, Evanston, Ill.

Frisch, Stefan, Janet B. Pierrehumbert and Michael B. Broe (2004). Similarity avoidance and the OCP. *Natural Language and Linguistic Theory* 22: 179–228.

Ghini, Mirco (2001a). *Asymmetries in the phonology of Miogliola*. Berlin: Mouton de Gruyter.

(2001b). Place of articulation first. In T. A. Hall (2001), 147–76.

Goldsmith, John A. (1976). Autosegmental phonology. Doctoral dissertation, MIT, Cambridge, Mass. Published 1979 by Garland Press, New York.

(ed.) (1995). *Handbook of phonological theory*. Oxford: Blackwell.

Guion, Susan G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica* 55: 18–52.

Hale, Mark, and Charles Reiss (1998). Formal and empirical arguments concerning phonological acquisition. *Linguistic Inquiry* 29: 656–83.

Hall, Beatrice L., and R. M. R. Hall (1980). Nez Perce vowel harmony: an Africanist explanation and some theoretical consequences. In Robert M. Vago (ed.), *Issues in vowel harmony,* 201–36. Amsterdam: John Benjamins.

Hall, Daniel Currie (1998). Contrastive specification for voicing in Czech. Paper presented at the MOT Phonology Workshop, University of Ottawa. Ms., Department of Linguistics, University of Toronto.

(1999). On the geometry of auditory space. In Carolyn Smallwood and Susana Béjar (eds.), *Toronto Working Papers in Linguistics* 17: 134–57. Toronto: Department of Linguistics, University of Toronto.

(2004). Information theory and the wrong approach to phonological feature specification. Paper presented at the Third North American Phonology Conference (NAPhC 3), Concordia University, Montreal, May 2004. Ms., Department of Linguistics, University of Toronto.

(2007). The role and representation of contrast in phonological theory. Doctoral dissertation, University of Toronto.

(2008). Prophylactic features and implicit contrast. In Avery, Dresher and Rice (2008), 35–54.

Hall, T. Alan (ed.) (2001). *Distinctive feature theory*. Berlin: Mouton de Gruyter.

(2007). Segmental features. In de Lacy (2007), 311–34.

Halle, Morris (1954). The strategy of phonemics. *Word* 10: 197–209. Reprinted in Makkai (1972), 333–42.

(1959). *The sound pattern of Russian: a linguistic and acoustical investigation.* The Hague: Mouton. Second printing, 1971.

(1970). On markedness. In *Proceedings of the 6th International Congress of Phonetic Sciences*, 61–71. Prague: Academia.

(1995). Feature geometry and feature spreading. *Linguistic Inquiry* 26: 1–46.

Halle, Morris, Bert Vaux and Andrew Wolfe (2000). On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31: 387–444.

Hancin-Bhatt, Barbara Jean (1994). Segment transfer: a consequence of a dynamic system. *Second Language Research* 10: 241–69.

Hansson, Gunnar Ólafur (2001). Theoretical and typological issues in consonant harmony. Doctoral dissertation, University of California, Berkeley.

Hayes, Bruce (1984). The phonetics and phonology of Russian voicing assimilation. In Mark Aronoff and. Richard T. Oerle (eds.), *Language sound structure: studies in phonology presented to Morris Halle*, 318–28. Cambridge, Mass.: MIT Press.

Hayes, Bruce, Robert Kirchner and Donca Steriade (eds.) (2004). *Phonetically based phonology*. Cambridge: Cambridge University Press.

Heffernan, Kevin (2005). Phonetic similarity and phonemic contrast in loanword adaptation. In Chiara Frigeni, Manami Hirayama and Sara Mackenzie (eds.), *Toronto Working Papers in Linguistics (Special Issue on Similarity in Phonology)* 24: 117–23. Toronto: Department of Linguistics, University of Toronto.

(2007). The role of phonemic contrast in the formation of Sino-Japanese. *Journal of East Asian Linguistics* 16: 61–86.

Herd, Jonathon (2005). Loanword adaptation and the evaluation of similarity. In Chiara Frigeni, Manami Hirayama and Sara Mackenzie (eds.), *Toronto Working Papers in Linguistics (Special Issue on Similarity in Phonology)* 24: 65–116. Toronto: Department of Linguistics, University of Toronto.

Hirayama, Manami (2003). Contrast in Japanese vowels. In Daniel Currie Hall (ed.), *Toronto Working Papers in Linguistics (Special Issue on Contrast in Phonology)* 20: 115–32. Toronto: Department of Linguistics, University of Toronto.

Hockett, Charles F. (1955). *A manual of phonology.* Baltimore: Waverly Press.

Holt, D. Eric (1997). The role of the listener in the historical phonology of Spanish and Portuguese: an Optimality-Theoretic account. Doctoral dissertation, Georgetown University.

Hulst, Harry van der (1995). Radical CV phonology: the categorial gesture. In Jacques Durand and Francis Katamba (eds.), *Frontiers of phonology*, 80–116. Essex: Longman.

(1996). Radical CV phonology: the segment–syllable connection. In Jacques Durand and Bernard Laks (eds.), *Current trends in phonology: models and methods*, 333–61. Salford: European Studies Research Institute (ESRI).

(2005). The molecular structure of phonological segments. In Philip Carr, Jacques Durand and Colin J. Ewen (eds.), *Headhood, elements, specification and contrastivity: phonological papers in honour of John Anderson*, 193–234. Amsterdam: John Benjamins.

Hulst, Harry van der, and Jeroen van de Weijer (1995). Vowel harmony. In Goldsmith (1995), 495–534.

Hutton, John (1996). Optimality Theory and historical language change. Paper presented at the 4th Phonology Workshop, May 1996, Manchester, England.

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

(1970b). The role of borrowing in the justification of phonological grammars. *Studies in African Linguistics* 1: 1–48.

(1973). Nupe three years later. *Language* 49: 447–52.

(2001). Vowel Harmony in Gunu. *Studies in African Linguistics* 30: 146–70.

(2002). On the limits of phonetic determinism in phonology: *NC revisited. In Elizabeth Hume and Keith Johnson (eds.), *The role of perception in phonology*, 141–85. New York: Academic Press.

(2003). 'Abstract' vowel harmony in Kàlɔ̀ŋ: a system-driven account. In Patrick Sauzet and Anne Zribi-Hertz (eds.), *Typologie des langues d'Afrique et universaux de la grammaire*, 85–112. Paris: l'Harmattan.

Ingram, David (1988). Jakobson revisited: some evidence from the acquisition of Polish phonology. *Lingua* 75: 55–82.

(1989). *First language acquisition: method, description and explanation*. Cambridge: Cambridge University Press.

Itô, Junko, and Armin Mester (1986). The phonology of voicing in Japanese: theoretical consequences for morphological accessibility. *Linguistic Inquiry* 17: 49–73.

Itô, Junko, Armin Mester and Jaye Padgett (1995). Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26: 571–613.

Jacobs, Haike and Carlos Gussenhoven (2000). Loan phonology: salience, the lexicon and Optimality Theory. In Joost Dekkers, Frank van der Leeuw and Jeroen van de Weijer (eds.), *Optimality Theory: phonology, syntax, and acquisition*, 193–210. Oxford: Oxford University Press.

Jacobsen, William H., Jr (1968). On the prehistory of Nez Perce vowel harmony. *Language* 44: 819–29.

Jakobson, Roman (1941). *Kindersprache, Aphasie, und allgemeine Lautgesetze.* Uppsala: Uppsala Universitets Arsskrift. Translated by A. R. Keiler as *Child language, aphasia, and phonological universals.* The Hague: Mouton, 1968.

(1949). On the identification of phonemic entities. *Travaux du Cercle Linguistique de Copenhague* 5: 205–13. Reprinted in Jakobson (1962c), 418–25. Also in Makkai (1972), 318–22.

(1962a [1929]). Remarques sur l'évolution phonologique du Russe comparée à celle des autres langues slaves. *Travaux du Cercle Linguistique de Prague* 2. Reprinted in Jakobson (1962c), 7–116.

(1962b [1931]). Phonemic notes on Standard Slovak. In Jakobson (1962c), 221–30. [Published in Czech in *Slovenská miscellanea (Studies presented to Albert Pražak)*. Bratislava, 1931.]

(1962c). *Selected writings I. Phonological studies.* The Hague: Mouton.

Jakobson, Roman, C. Gunnar M. Fant and Morris Halle (1952). *Preliminaries to speech analysis.* MIT Acoustics Laboratory, Technical Report, No. 13. Reissued by MIT Press, Cambridge, Mass., eleventh printing, 1976.

Jakobson, Roman, and Morris Halle (1956). *Fundamentals of language.* The Hague: Mouton.

Jakobson, Roman, and John Lotz (1949). Notes on the French phonemic pattern. *Word* 5: 151–8.

Ji, Yonghai, Zhizhong Zhao and Liyuan Bai (1989). *Xiandai Manyu Babaiju* [Eight Hundred Sentences of the Modern Manchu]. Beijing: Zhongyang Minzu Xueyuan Chubanshe.

Jiménez, Jesús (1998). Valencian vowel harmony. *Rivista di Lingüística* 10: 137–61.

Joos, Martin (1957). *Readings in linguistics*, Vol. I. Washington: American Council of Learned Societies.

Kabak, Bariş (2004). Acquiring phonology is not acquiring inventories but contrasts: the loss of Turkic and Korean primary long vowels. *Linguistic Typology* 8: 351–68.

Kaisse, Ellen M., and Patricia A. Shaw (1985). On the theory of Lexical Phonology. *Phonology Yearbook* 2: 1–30.

Kang, Yoonjung (2007). Interlanguage segmental mapping as evidence for the nature of lexical representation. *Language and Linguistics Compass* 1: 1–16.

Kaplan, Lawrence D. (1981). *Phonological issues in North Alaskan Inupiaq*. Fairbanks, AK: Alaska Native Language Center.

(1990). The language of the Alaskan Inuit. In Dirmid R. F. Collis (ed.), *Arctic languages: an awakening*, 131–58. Paris: UNESCO.

Kaun, Abigail Rhoades (1995). The typology of rounding harmony: an Optimality Theoretic approach. Doctoral dissertation, University of California, Los Angeles.

Kaye, Jonathan, Jean Lowenstamm and Jean-Roger Vergnaud (1985). The internal structure of phonological elements: a theory of charm and government. *Phonology Yearbook* 2: 305–28.

Kean, Mary-Louise (1980). The theory of markedness in generative grammar. Doctoral dissertation, MIT, Cambridge, Mass., 1975. Reproduced by the Indiana University Linguistics Club, Bloomington, Ind.

Keating, Patricia A. (1988). Underspecification in phonetics. *Phonology* 5: 275–92.

Kenstowicz, Michael (1994). *Phonology in generative grammar*. Oxford: Blackwell.

(2003). Salience and similarity in loanword adaptation – a case study of Fijian. Ms., Department of Linguistics and Philosophy, MIT, Cambridge, Mass. [To appear in *Language Sciences*.]

Kenstowicz, Michael, and Charles Kisseberth (1977). *Topics in phonological theory.* San Diego, Calif.: Academic Press.

Kingston, John (2007). The phonetics–phonology interface. In de Lacy (2007), 401–34.

Kingston, John, and Randy Diehl (1994). Phonetic knowledge. *Language* 70: 419–53.

Kiparsky, Paul (1965). Phonological change. Doctoral dissertation, MIT, Cambridge, Mass.

(1973). Phonological representations. In Osamu Fujimura (ed.), *Three dimensions of linguistic theory*, 1–135. Tokyo: TEC Company.

(1982). From cyclic to Lexical Phonology. In Harry van der Hulst and Norval Smith (eds.), *The structure of phonological representations (Part I),* 131–76. Dordrecht: Foris.

(1985). Some consequences of Lexical Phonology. *Phonology Yearbook* 2: 85–138.

(2000). Opacity and cyclicity. *Linguistic Review* 17: 351–65.

(2002). Syllables and moras in Arabic. In Caroline Féry and Ruben van de Vijver (eds.), *The syllable in Optimality Theory,* 147–82. Cambridge: Cambridge University Press.

(forthcoming). *Paradigm effects and opacity*. Stanford, Calif.: CSLI.

Kirchner, Robert (1997). Contrastiveness and faithfulness. *Phonology* 14: 83–111.

(1998). An effort-based approach to consonant lenition. Doctoral dissertation, University of California, Los Angeles.

Korn, David (1969). Types of labial vowel harmony in the Turkic languages. *Anthropological Linguistics* 11: 98–106.

Kuhl, Patricia, K. A. Williams, Francisco Lacerda, Kenneth N. Stevens and Björn Lindblom (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255: 606–8.

Kuroda, S. Y. (1967). *Yawelmani phonology* (Research monograph no. 43). Cambridge, Mass.: MIT Press.

LaCharité, Darlene, and Carole Paradis (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics* 33: 379–430.

Lahiri, Aditi, and Henning Reetz (2002). Underspecified recognition. In Carlos Gussenhoven and Natasha Werner (eds.), *Laboratory phonology*, Vol. VII, 637–75. Berlin: Mouton.

Leben, William R. (1973). Suprasegmental phonology. Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, Mass.

Levelt, Clara C. (1989). An essay on child phonology. MA thesis, Leiden University.

Li, Shulan, and Qian Zhong (1986). *Xiboyu Jianzhi* [A Brief Introduction to the Xibe language]. Beijing: Minzu Chubanshe.

Lightner, T. M. (1963). A note on the formation of phonological rules. *Quarterly progress report* (Research Laboratory of Electronics, MIT): 68, 187–9.

Liljencrants, Johan, and Björn Lindblom (1972). Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48: 839–62.

Lindblom, Björn (1986). Phonetic universals in vowel systems. In John J. Ohala and Jeri J. Jaeger (eds.), *Experimental phonology*, 13–44. New York: Academic Press.

Lloret, Maria-Rosa (2008). On the nature of vowel harmony: spreading with a purpose. In Antonietta Bisetto and Francesco E. Barbieri (eds.), *Proceedings of the XXXIII Incontro di Grammatica Generativa*, 15–35. Bologna: University of Bologna.

Lombardi, Linda (2002). Coronal epenthesis and markedness. *Phonology* 19: 219–51.

Mackenzie, Sara (2002). Contrast, inventories, and vowel harmony in OT. Paper presented at the 2002 annual conference of the Canadian Linguistic Association, Toronto. Ms., Department of Linguistics, University of Toronto.

(2005). Similarity and contrast in consonant harmony systems. In Chiara Frigeni, Manami Hirayama and Sara Mackenzie (eds.), *Toronto Working Papers in Linguistics (Special Issue on Similarity in Phonology)* 24: 169–82. Toronto: Department of Linguistics, University of Toronto.

(2009). Contrast and similarity in consonant harmony systems. Doctoral dissertation, University of Toronto.

Mackenzie, Sara, and B. Elan Dresher (2004). Contrast and phonological activity in the Nez Perce vowel system. In Pawel M. Nowak, Corey Yoquelet and David Mortensen (eds.), *Proceedings of the Twenty-ninth Annual Meeting of the Berkeley Linguistics Society, February 14–17, 2003*, 283–94. Berkeley, Calif.: Berkeley Linguistics Society.

Maddieson, Ian (1984). *Patterns of sounds.* Cambridge: Cambridge University Press.

Makkai, Valerie Becker (ed.) (1972). *Phonological theory: evolution and current practice.* New York: Holt, Rinehart and Winston.

Martinet, André (1960). *Éléments de linguistique générale.* Paris: Librairie Armand Colin.

(1964). *Elements of general linguistics.* With a foreword by L. R. Palmer. Translated by Elisabeth Palmer. Chicago: University of Chicago Press.

Maye, Jessica (2000). The acquisition of speech sound categories on the basis of distributional information. Doctoral dissertation, University of Arizona, Tucson.

Maye, Jessica, Janet F. Werker and LouAnn Gerken (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82: B101–B111.

McCarthy, John J. (1988). Feature geometry and dependency: a review. *Phonetica* 45: 373–418.

Mercado, Raphael (2002). Contrastive theory and Québécois French oral vowels. Paper presented at the MOT Phonology Workshop, February, McGill University.

Mohanan, K. P. (1986). *Lexical Phonology*. Dordrecht: Reidel.

(1991). On the bases of Radical Underspecification. *Natural Language and Linguistic Theory* 9: 285–325.

Morén, Bruce (2003). The Parallel Structures Model of feature geometry. In Johanna Brugman and Anastasia Riehl (eds.), *Working Papers of the Cornell Phonetics Laboratory*, Vol. XV, 194–270. Ithaca, NY: Department of Linguistics, Cornell University.

(2006). Consonant–vowel interactions in Serbian: features, representations and constraint interactions. *Lingua* 116: 1198–244.

Moulton, Keir (2003). Deep allophones in the Old English laryngeal system. In Daniel Currie Hall (ed.), *Toronto Working Papers in Linguistics (Special Issue on Contrast in Phonology)* 20: 157–73. Toronto: Department of Linguistics, University of Toronto.

Nevins, Andrew Ira (2004). Conditions on (dis)harmony. Doctoral dissertation, MIT, Cambridge, Mass.

(2005). Microvariations in harmony and value-relativized parametrization. *Linguistic Variation Yearbook* 5: 187–232.

Newman, Stanley (1944). *Yokuts language of California* (VFPA 2). New York: The Viking Fund Publications in Anthropology.

Ní Chiosáin, Máire, and Jaye Padgett (2001). Markedness, segment realization and locality in spreading. In Linda Lombardi (ed.), *Constraints and representations: segmental phonology in Optimality Theory*, 118–56. Cambridge: Cambridge University Press.

Padgett, Jaye (1991). Stricture in feature geometry. Doctoral dissertation, University of Massachusetts, Amherst.

(2002). Feature classes in phonology. *Language* 78: 81–110.

(2003a). Contrast and post-velar fronting in Russian. *Natural Language and Linguistic Theory* 21: 39–87.

(2003b). The emergence of contrastive palatalization in Russian. In D. Eric Holt (ed.), *Optimality Theory and language change* (Studies in Natural Language and Linguistic Theory 56), 307–35. Dordrecht: Kluwer Academic Publishers.

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

Paradis, Carole, and Jean-François Prunet (2000). Nasal vowels as two segments: evidence from borrowings. *Language* 76: 324–57.

Pater, Joe (2003). The perceptual acquisition of Thai phonology by English speakers: task and stimulus effect. *Second Language Research* 19: 209–23.

(2004). Bridging the gap between perception and production with minimally violable constraints. In René Kager, Joe Pater and Wim Zonneveld (eds.), *Constraints in phonological acquisition*, 219–44. Cambridge: Cambridge University Press.

Pater, Joe, Christine L. Stager and Janet F. Werker (2004). The perceptual acquisition of phonological contrasts. *Language* 80: 384–402.

Penny, Ralph (1969). *El habla Pasiega. Ensayo de dialectología montañesa*. London: Tamesis Books.

(1970). Mass-nouns and metaphony in the dialects of North-Western Spain. *Archivum Linguisticum* 1: 21–30.

Pesetsky, David (1979). Russian morphology and lexical theory. Ms., Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Pham, Hoa (1997). The coronal–velar relationship in Vietnamese dialects. MA thesis, University of Toronto.

(1998). The coronal–velar relationship in Vietnamese: a prosodic account. *Asia-Pacific Language Research* 1.

Pierrehumbert, Janet B., Mary E. Beckman and D. R. Ladd (2000). Conceptual foundations of phonology as a laboratory science. In Noel Burton-Roberts, Philip Carr and Gerard Docherty (eds), *Phonological knowledge: conceptual and empirical issues*, 273–303. Oxford: Oxford University Press.

Piggott, Glyne (1992). Variability in feature dependency: the case of nasality. *Natural Language and Linguistic Theory* 10: 33–77.

Pinker, Steven (1994). *The language instinct: how the mind creates language*. New York: William Morrow.

Polański, Kazimierz (1993). Polabian. In Bernard Comrie and Greville G. Corbett (eds.), *The Slavonic languages*, 795–824. London: Routledge.

Polka, Linda, Connie Colantonio and Megha Sundara (2001). A cross-language comparison of /d/ – /ð/ perception: evidence for a new developmental pattern. *Journal of the Acoustical Society of America* 109: 2190–201.

Prince, Alan, and Paul Smolensky (2004). *Optimality Theory: constraint interaction in generative grammar.* Oxford: Basil Blackwell. [First published as Technical Report TR-2, Rutgers Center for Cognitive Science, Rutgers University, New Brunswick, NJ, 1993.]

Pulleyblank, Douglas (1986). *Tone in Lexical Phonology*. Dordrecht: Reidel.

Pye, Clifton, David Ingram and Helen List (1987). A comparison of initial consonant acquisition in English and Quiché. In Keith E. Nelson and Ann Van Kleeck (eds.), *Children's language*, Vol. VI, 175–90. Hillsdale, NJ: Erlbaum.

Radišić, Milica (2007). Velar /g/ in Serbian – stop or fricative: a contrastive account. Ms., Department of Linguistics, University of Toronto.

Reh, Mechthild (1996). *Anywa language: description and internal reconstruction*. Cologne: Ruddiger Koppe.

Rice, Keren (1993). A reexamination of the feature [sonorant]: the status of 'sonorant obstruents'. *Language* 69: 308–44.

(1994). Peripheral in consonants. *Canadian Journal of Linguistics / Revue canadienne de linguistique* 39: 191–216.

(1996). Default variability: the coronal–velar relationship. *Natural Language and Linguistic Theory* 14: 493–543.

(1997). Japanese NC clusters and the redundancy of postnasal voicing. *Linguistic Inquiry* 28: 541–51.

(2002). Vowel place contrasts. In Mengistu Amberber and Peter Collins (eds.), *Language universals and variation,* 239–70. Westport, Conn.: Praeger.

(2003). Featural markedness in phonology: variation. In Lisa Cheng and Rint Sybesma (eds.), *The 2nd SOTA book*, 387–427. Berlin: Mouton de Gruyter.

(2005). Sequential voicing, postnasal voicing, and Lyman's Law revisited. In Jeroen van de Weijer, Kensuke Nanjo and Tetsuo Nishihara (eds.), *Voicing in Japanese,* 25–45. Berlin: Mouton de Gruyter.

(2007). Markedness in phonology. In de Lacy (2007), 79–97.

(forthcoming). *Featural markedness*. Cambridge: Cambridge University Press.

Rice, Keren, and Peter Avery (1993). Segmental complexity and the structure of inventories. In Carrie Dyck (ed.), *Toronto Working Papers in Linguistics (Special Issue on Complexity in Phonological Representations)* 12,2: 131–54. Toronto: Department of Linguistics, University of Toronto.

(1995). Variability in a deterministic model of language acquisition: a theory of segmental elaboration. In John Archibald (ed.), *Phonological acquisition and phonological theory*, 23–42. Hillsdale, NJ: Lawrence Erlbaum.

Rice, Keren, and Sharon Hargus (eds.) (2005). *Athabaskan prosody*. Philadelphia: John Benjamins.

Rigsby, Bruce J., and Michael Silverstein (1969). Nez Perce vowels and Proto-Sahaptian vowel harmony. *Language* 45: 45–59.

Rohany Rahbar, Elham (2006). Vowel height in Persian. Ms., Department of Linguistics, University of Toronto.

(2008). On distinguishing Persian vowels. In Susie Jones (ed.), *Proceedings of the 2008 annual conference of the Canadian Linguistic Association*, www.chass.utoronto.ca/∼cla-acl/actes2008/CLA2008_Rahbar.pdf.

Rose, Sharon (1993). Coronality and vocalic underspecification. In Carrie Dyck (ed.), *Toronto Working Papers in Linguistics (Special Issue on Complexity in Phonological Representations)* 12,2: 155–77. Toronto: Department of Linguistics, University of Toronto.

Rose, Sharon, and Rachel Walker (2004). A typology of consonant agreement as correspondence. *Language* 80: 475–531.

Rubach, Jerzy (2000). Glide and glottal stop insertion in Slavic languages: a DOT analysis. *Linguistic Inquiry* 31: 271–317.

(2003). Polish palatalization in Derivational Optimality Theory. *Lingua* 113: 197–237.

Sagey, Elizabeth (1986). The representation of features and relations in non-linear phonology. Doctoral dissertation, MIT, Cambridge, Mass.

Sapir, Edward (1925). Sound patterns in language. *Language* 1: 37–51. Reprinted in Joos (1957), 19–25.

(1933). La réalité psychologique des phonèmes. *Journal de Psychologie Normale et Pathologique* (Paris) 30: 247–65. English version published as The psychological reality of phonemes in D. G. Mandelbaum (ed.), *Selected writings of Edward Sapir in language, culture and personality*, 46–60. Berkeley: University of California Press, 1949.

Saussure, Ferdinand de (1972 [1916]). *Cours de linguistique générale. Publié par Charles Bally et Albert Sechehaye; avec la collaboration de Albert Riedlinger. Éd. critique préparée par Tullio de Mauro.* Paris: Payot.

(1986). *Course in general linguistics.* Translation of Saussure (1916) by Roy Harris. La Salle, Ill: Open Court.

Shannon, Claude Elwood (1993 [1956]). The zero error capacity of a noisy channel. In N. J. A. Sloane and Aaron D. Wyner (eds.), *Claude Elwood Shannon: collected papers*, 221–38. Piscataway, NJ: IEEE Press. Originally published in *IRE Transactions on Information Theory*, IT-2.

Shevelov, George Y. (1964). *A prehistory of Slavic: the historical phonology of Common Slavic.* Heidelberg: Carl Winter.

Silverman, Daniel (1992). Multiple scansions in loanword phonology: evidence from Cantonese. *Phonology* 9: 289–328.

Smith, Eric J. M. (2007). [-ATR] harmony and the vowel inventory of Sumerian. *Journal of Cuneiform Studies* 59: 19–38.

Smith, Neil V. (1973). *The acquisition of phonology: a case study*. London: Cambridge University Press.

Stager, Christine L., and Janet F. Werker (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature* 388: 381–2.

Stanley, Richard (1967). Redundancy rules in phonology. *Language* 43: 393–436.

Steriade, Donca (1987). Redundant values. In Anna Bosch, Barbara Need and Eric Schiller (eds.), *CLS 23: Papers from the 23rd Annual Regional Meeting of the Chicago Linguistic Society. Part Two: Parasession on Autosegmental and Metrical Phonology*, 339–62. Chicago: Chicago Linguistic Society.

(1995). Underspecification and markedness. In Goldsmith (1995), 114–74.

(1997). Phonetics in phonology: the case of laryngeal neutralization. Ms., Department of Linguistics, University of California, Los Angeles.

(2009). The phonology of perceptibility effects: the P-Map and its consequences for constraint organization. In Kristin Hanson and Sharon Inkelas (eds.), *The nature of the word: studies in honor of Paul Kiparsky*, 151–79. Cambridge, Mass.: MIT Press.

Stevens, Kenneth N., and Samuel Jay Keyser (1989). Primary features and their enhancement in consonants. *Language* 65: 81–106.

Stevens, Kenneth N., Samuel Jay Keyser and Haruko Kawasaki (1986). Toward a phonetic and phonological theory of redundant features. In Joseph S. Perkell and Dennis H. Klatt (eds.), *Symposium on invariance and variability of speech processes*, 432–69. Hillsdale, NJ: Lawrence Erlbaum.

Svantesson, Jan-Olaf (1985). Vowel Harmony Shift in Mongolian. *Lingua* 67: 283–327.

Timberlake, Alan (2004). *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Trehub, Sandra E. (1976). The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47: 466–72.

Trubetzkoy, N. S. (1939). *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.

(1969). *Principles of phonology*. Translation of Trubetzkoy (1939) by Christiane A. M. Baltaxe. Berkeley: University of California Press.

(2001 [1936]). A theory of phonological oppositions. In Anatoly Liberman (ed.), *N. S. Trubetzkoy: studies in general linguistics and language structure*, 14–21. Translated by Marvin Taylor and Anatoly Liberman. Durham, NC: Duke University Press. [First published as Essai d'une théorie des oppositions phonologiques. *Journal de Psychologie Normale et Pathologique* 33 (1936): 5–18.]

Tryon, D. T. (1970). *An introduction to Maranungku* (Pacific Linguistics Series B 14). Canberra: Australian National University.

Tucker, Archibald (1994). *A grammar of Kenya Luo (Dholuo)*. Cologne: Ruddiger Koppe.

Twaddell, W. Freeman (1935). On defining the phoneme. *Language Monograph*, 16. Reprinted in Joos (1957), 55–80.

Underhill, Robert (1976). The case for an abstract segment in Greenlandic. *International Journal of American Linguistics* 42: 349–58.

Vaux, Bert (2000). Disharmony and derived transparency in Uyghur vowel harmony. *North East Linguistic Society* (Amherst, Mass.: GLSA) 30: 671–98.

Vilks, Baiba (2002). The contrast hierarchy in the Latvian consonant inventory. Paper presented at the MOT Phonology Workshop, McGill University. Poster presented at the Second International Conference on Contrast in Phonology, University of Toronto.

Walker, Rachel (1993). A vowel feature hierarchy for contrastive specification. In Carrie Dyck (ed.), *Toronto Working Papers in Linguistics (Special Issue on Complexity in Phonological Representations)* 12,2: 179–97. Toronto: Department of Linguistics, University of Toronto.

(2001). Round licensing, harmony, and bisyllabic triggers in Altaic. *Natural Language and Linguistic Theory* 19: 827–78.

(2005). Weak triggers in vowel harmony. *Natural Language and Linguistic Theory* 23: 917–89.

Weiss, Daniel J., and Jessica Maye (2008). The role of contrast in the acquisition of phonetic systems. In Avery, Dresher and Rice (2008), 219–29.

Werker, Janet F., Christopher T. Fennell, Kathleen M. Corcoran and Christine L. Stager (2002). Age and vocabulary size influences on the phonological representation of newly learned words in infants aged 14 to 20 months. *Infancy* 3: 1–30.

Werker, Janet F., John H. V. Gilbert, Keith Humphrey and Richard C. Tees (1981). Developmental aspects of cross-language speech perception. *Child Development* 52: 349–55.

Werker, Janet F., and Richard C. Tees (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7: 49–63.

Wheeler, Max W. (2005). *The phonology of Catalan*. Oxford: Oxford University Press.

Williamson, Kay (1965). *A grammar of the Kolokuma dialect of Ijo* (West African Language Monographs 2). Cambridge: Cambridge University Press.

Wu, Yuwen (1994). Mandarin segmental phonology. Doctoral dissertation, University of Toronto.

Yip, Moira (1993). Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* 2: 261–91.

Zhang, Xi (1996). Vowel systems of the Manchu-Tungus languages of China. Doctoral dissertation, University of Toronto.

Zhang, Xi, and B. Elan Dresher (1996). Labial harmony in Written Manchu. *Saksaha*: *A Review of Manchu Studies* 1: 13–24.

(2004). Tongue root harmony in Written Manchu. In Carsten Naher (ed.), *Proceedings of the First International Conference on Manchu-Tungus Studies (Bonn, August 28 – September 1, 2000) II: Trends in Tungusic and Siberian linguistics,* 161–90. Wiesbaden: Harrassowitz.

Zhao, Jie (1989). *Xiandai Manyu Yanjiu* [Studies on the Modern Manchu Language]. Beijing: Minzu Chubanshe.

Zhou, Hong (1999). Vowel systems in Mandarin languages. Doctoral dissertation, University of Toronto.

Zwicky, Arnold M. (1971). More on Nez Perce: an alternative analysis. *International Journal of American Linguistics* 37: 122–6.

Index of names

Subject index