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	v̪	r	z	ɣ	ʁ
	|	|	|	|	|	|	
	r - l	f	r̪	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.