Feature Theory

Mark Tatham

Tatham, Mark (2005) Feature Theory. In *Encyclopaedia of Linguistics*, 2 vols., ed. Philipp Strazny. New York: Fitzroy Dearborn

---

Introduction

Classical Phonetics felt the need for a classification system for speech segments to give some indication of how sounds are made and at the same time to identify relationships between them. It used the place-manner system for consonants and the high-low / front-back system for vowels. The main purpose here was clearly to enable the phonetician to specify how particular sounds were made with respect to their articulation. There was, however, an important spin-off: it became possible to use the features or parameters of the classification system to label whole sets of sounds or articulations. Thus we might refer to: 'the set of all plosives' (seven in English: *p, t, k, b, d, g, and the glottal stop*), or 'the set of all voiced plosives' (three in English: *b, d, g*), or 'the set of all voiced alveolar plosives' (one only in English: *d*) – and so on, cutting horizontally and vertically around the consonant matrix. Similarly, for vowels, 'the set of all front vowels' (English: *i, e, æ*), or 'the set of all rounded vowels' (English: *u, o*), and so on.

For the most part the categories or features used by Classical Phonetics reflected what phoneticians regarded as the salient parameters of articulation: place features such as bilabial, dental, alveolar, palatal, velar, pharyngeal, laryngeal, etc., and manner features such as plosive, fricative, continuant, etc. This is not an exhaustive list, but the idea was to characterise for each consonantal sound *where* in the vocal tract the sound was made and *how* it was made. In the case of vowels the relative position of the relevant part of the tongue on the high-low and front-back axes of a two dimensional co-ordinate system within the oral cavity would be sufficient, using front-back features like such as front, central and back, and high-low features such as high, mid and low.

It was sometimes necessary to add to these basic classificatory features to reflect details of the articulation, often depending on how finely the phonetician needed to specify articulations. In the case of vowels some additional features were necessary to expand on the two dimensional co-ordinate systems: vowels could in addition be spread or round, nasal or nasalised, and long or short; and consonants might be either tense or lax. For example, in English the high front vowel [i] is said to be long, whereas the low front vowel [æ] is said to be short; [t] and [s] are tense sounds, but [b] and [v] are not tense.

As a spin-off of being able to label sets of sounds or articulations in this way it became possible to describe the *behaviour* of various sets – that it, the behaviour common to members of now identifiable sets could be characterised. So, for example, it was possible to say that the set of voiced plosives devoice in word-final position or that all vowels lengthen before voiced plosives in the same syllable, and so on. So rules no longer had to be about the contextual behaviour of individual sounds – but in terms of how sets or classes of sounds behave. Phoneticians now had the ability to capture and express generalisation – an important theoretical principle in linguistics: generalisations *must* be expressed whenever possible. Capturing generalisations is the basis of predictive modelling because it means that we can ask the question: What about any other segment which has these same features? – even though we may not yet have analysed it. Generalisation leads to predictions and hypotheses, and these form a strong basis for scientific advance.

For all its potential as the basis of a truly scientific study of speech, the feature classification system afforded by Classical Phonetics however rested in fact on very weak theory. For example, a given sound might have its relevant features identified, but phoneticians did not yet have a way of saying why there were gaps in how features could be combined, nor was it possible to predict why certain sounds could not in fact ever occur. Why can we not have a bilabial plosive which is also a fricative? Why is it not possible to have a voiced glottal fricative? We can easily supply answers to these questions – but it is hard to construct a theory around this system that predicts all constraints and at the same time all possibilities in an *explanatory* way – which is ultimately what we're after. We want to know *why* things are the way they are, as well simply being able to characterise them *as* they are.

It was not until Transformational Generative Grammar came along, though, that these generalisations became formalised in recent phonological theory. Morris Halle's *Sound Pattern of Russian* (1959) was really the first influential textbook in modern phonological theory (just two years after Noam Chomsky's *Syntactic Structures* (1957), the first influential textbook in modern syntactic theory). The Generative Phonologists adopted the theory of distinctive features from the earlier Prague School of Linguistics (Nikolai Trubetskoy 1958 [1939]) – a much more formal representational system than that of the classical phoneticians.

Distinctive Feature Theory in Phonology

The use of distinctive features in phonology enables us to capture *natural classes*, and, by extension, to generalise regularly occurring phenomena and to formulate predictions about the behaviour of class members. If we wanted to hypothesise about human processing of phonology we would use this idea to suggest that human beings process the patterns of phonology as part of speech planning in terms of these classes rather than in terms of individual segments. The regularity of patterning in phonology is part of the evidence for this claim – but the claim is more solid when based on the evidence that when the users of a language make up new words they do so by producing utterances which obey the rules of the natural classes their sounds fall into.

There have been various sets of distinctive features proposed as the parameters of segment description and classification. The original set appeared in Jakobson, Fant and Halle (1952), and consisted of around 14 features; Jakobson and Halle (1956) had 12. Chomsky and Halle (1968) had around 45 features, explaining that they found the original set of 14 somewhat inappropriate for characterising some subtleties in phonology.

Most modern phonologists argue for a binary system of indexing features: a segment either possesses or does not any one particular features. The point here is that whereas when describing the physical world of phonetics it may be useful to have a system capable of capturing multi-valued features (an *n*-ary system), in the cognitive or perceptual world of phonology the binary system is preferable. For example, although we might find that there is somewhat more coarticulation between [–nasal] vowels and surrounding [+nasal] consonants (in words like *man* or *moon*) in some US accents of English than in most British accents of English, phonologically this observation is not relevant. There are no nasal vowels in English, though we may want to note that in some accents they are nasalised in some environments – degree of nasalisation being of no consequence.

Clearly, with a binary system of indexing the maximum number of features needed to uniquely classify the sounds of a language like English (with around 45 phonemes) would be six, giving us 2^6 or 64 segments. More would be needed to uniquely classify the sounds of all the languages of the world or indeed all possible human languages. Larger sets of features were chosen because it was felt appropriate to sacrifice mathematical simplicity or elegance in favour of a feature labelling system which appeared to related these phonological features with the phonetic set of Classical Phonetics. Thus the *meaning* of the features became more transparent.

These ideas are embodied in three principles surrounding the distinctive feature set. They should be able to

1. characterise all *contrasting segments* in human languages;
2. capture *natural classes* in a clear fashion;
3. be transparent with regard to *phonetic correlates*.

A claim inherent in the *first principle* is that the feature set somehow might embody the universal humanness of features rather than any language-*specific* nature. It is predicted that if this set is correctly specified no other features will be needed even for future languages, so long as human beings do not change how they make and handle language – that is, so long as human beings remain human.

The *second principle* refers to not just classes, but *natural* classes. The idea here is that the classes themselves reveal something of what is natural in human language behaviour, once again referring to the fact that phonological processing is a human activity, and will therefore contain elements which are truly universal.

The *third principle* enables us to establish phonetic similarity – that is, to group sounds which are phonetically similar by feature. In the end there is a very good reason for doing this: it becomes possible to explain some phonological processes in terms of the behaviour of their phonetic correlates.

The distinctive feature set most usually found these days is approximately that of Halle and Clements (1983), which is based on the Chomsky and Halle (1968) set. A detailed description can be found in Gussenhoven and Jacobs (1998). Chomsky and Halle have a lengthy description of their own set.

Feature sets

[The tables in this section are based on Appendix 2 of Clark and Yallop (1990).]

The 12 features found in Jakobson and Halle (1956) can be related to what it is they characterise articulatory and acoustically. These are abstract, phonological features so their relationship with phonetic characterizations is one of relative correlation, rather than absolute definition. The physical correlates are only approximate – there are many variants on these simple descriptions both between languages and dialects, and between different speakers.

feature	***opposing feature***	***articulatory correlation***	***acoustic correlation***
vocalic	non-vocalic	vocal cord vibration, relatively unobstructed vocal tract	periodic vocal cord excitation and clear formant structure
consonantal	non-consonantal	partial or complete vocal tract constriction	overall energy relatively low
compact	diffuse	front ‘resonance chamber’ dominates	energy focused toward the centre of the spectrum
tense	lax	vocal organs relatively tense or the entire tract voluntarily greatly distorted	high energy, spread throughout the spectrum
voiced	voiceless	vocal cord vibration present	periodic (laryngeal) excitation
nasal	oral	nasal cavity brought into play – the velum is lowered	additional nasal formant(s) present
discontinuous	continuous	vocal tract rapidly closing and opening	interruption of the acoustic ‘flow’
strident	mellow	turbulence created at the place of articulation	temporally unstructured noise at a relatively high frequency
checked	un-checked	glottalised	abrupt energy onset/offset
grave	acute	marginal within the vocal tract (not central)	energy focused on the lower part of the spectrum
flat	plain	constricted aperture (example is lip-rounding)	upper frequencies attenuated
sharp	plain	constriction of the upper oral cavity, but relative widening or the lower cavity or pharynx	lower frequencies attenuated

The Chomsky and Halle feature set is more comprehensive. There are some 27 basic articulatory features, though each has particular acoustic correlates. One point about these features is that they can be used at an abstract phonological or perceptual level, in which case they take on binary values, or the can in a fairly limited way be used at a physical phonetic level, in which case they can be multi-valued. It is important to realise, though, that the correlation is not necessarily linear or one-to-one.

feature	***articulatory correlation***
major class features	
1. sonorant	vocal cord vibration possible and usually present
2. vocalic (or syllabic)	vocal cord vibration possible, but vocal tract constriction restricted to vowel positions
3. consonantal	significant constriction present in the vocal tract
cavity features	
4. coronal	blade of the tongue raised
5. anterior	constriction in front of the palato-alveolar place
6. high	raised tongue body
7. low	lowered tongue body
8. back	retracted tongue body
9. rounded	labial rounding
10. distributed	extended place of articulation
11. covered	narrow and tense pharynx or raised larynx
12. glottal constriction	constriction produced by the vocal cords
13. nasal	velar port open
14. lateral	tongue sides lowered
manner of articulation features	
15. continuant ([–continuant] = stop)	relatively unimpeded air flow
16. instantaneous release	sudden release (as in plosives)
17. velar suction	velar closure (used in clicks, for example)
18. implosive	glottal closure produces inward air flow
19. velar pressure	*(unclear)*
20. ejective	glottal closure and increased pressure
21. tense ([–tense] = lax)	musculature is highly contracted
source features	
22. raised sub-glottal pressure	increased muscular contraction to raise sub-glottal pressure

23. voiced ([–voiced] = voiceless)	the vocal cords are vibrating
24. strident	air flow turbulence produced at point of articulation
prosodic features	
25. stressed	*(unclear)*
26. pitch (relative, scalar)	*(unclear)*
27. length (relative, scalar)	*(unclear)*

Example use of Distinctive Feature Theory: Redundancy

Redundancy is an important aspect of phonology which is captured by the use of distinctive features. Consider for example the fact that all segments in English which are [+nasal] are also [+voice]. We could say that to specify [+voice] for segments like [m] and [n] is to fail to capture this redundancy. The main distinctive feature here is the nasality – the voicing is secondary and entirely predictable: all nasal consonants are voiced – remember we are talking abstract phonology, not phonetics.

One of the principles involved in Distinctive Feature Theory is to set up a system to capture all the segmental *contrasts* in the world's languages. This can be done, *and* we can also show where there is *no* contrast: there is no contrast, nor possibility of contrast, where there is redundancy. If nasals are always voiced, then there cannot be a contrast involving voiceless nasals. Two things follow from this in the way we use features in the theory:

- we need only indicate those feature markings which contribute to the contrasts in a particular language;
- we can capture the redundancies in a separate table of *metarules* – rules outside the Distinctive Feature specification.

Omitting feature markings where there is redundancy means literally leaving the redundant cells blank in the distinctive feature matrix formed when segments have their feature specification characterised. The fact of the redundancy is captured by separate rules which take the general form:

if X then Y

or, in our specific example:

if [+nasal] then [+voice].

But why would we want to capture this redundancy, except to show that it is a regularity in the way segmental features pattern? We want to do this because speakers of a language *know* about the redundancy. Let's look at an example: there are three nasals in English, the nasal alveolar stop [n], the bilabial nasal stop [m] and the velar nasal [ŋ]. If we ask an English speaker to *invent* a new nasal – say, a palatal nasal like the one found in the French word *angeau* – they will also automatically make it [+voice]. It's as though they know that nasals must be voiced – which is another way of stating the rule above.

The very early original important text on this point was Richard Stanley's *Redundancy rules in phonology*, published in the journal *Language* in 1967 (vol. 43). These redundancy rules were called *segment structure rules* to contrast them with another type: *sequence structure rules*. The latter capture a speaker's knowledge of redundancy in the specification of segments themselves in patterned sequences. Thus if we have the sequence at the beginning of a syllable in English: CCCV... (with C=consonant, V=vowel), then the first C must be [s] – a completely redundant situation, since all we need to know is that there is a consonant there, followed by two others. In fact there are heavy constraints also, of course, on the remaining two consonants: the second one must be a plosive and the third must be liquid ([r, l])of some sort or a semi-vowel ([y, w]). Some phonologists have pointed out that the onset of a syllable strings consonants together with increasing sonority until you get to the vowel nucleus (the supreme sonorous segment), followed by a coda of consonants of decreasing sonority – though there are exceptions to this principle.

Stanley's work, which like much of the earlier work in modern phonology, adopted an essentially *linear* approach to analysis. Segments were characterised in terms of how they behaved in respect of their neighbouring segments in the surface string which represented the underlying phonological makeup of utterances. These days it is better to re-construct such analysis in a *non-linear* or hierarchical way, since this is much more revealing of the underlying constraints on the sequencing of segments on the surface. Nevertheless whatever the approach the focus is on structure in phonology and the way such structure is particularly well expressed using sets of constraints on what can and cannot happen in a phonological system.

Explanation is the ultimate goal, and for this to be fully effective it must be sourced from outside the domain of linguistics itself. So for example, we could use Distinctive Feature Theory to express an observed regularity that in many languages utterance final obstruents that are [+voice] often become [–voice], and this regularity is *explained* by reference to the failure of vocal cord vibration as the trans-glottal airflow decreases with falling sub-glottal air pressure – that is, the vocal cord vibration tends to stop as we run out of breath toward the end of an utterance. The explanation for the phonological regularity captured by regularities expressed through Distinctive Feature Theory lies *outside* phonology – in the domain of aerodynamics.

Distinctive Feature Theory makes a distinction then between the use of features for characterising the contrastive properties of phonological segments, and using them to indicate redundancy. An *incomplete* distinctive feature matrix uses blanks to indicate redundancy (and let you know where cells are the subject of redundancy rules), whereas a *fully specified* distinctive feature matrix has all cells filled with either a + or a – to indicate exhaustively just which features are present or absent for this particular segment.

Conclusion

There is much more to Distinctive Feature Theory than there is space to deal with here. It will be helpful to consult some of the readings below to find a fuller treatment. The main thing to remember about Distinctive Feature Theory is that it has been a significant step forward in classification from the rather crude phonetically-based ideas of Classical Phonetics. Remembering, however, that it is essentially a concept in abstract phonology (rather than phonetics), its principal importance lies in how it lends itself to capturing the generality of phonological processes and the principles underlying the structure of phonological segments.

References

Chomsky, Noam (1957) Syntactic Structures. The Hague: Mouton

Chomsky, Noam and Halle, Morris (1968) The Sound Pattern of English. Cambridge, Mass.: MIT Press

Clark, John and Yallop, Colin (1990) An Introduction to Phonetics and Phonology. Oxford: Blackwell

Gussenhoven, Carlos and Jacobs, Haike (1998) Understanding Phonology. London: Arnold

Halle, Morris (1959) The Sound Pattern of Russian. The Hague: Mouton

Halle, Morris and Clements, George .N. (1983) Problem Book in Phonology. Cambridge, Mass.: MIT Press

Jakobson, Roman and Halle, Morris (1956) Fundamentals of Language. The Hague: Mouton

Jakobson, Roman, Fant, Gunnar and Halle, Morris (1952) Preliminaries to speech analysis: the distinctive features and their correlates. Cambridge, Mass.: MIT Press

Stanley, Richard (1967) Redundancy rules in phonology. Language Vol. 43

Trubetskoy, Nikolai.S. (1958) Grundzüge der Phonologie, Göttingen: Vandenhoeck and Ruprecht (reprinted from Travaux du Cercle Linguistique de Prague (1939)