THE DEVIL IS IN THE DETAIL

Francis Nolan
Department of Linguistics, University of Cambridge, UK

ABSTRACT

I start with cosmology, where recent observations of supernovae suggest that the expansion of the universe is accelerating, whereas the dominant theory predicts that it should be slowing down. Where are the supernovae data of phonetics? Perhaps, I suggest, they are to be found in the linguistic phonetic details of speech. I go through a number of cases of what is sometimes called 'extrinsic' phonetic detail – detail which is not (contrastively) phonological, but which does not, either, emerge from physical principles. I deal with both segmental and suprasegmental examples. The details are intriguing, and I predict that whatever the attractions of elegant, general theories of speech, phoneticians will increasingly be bedevilled by the detail and diversity of phonetic realisation.

1. INTRODUCTION

It was recently widely publicised [32, 14] that astronomers, led by Saul Perlmutter in the Lawrence Berkeley Laboratory, have discovered one of the most distant supernova yet, 10 billion light years away. That figure echoed round my mind. The cataclysmic event whose consequences were being observed had happened ten thousand million years ago, long before our solar system was formed. During all that time radiation had journeyed across the unimaginably vast distances of space before arriving here. The scale of the universe, beyond normal comprehension, is perhaps best expressed in that classic work *The Hitch Hiker's Guide to the Galaxy* [1]: 'Space is big. Really big. You just won't believe how vastly hugely mindbogglingly big it is.'

The point of searching for not one, in fact, but tens and tens of supernovae, which appear fleetingly for a few weeks in the distant galaxies of the universe, is that they are known to all produce light of the same high intensity. Because of this, the speed at which they are moving away, or rather were moving away at the different times in the history of the universe from which their light reaches us, can be calculated; and from this the rate of expansion of the universe can be estimated. To everyone's astonishment, and contrary to received theories of the universe, the result appears to be that the expansion of the universe is speeding up, not slowing down. Astrophysics may be turned on its head, and the search is on for the hidden particles in space whose energy might be counteracting the force of gravity.

Probably all civilisations have been drawn to construct theories about what the stars are and why they are there. The theories provided a framework of thought which helped people to make sense of what they could see; but to an extent perhaps inhibited the search for new data. Once one has a model that appears to work, the urge to find more facts is inclined to be felt less strongly. The status quo is perhaps most likely to broken by technological developments (the telescope, for instance) than by pressures from within an apparently adequate theoretical framework. Seeing further, and more clearly into space, and having more sophisticated conceptual tools, has at various stages pushed the prevailing cosmology to, and beyond, its limits.

What has this got to do with the phonetic sciences? You will perhaps already have drawn your own parallels. Phonetics, as an adjunct to linguistics, has traditionally been dominated by a language-based perspective. In particular, the phonological notion of a system of contrastive elements, each defined by a distinctive difference between it and each other element, for a long time constituted the 'cosmology' of phonetics. Languages have systems of contrasting phonological units which are abstract or mental, and these are realised or implemented in the physical world. The linguist's interest is in the systems. But in the century which is just drawing to a close, technology and methodology have made data available which have pushed this view to its limits.

While astronomy and cosmology have looked ever further outwards, and found phenomena of ever increasing orders of magnitude, thanks in part to telescopes, phoneticians are more like scientists who have been given ever better microscopes. We are discovering, in contrast to the lean and minimal systems which constitute languages, a wealth of detail and richness in the implementation of those language systems. The theme of this talk, then, is that the more detailed our knowledge of the properties of speech becomes, the more difficult it is to sustain our simplifying assumptions, our models which help us comprehend our universe. I will look specifically at a number of cases of troublesome detail which may force us to revise the way we think.

2. A NEW TRADITIONAL SOURCE OF DETAIL

Our microscope is not purely a technological one. I have recently been involved, together with John Esling, the Secretary of the International Phonetic Association, in the final stages of editing the long-awaited *Handbook of the International Phonetic Association* [15]. I hope I can be forgiven, at a conference whose hosts include the President of the International Phonetic Association, John Ohala, for shamelessly 'plugging' this work, which contains a wealth of information useful to those who work in the speech sciences and linguistics, including a tutorial on the use of the International Phonetic Alphabet, and full listings of phonetic symbols, their names, and computer codings. Most impressively, I think, it also contains a set of nearly thirty 'Illustrations' which have appeared in the *Journal of the International Phonetic Association*. For those who are not familiar with these, they are a concise phonetic description of a language, including a transcription of a recorded passage. These Illustrations have been contributed over the last few years by members of the International Phonetic Association, and carry on a tradition of close phonetic observation which has a venerable history.

Like other surveys of the sounds and sound systems of languages such as [26] and [22], the Illustrations are testimony to the richness and diversity of the ways in which languages use sound. They provoke the question of just why languages are so diverse in their use of sounds. There are, of course, well known regularities and constraints: we know that if a vowel system 'wants' five contrasting vowels, there are good ways and bad ways of realising that ambition. A good way, as in Figure 1, is to space them out nicely in the available space, with front and back close and mid vowels and an open vowel. Hebrew [15:104] is a language exemplifying this sort of pattern.

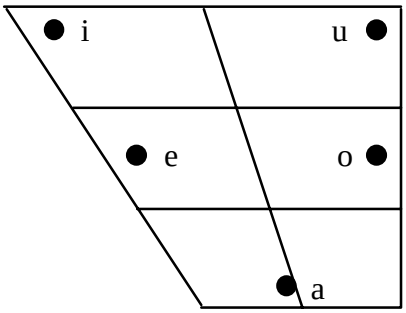

Figure 1. A good five-vowel system

A hypothetical bad way is shown in Figure 2. Four close vowels, the front one rounded and the back one unrounded, and an open front rounded vowel is not a good solution, to judge by what is found in the world's languages. And we have quite a good understanding of why the first system is a 'better' one than the second system, based on the general principle that vowels should spread themselves out in the available perceptual, and perhaps articulatory, space. More formally, attempts such as those of Lindblom [24, 25] to use the principle of maximising perceptual distance to model vowel systems had considerable success in explaining 'plausible' vowel systems.

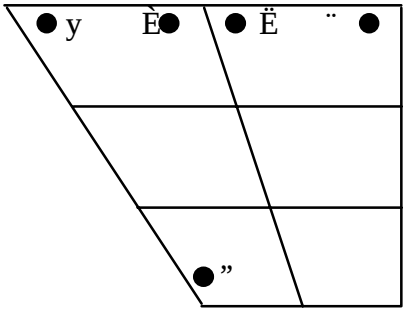

Figure 2. A not-so-good five-vowel system

This is one tiny corner of the universe of phonetic diversity where there are apparently regularities and we can explain them. But as has been pointed out (e.g. by Lass [23:134ff]) there is a danger of wishing too much regularity into the phonetic data.

Figure 3 shows the five-vowel system of Taba [15:150], an Austronesian language spoken in Maluku province, Indonesia. This still conforms to the 'ideal' distribution of vowels, but their spacing is imperfect.

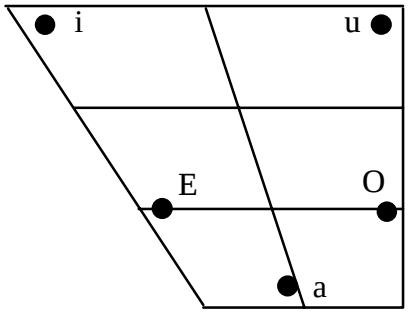

Figure 3. Five-vowel system of Taba

Tukang Besi [15:159], another Austronesian language spoken in south-east Sulawesi, Indonesia, deviates further from the 'ideal', showing a more skewed system.

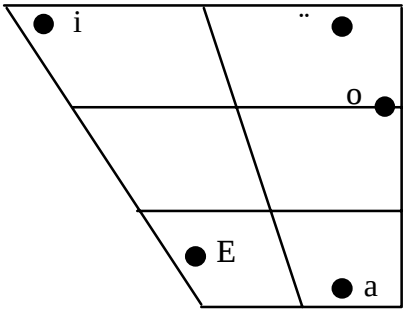

Figure 4. Five-vowel system of Tukang Besi

We have, then a few problems with phonetic detail even before we bring a technology to bear on our subject matter. Impressionistic phonetic observation alone gives us plenty to ponder on.

3. ARTICULATORY PHONOLOGY

Over the last decade at least one very influential school of thought has moved away from phonetic detail. Not in the sense that the data have been downplayed; on the contrary, the followers of this school have been among the most active in providing highly detailed quantitative descriptions of articulation. Browman and Goldstein (e.g. [3, 4, 5]) and the many researchers who have been inspired by their work have used techniques such as x-ray microbeam scanning, electropalatography, and electromyography to enhance greatly our understanding of the timing and coordination of the articulators. They have moved away from phonetic detail rather in the sense that some of it is regarded as an artefact of the way we describe speech, and in particular of our traditional segmentally and phonemically oriented view of speech.

Articulatory Phonology has been seen as providing an alternative to 'translation', the process implicit in the traditional phonetic world-view of resolving the dichotomy between invariant timeless phonological entities and the time-varying physical continuum of speech.

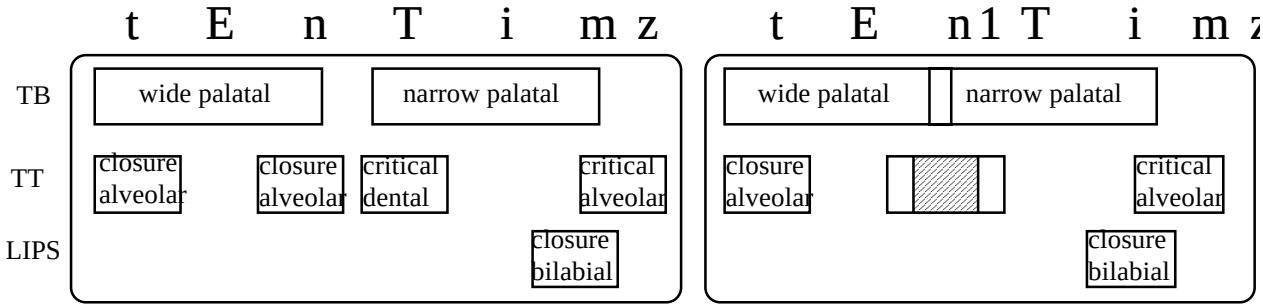

Figure 5. Schematic gestural scores for an English phrase spoken slower (left) and faster (right)

Articulatory Phonology rejects the view that speech activity is in a different domain from the phonological representation underlying speech. As is now widely familiar, one of the basic premises of Articulatory Phonology is that if a phonological representation is couched in appropriate terms, the disparity between 'plan' and 'implementation', between 'phonological representation' and 'phonetic interpretation' disappears.

The way in which this premise has been explored in Articulatory Phonology is to use 'gestures' as the primitives in phonetic-phonological representations. A gesture is, according to Browman and Goldstein [3:206]

> an abstract characterisation of coordinated task-directed movements of articulators ....[and is]...precisely defined in terms of the parameters of a set of equations for a 'task-dynamic' model...

Task-dynamics is a general model of skilled movement control based on dynamic equations such as might be used to describe the action of a mass attached to a spring (for an introduction see [12], and for a more technical account, [17]).

Figure 5, after [3], is a now-familiar representation, the 'gestural score', which shows the coordination of gestures making up a word. It is important to remember, however, that this is more than a superficial 'parametric diagram' of the articulators; the gesture 'boxes' merely stand in for their dynamic descriptions. Thus for instance when greater overlap occurs between gestures competing for the same articulator, as at the word boundary in the faster version on the right, a computationally specifiable blending of the trajectories resulting from the two gestures will occur – in this case, an articulation intermediate between alveolar and dental.

Since gestures are, in a sense, physical-world entities, and interact with each other particularly in the demands they make on the articulators, much of the allophonic variation of continuous speech is already inherent in the phonological representation, the 'gestural score'. In this way, the hope is that the problem of phonetic interpretation, as other than an expression of the physical laws of the universe, is defined away. It is no longer necessary to 'realise' or 'implement' a phonological velar as being a more fronted articulation before a front vowel and a more backed articulation before a back vowel; the velar automatically blends with the respective vowel articulations by virtue of the dynamic forces involved. The phonetic detail of speech is taken care of, given a sufficiently rich prosodic structure to account for (for instance) the fact that the magnitude and phasing of gestures varies according to syllable position.

4. A COMPLICATED ACCOMMODATION

Articulatory Phonology potentially provides a seamless, complete model from lexicon to acoustics. It has stimulated a great deal of valuable research, and it has redefined the way we think about phonetic detail, accounting for much of it according to general principles, essentially the laws of physics. The message, put in slightly biblical terms, is that we have been led astray by the pre-quantitative phonetician's allophones and narrow transcriptions, and we should now render unto the dynamicist what is the dynamicist's, and render unto the phonetician what is the phonetician's – if anything remains!

However there are hints that not everything can be rendered unto the dynamicist. In the case of segmental accommodation involving adjacent places of articulation and the same articulator, Browman and Goldstein [3: 220] predict gestural blending and an intermediate articulation; e.g. for dental and alveolar:

> the location of the constriction should not be identical to that of either an alveolar or a dental, but rather should fall somewhere in between.

Holst and Nolan [13] and Nolan, Holst & Kühnert [30] present evidence that this general account does not cover all observed possibilities. In an experiment [13] dealing with accommodation at boundaries of different types, acoustic analysis of abutting [s–S] at word boundaries (e.g. 'drinks sherry') yielded acoustic forms categorisable into the spectrographic schemata in Figure 6. At the bottom, [s]- and [S]-gestures are abstractly represented as time functions, with increasing overlap broadly corresponding to increasing speech rate. A, B, and C provide no problem for Articulatory Phonology. When the gestures are maximally separated, the resultant fricatives show their respective spectrographic patterns, with a lower cut-off frequency for the turbulent energy for [S]. Increasing overlap of the two gestures leads to a more transitional acoustic effect.

In D, the shaded form is the one which should be predicted by Articulatory Phonology. The duration should be no greater than for one fricative, and the acoustic pattern should have a cut-off intermediate between those of the two canonical fricatives In fact, the type 'D' friction observed (unfilled outline) was identical to that of [S] for each speaker; and duration was greater than that of initial [S]. The identity of the articulation to that of 'canonical' [S] is supported by EPG analysis on other speakers [30]. In short, if the magnitude of the [s] gesture has been reduced to zero the [S] ought to have the duration of an initial [S]; if the magnitude of the [s] gesture

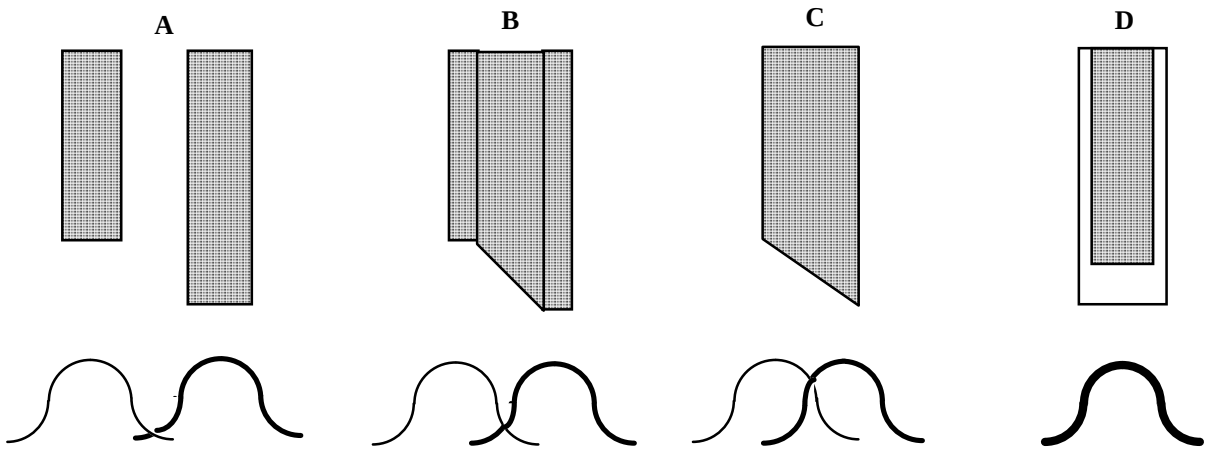

Figure 6. Schematised results of an experiment on fricative assimilation; above, spectrographic schemata; below, gestures.

has not been reduced to zero, then the acoustic effect should be seen. [30] also weakens the objection of Browman [2] that the acoustic results might be the result of quantal effects disguising the articulatory patterns, by showing that the relevant acoustic properties do very continuously between the two articulations.

Holst and Nolan [13, 30] argue that only if the speaker has intended to change the phonetic target of the first fricative (to [S]) can the long post-alveolar fricative be accounted for. Here, then, is a phonetic detail which cautions us against rendering all connected speech effects to the dynamicist.

On its own this is a small and controversial piece of evidence. But it should not surprise us, since today's connected speech processes, however dynamically motivated, are tomorrow's phonology. We are not tempted to think of *church* as having velar initial and final consonants because the Old English front vowels in *cirice,* which plausibly gave rise to palatal stops by gestural blending, and thence perceptual reinterpretation to post-alveolar affricates, have long since gone. The dynamic process has been phonologised. But in the synchronic state of a language, undergoing as any language does continual phonological change, it is axiomatic that effects must coexist resulting from connected speech processes at different stages of phonologisation. Not surprising, then, that close attention to phonetic detail has, arguably, teased apart two alternative kinds of sibilant assimilation. The general point is that only by probing with an open mind the phonetic details of processes which are apparent candidates for a gestural, dynamic, treatment will we establish the limits of the theory.

5. DIALECT VARIATION: LIQUIDS

As well as dealing elegantly with many boundary phenomena such as assimilation, Articulatory Phonology holds out the promise of providing a neat account of allophonic variation (as, in fact, in the *cirice* example above where the classic fronting of a velar before a close front vowel took place). One phenomenon in English which has come under particular scrutiny is that most famous of all allophonic alternations, clear and dark /l/. Sproat and Fujimura [33] claimed that the difference did not involve different gestures, in Articulatory Phonology terms, but a different phasing of two gestures depending on syllable position. The two gestures are tongue tip (or blade) closure (in the mid-line of the vocal tract) and tongue body retraction. The former can be regarded as 'consonantal', and the latter as 'vocalic'. In syllable-initial position, it is claimed in [33], the blade gesture is phased ahead of the tongue body gesture, and in syllable-final position, it is the tongue body gesture which is initiated first.

This work was carried out on American English. In American English, as has often been pointed out, all laterals tend to be relatively dark. It is interesting to consider, as Carter [6] has done, dialects of English exhibiting a variety of relationships between clearness and darkness in their liquids. He starts from an observation by Kelly and Local [16] to the effect that speakers with a clear initial lateral would have a dark rhotic, and speakers with a dark initial lateral would have a clear rhotic. In a rather neat way, then, the claim perhaps was that the secondary articulation supported contrastivity, but in a way independent of the primary articulation.

Carter took two 'non-rhotic' accents of English (non-rhotic in the sense of phonotactically prohibiting /r/ except where a vowel follows) and two 'rhotic' accents (where /r/ is in contrast with /l/ in all positions. Within each group one accent exemplified 'clear initial lateral' and the other 'dark initial lateral'. The dialects are summarised in Table 1. A formant analysis was then carried out on the initial and final rhotics.

	Non-rhotic	**Rhotic**
Clear initial lateral	Sunderland	County Tyrone
Dark initial lateral	Manchester	Fife

Table 1. Dialects used by Carter [6]

The situation which Carter uncovered is quite a complex one. In the case of the non-rhotic dialects, Sunderland had a darker initial /r/ than /l/, while the final /l/ was darker than the initial one. Manchester had a clearer initial /r/ than /l/, while the final /l/ was dark. On the other hand in the rhotic varieties, the initial rhotic was always darker than the initial lateral (even where, as in Fife, the initial lateral was dark), and final laterals were darker than initial laterals. Final rhotics, though, were

clearer than final laterals. It seems that the generalisation accounted for by [33] that final laterals are darker than initial ones is not refuted by these data; on the other hand the variation in degree of darkness of laterals and rhotics is more than could be accounted for by general principles.

In particular, from the point of view of Articulatory Phonology, although the general trend for laterals to be darker syllable-finally than syllable-initially is susceptible to a generalisation in terms of the phasing and magnitude of gestures, the details of the implementation of the contrasting liquids across dialects is clearly not susceptible of explanation in dynamic terms, since it is part of the language- (or dialect-) specific information which must be part of what is volitionally variable.

Carter concludes that

> The arrangement of gestures is not intrinsic to the phonology of syllable structure since it is dependent not only on position in syllable structure but also on dialect-specific phonetic interpretation.

That is, there is a need for what he terms 'extrinsic phonetic interpretation'. Carter is working within the conceptual framework of declarative phonology, but the data is interesting whatever the framework. The problem is how to account for phonetic detail which is neither contrastive (and hence has no place in a phonological representation as traditionally conceived) nor predictable from the physiology, dynamics, and so on of the speech mechanism. The issue is not new, and a distinction between intrinsic and extrinsic allophones, thus termed, is one drawn at least as early as 1972 by Ladefoged [20].

6. DIALECT VARIATION: VOICELESS STOPS

The point is reinforced if we look at another area of dialect variation in English, the treatment of voiceless alveolar stops in different varieties. This is perhaps where the greatest diversity in the English consonant system is found. A summary of some of this variation is given in Table 2.

In most accents of English /t/ at the start of a stressed syllable will be realised as an aspirated alveolar stop, sometimes with affrication. In a few accents, such as those of some parts of the north of England (here Yorkshire), there is relatively little aspiration.

The realisation of non stressed-syllable initial /t/ is articulatorily and acoustically diverse across the different accents, and also in different phonetic environments. Famously in much American English, foot-medially before a vowel, it will be a short voiced stop, flap, or tap (I shan't enter the debate here as to which of these it is, except to note that my transcription indicates my belief that the sound in question has neither the aerodynamic properties to qualify as a true tap nor the ballistic nature to qualify as a true flap, but is rather just a very short stop). This is a very tractable allophone for an Articulatory Phonology account; a reduction in the magnitude of the alveolar closure gesture will shorten the occlusion phase, which in turn reduces the amount of time that air is impeded from flowing through the glottis. If that were not enough, no doubt voicing would easily arise from greater overlap of the glottal voicing gesture for the vowels on either side of the stop. The phonetic description of the lenited forms in Southern Irish English, which are picked out in Table 2 by double lines around the boxes, is discussed in detail in [31]; broadly they are slit (as opposed to grooved) alveolar fricatives. These might also be the output of gesture magnitude reduction, as the alveolar closure is not quite achieved. It is not clear, however, what general dynamic principle could be appealed to in order to explain why in these three cases (American, Yorkshire, and Southern Irish) three distinctly different behaviours should arise.

With glottalisation we come, in my view, to the point where an account in terms of gestural reduction and overlap ceases to be plausible (note that the shaded boxes in Table 2 show forms with glottalisation). With the majority of voiceless consonants the vocal cords are moved apart, most dramatically so for fricatives which require high airflow through the vocal tract, but also for stops, and the aerodynamic effects which produce voicing are weakened. Voicing is inclined to cease, though not instantaneously because the opening is not instantaneous and (presumably) because of momentum in the vibrating cords. On spectrograms voicing can often be seen dying away during the closure phase of a perceptually voiceless stop. The opening for the consonants is evident in coarticulated breathy voice on the vowel preceding the stop or even in preaspiration [27]. The glottis-opening strategy is generally effective at producing voicelessness, but the facts that the voicing dies out gradually and the glottal opening has to be reversed very soon for a following vowel means that there is a danger of voicing continuing through a stop.

Another way of inhibiting voicing in a stop (but not a fricative) is to press the vocal cords together firmly enough that vibration ceases, i.e. a glottal stop. This will be reflected in some degree of creaky voice in the preceding vowel. The glottal closure can be simultaneous with an oral closure, giving a glottalised stop as in the North East of England (["bEt°/å]), but it seems that there is a strong temptation particularly with alveolars for a dialect to allow the glottal closure to take over the work of the oral closure, as in Cockney (["bE/å]).

Glottal opening and glottalisation (closure too tight for voicing) seem to me diametrically opposed strategies for solving the problem of how to achieve voicelessness in the predominantly voiced stream of speech. Opening is achieved by contracting the posterior cricoarytenoid muscles, closing by contracting the lateral cricoarytenoids and the interarytenoids. In Articulatory Phonology terms they have to be different gestures. When I have made this point before (e.g. [28]) there has been a polite murmuring of 'oh no they're not' from the articulatory phonologists, but I have yet to see a detailed account of how they can be other than diametrically opposed. The acoustic goal, in one sense, is the same – no voicing – but the ways of achieving it are mutually exclusive.

If glottalisation cannot arise from gestural weakening or overlap, which it seems to me it can't, then what are the alternatives for Articulatory Phonology? Notice that the problems do not arise simply from the view that the element we call in phonemic terms /t/ which occurs in different positions is 'the same thing'. We can abandon any notion of the segment, and think in terms of gestural scores for complete words, but what happens at word boundaries still needs to be taken care of. This would mean two different gestural scores for *get* in those dialects with glottalisation, or possibly three in North East England if the reported voicing assimilation process [18], which doesn't happen between vowels, is not simply the result

Accent	***tie***	***attain***	***date***	***better***	***get well***	***get a lot of***
RP	tÓaI	´"tÓeIn	deItÓ	"bEt“Ó‘´	gEt"•wE:	"gEtÓ´"lÅtÓ´v
SSB	tÓaI	´"tÓeIn	deI/°tÓ	"bEt“Ó‘´	gE/°t"wE:	"gEtÓ´"lÅtÓ´v
Cockney	tˢAI	´"tˢœIn	dœI/	"bE/å	gE/"wEØ	"gE/´"lÅ/´v
Yorkshire	taE	´"tE6…n	dE6…/°tÓ	"bE®´	gE/°t"wE:	gE®´"lA®´v
NE England	tÓœI	´"tÓe´	de´tÓ	"bEt°/	gEd"wEl	"gEt°/´"lÅ

of gestural overlap. This at least runs counter to the traditional phonological concern to capture generalisations and store the minimum number of forms in the lexicon.

The general point to be taken from the data here is that there is a considerable amount of volitionally controlled context sensitive variation in the realisation of English stops, a point already made, for instance, and supported at length in Docherty [7] with respect to voicing, and, implicitly, by many others who have worked specifically on glottalisation in English dialects. Glottalisation does not, I have argued, fall out automatically from the workings of the vocal mechanism; it is an alternative strategy for achieving voicelessness. Furthermore the patterning of glottalisation cannot follow entirely 'natural' principles, since it exhibits quite contrary patterns across dialects – see the shaded parts of Table 2, in particular the different pattern of North East English. Whereas glottalisation is generally most favoured when no vowel follows, in this dialect medial but not final position is favoured.

As far as theory is concerned, we are again forced back from the elegant Articulatory Phonology position that all connected speech processes fall out of dynamics. It is most unlikely that the gestural properties of the final /t/ of *get* differ, between say SSB (Standard Southern British) and Yorkshire in such a way as to predict that in the first case it becomes glottalised in *get a lot of* and in the second case it turns into an /r/. A similar point has been made recently by Ladefoged [21] on the basis of variation in /l/: 'Articulatory Phonology can explain many things, but it cannot explain all the phonological properties of languages.' Once more we have to conclude that between the lexicon and pronunciation there is a potentially underestimated amount of extrinsic phonetic detail to be added in. The gestural score is not (and cannot aspire to be) the sole phonological representation, but must itself be the output of quite complex processes of realisation or interpretation.

7. DIALECT VARIATION: INTONATION

In the past couple of decades our technology for analysing intonation, both electronic and in terms of models, has advanced dramatically. It is now standard to combine simultaneous listening and observation of the acoustic signal including its fundamental frequency derivative. A wealth of new suprasegmental phonetic detail has become more readily accessible, and we are only beginning to come to terms with it. Such detail has emerged from numerous single-variety studies, but it is brought more sharply into focus (as in the case of the segmental detail we examined above) when cross-variety studies are carried out. The successful analysis of the detail depends on recognising that intonation, just like segmental phonetics, has a phonological structure which (whatever terms one prefers to use) is subject to interpretation or realisation in a language-particular (and variety-particular) way.

One such aspect is embodied in the notion of 'truncation' and compression. Ladd [19:115-118] discusses the case of question intonation in Hungarian. Questions underlyingly have a rise-fall nucleus, L*HL in autosegmental terms. A rise fall in English (e.g. RP) tends to be used in marked contexts such as exclamations: *soda (indeed!), beer (indeed!)* and the full rise-fall will be 'compressed' onto even a monosyllable. In Hungarian, on the other hand, questions on the structurally parallel items *szódá?* ('water?') and *sör?* ('beer?') appear superficially to have different intonation patterns. Whilst *szódá* has a rise-fall, the monosyllabic *sör?* has a simple rise. The insight, however, is that while English 'compresses' the pitch movement defined by the three tones (here LHL) even onto a monosyllable, Hungarian will tolerate a maximum of two tones on any syllable, and here 'truncates' the rightmost leaving LH. This is illustrated schematically in Figure 7.

The phonetic forms make it look as if there are two alternative question intonations in Hungarian, a rise and a rise-fall. But their functional equivalence, and complementary distribution as far as syllable environment is concerned, mean that they can be regarded as phonologically equivalent. A difference in what Carter [6] would call extrinsic phonetic interpretation between the two languages results in the observed difference.

Grabe [9, 10] shows a similar difference between English and German falls (HL). In monosyllables in Standard Southern English, as the voiced segmental material becomes progressively shorter, the size of the fall is preserved by virtue of a more rapid rate of change. In comparable German utterances, however, a nuclear fall is effectively a high level tone. Underlyingly this is nonetheless HL, by the argument that as soon as longer segmental material is substituted in

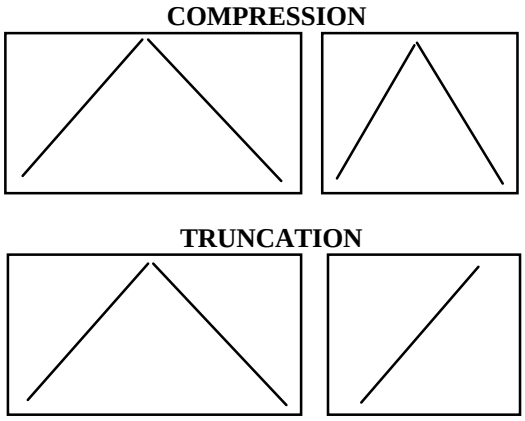

Figure 7. Distribution of an LHL pitch accent on longer and shorter phonetic material, showing the strategies of compression and truncation as in Hungarian.

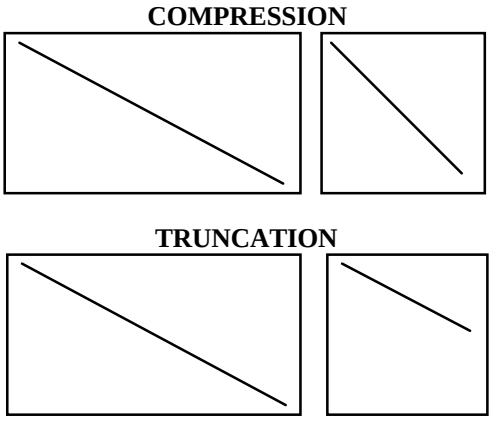

Figure 8. Distribution of an HL pitch accent on longer and shorter phonetic material, showing the strategies of compression and truncation as in German.

comparable contexts the fall emerges. This is summarised in Figure 8.

Grabe's data is less categorical than Ladd's description implies for Hungarian, since long German monosyllables show a degree of fall in Fo, suggesting that what is involved is not a categorical deletion of the final L but, perhaps, that the L is associated more abstractly with the latter part of a rhythmic foot template consisting of a stressed and unstressed syllable – whether or not the unstressed syllable is present. Rises, incidentally, are compressed in both languages.

In a project whose aim is to provide a systematic comparison of the intonation of several varieties of English in the British Isles [8], Grabe et al. [11] have discovered similar differences of compression and truncation behaviour between Standard Southern British, Leeds, Newcastle, and Belfast. Their experiment used names which varied in terms of the syllabic and segmental material available for voicing, e.g. 'Mr Sheafer' (disyllable), 'Mr Sheaf' (monosyllable, long vowel), and 'Mr Shift' (short vowel), in contexts eliciting the word as the intonation nucleus in a question or a statement. Speakers in three of the dialects used falling nuclei on the statements and rising nuclei on the questions. In Belfast, as is well known, the default statement intonation is a rise (more accurately a 'rise-plateau'), and this intonation was used by Belfast speakers in both contexts.

Truncation or compression was assessed primarily by determining the rate of Fo change in the voiced material associated with the word. Most intriguingly, it emerged that while Standard Southern British and Newcastle compress both rises and falls, Leeds truncates both rises and falls. Belfast, using only rises, truncates these. If we add German as a more distantly related variety, it exhibits a compromise between these two strategies, compressing rises and truncating falls (as noted above. Table 3 summarises this behaviour, with compression highlighted.

Truncation versus compression, then, is a strong candidate for consideration as an extrinsic phonetic difference, this time in the suprasegmental domain.

Another candidate which has emerged from [8] is the alignment of the peak associated with H* pitch accents at the beginning of an intonational phrase. Nolan and Farrar [29] show that all four of the dialects above are inclined to align the peak after the initial stressed syllable with which the accent is

	RISE	FALL
SSB	compresses	compresses
Newcastle	compresses	compresses
German	*compresses*	*truncates*
Leeds	truncates	truncates
Belfast	truncates	–

Table 3. Truncation / compression in different varieties.

phonologically associated, this tendency being much stronger when there are no unstressed syllables preceding the initial accented syllable. That is, unless there is 'anacrusis', the pitch peak will lag after the accented syllable. However, as shown in Figure 9, dialects differ in their propensity to lag.

For instance, Cambridge speakers achieve 50% of their Fo peaks on the phonologically associated syllable, and lag very few beyond the first following unstressed syllable ('extreme lag'), whereas Newcastle speakers have extreme lag almost as often as they match alignment with phonological association. Phonetic details such as these are undoubtedly important for our perceptual identification and characterisation of dialects

8. CONCLUSIONS

I hope in this paper to have provided some thought-provoking phonetic details which fall between the contrastively phonological and the physical. I predict that focus of the phonetic sciences will increasingly be on such detail. Its richness and diversity, between languages and within a language, has long been known, but its fine structure is only now becoming apparent through the application of new technology and new models. Contrastive phonological systems are well understood, for all that the rate of turnover of new formal models for capturing them is quite high; and the physical properties of speech in their own right are gradually being unfolded. But we do not have models which adequately accommodate the kinds of linguistic-phonetic variability which I have discussed.

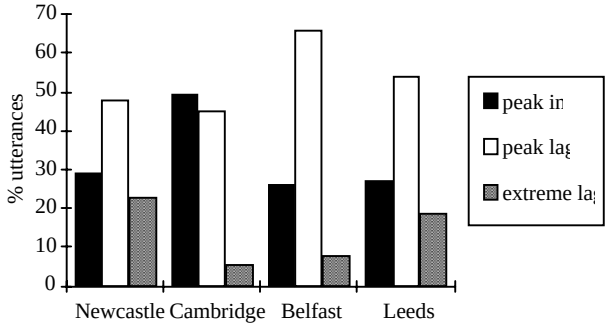

Figure 9. Peak alignment in different dialects.

I cannot promise that a revolution lies around the corner in the phonetic sciences of the kind which is heralded for cosmology by the claim of the 'accelerating universe'. But I do think that just as the attention of cosmologists will shift to confirming and explaining this claim, so the attention of the phonetic sciences will undergo a shift. The shift will be from hoping to find an elegant model based on contrastive phonology and implementation based on physical principles, to the acknowledgement of extrinsic phonetic interpretation and its incorporation into an overall theory.

REFERENCES

[1] Adams, D. 1979. *The Hitch Hiker's Guide to the Galaxy*. London: Pan (Macmillan).

[2] Browman, C.P. (1995) Assimilation as gestural overlap: comments on Holst and Nolan. In B. Connell and A. Arvaniti (eds), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV*. Cambridge: Cambridge University Press, 334-342.

[3] Browman, C.P. and Goldstein, L. 1989. Articulatory gestures as phonological units. *Phonology* 6(2): 201-231.

[4] Browman, C.P. & Goldstein, L. 1990. Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M.E. Beckman (eds) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. Cambridge: Cambridge University Press, 341–376.

[5] Browman, C.P. & Goldstein, L. 1992. Articulatory Phonology: an overview. *Phonetica* 49, 155–180.

[6] Carter, P. Forthcoming. Extrinsic phonetic interpretation: spectral variation in English liquids. To appear in J. Local, R. Ogden and R. Temple (eds.), *Papers in Laboratory Phonology VI.* Cambridge: Cambridge University Press.

[7] Docherty, G.J. 1992. *The Timing of Voicing in British English Obstruents.* The Hague: Mouton de Gruyter.

[8] English Intonation in the British Isles. Project funded by grant no. R000237145 from the Economic and Social Sciences Research Council. Web address: http://www.mml.cam.ac.uk/ling/intoproj.htm

[9] Grabe, E. 1998. *Comparative Intonational Phonology: English and German*. Doctoral Dissertation, University of Nijmegen & Max Planck Institute for Psycholinguistics. MPI Series in Psycholinguistics 7.

[10] Grabe, E. 1998. Pitch accent realisation in English and German. *Journal of Phonetics* 26, 129-143.

[11] Grabe, E., Nolan, F., and Farrar, K. Submitted. Pitch accent realisation in four varieties of British English. Submitted to *Journal of Phonetics.*

[12] Hawkins, S. 1992. An introduction to task dynamics. In G.J. Docherty and D.R. Ladd (eds), *Papers in Laboratory Phonology II.* Cambridge: Cambridge University Press, 9-25.

[13] Holst T. and Nolan F. 1995. The influence of syntactic structure on [s] to [S] assimilation. In B. Connell and A. Arvaniti (eds), *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV.* Cambridge: Cambridge University Press, 315-333.

[14] Horizon. 1999. BBC2 science documentary. Script archived at http://www.bbc.co.uk/horizon/infinity.shtml.

[15] International Phonetic Association. 1999. *The Handbook of the International Phonetic Association.* Cambridge: Cambridge University Press.

[16] Kelly, J. and Local, J.K. 1986. Long-domain resonance patterns in English. *Proceedings of the International Conference on Speech Input / Output, Institute of Electronic Engineers*, 304-308.

[17] Kelso, J.A.S., Saltzman, E.L. and Tuller, B. 1986. The dynamical perspective on speech production: data and theory. *Journal of Phonetics* 14, 29-59.

[18] Kerswill, P. 1987. Levels of linguistic variation in Durham. *Journal of Linguistics* 23, 25-49.

[19] Ladd, D.R. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.

[20] Ladefoged, P. 1972. Phonetic prerequisites for a distinctive feature theory. In A. Valdman (ed.), *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre.* The Hague: Mouton.

[21] Ladefoged, P. Forthcoming. Some thoughts on syllables: an old fashioned interlude. Commentary on papers by Carter, de Jong, and Dick. To appear in J. Local, R. Ogden and R. Temple (eds.), *Papers in Laboratory Phonology VI.* Cambridge: Cambridge University Press.

[22] Ladefoged, P. and Maddieson, I. 1996. *The Sounds of the World's Languages*. London: Blackwell.

[23] Lass, R. 1984. *Phonology*. Cambridge: Cambridge University Press.

[24] Liljencrants, J. and Lindblom, B. 1972. Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48, 831-852.

[25] Lindblom, B. 1986. Phonetic universals in vowel systems. In: J.J. Ohala and J.J. Jaeger (eds) *Experimental Phonology*, 13-44.

[26] Maddieson, I. 1984. *Patterns of Sounds.* Cambridge: Cambridge University Press.

[27] Ní Chasaide, A. and Gobl, C. 1993. Contextual variation of the vowel voice source as a function of adjacent consonants. *Language and Speech* 36, 303-330.

[28] Nolan, F. Forthcoming. Phonological representation and phonetic interpretation in intonation analysis. To appear in J. Local, R. Ogden and R. Temple (eds.), *Papers in Laboratory Phonology VI.* Cambridge: Cambridge University Press.

[29] Farrar, K. and Nolan, F. 1999. Timing of Fo peaks and peak lag. *Proc. 14th International Congress of Phonetic Sciences.* University of California, Berkeley.

[30] Nolan, F., Holst, T. and Kühnert B. 1996. Modelling [s] to [S] accommodation in English. *Journal of Phonetics* 24, 113-137.

[31] Pandeli, H., Eska, J.F., Ball, M.J., and Rahilly, J. 1997. Problems of phonetic transcription: the case of the Hiberno-English slit-t. *Journal of the International Phonetic Association* 27, 65-75.

[32] Science. 1998. The accelerating universe. *Science*, December 1998.

[33] Sproat, R. and Fujimura, O. 1993. Allophonic variation in English /l/. *Journal of Phonetics* 21, 291-311.