ABSTRACT

Evidence is discussed which perturbs the segmental, categorial foundation of descriptive phonetics. EPG studies showed that in cases which would be treated in auditory phonetic analysis and in phonological description as place assimilation, there is often a residual gesture towards the 'underlying' segment. Such results underline that the performance of segmental contrasts is neither discretely segmental in time, nor categorial in the sense of involving an inventory of discretely different elements. Segmentalised phonetic description is further challenged by instrumental evidence that neutralisation may be phonetically incomplete; and that segmental contrast may be cued over domains as large as the stress foot. Phonetics needs a more explicit statement of the relation of segments to articulation and to perception.

INTRODUCTION

Throughout the history of modern phonetics the phone-sized segment has been crucial. True, other elements such as syllables and features have periodically competed for attention; but the centrality of the phone is such that even these alternative elements are often talked of as combinations of, or decompositions of, phones.

The phone-sized segment remains at the heart of phonetic description despite decades of instrumental research into articulation and acoustics demonstrating beyond doubt that discrete phones do not exist in a straightforward way in the speech event, at least as superficially observable. An x-ray film of speech, or a dynamic palatographic record, shows gestures for different segments, overlapping and blending. And from the earliest speech synthesis it has been known, for instance, that the perceptual cues to a consonant are distributed at least over the adjacent vowels.

But the survival of the phone is not hard to explain. It is the basis of our only extensive model of general phonetic description, as embodied for instance in the alphabet of the International Phonetic Association. This in turn reflects the fact that phone- or phoneme-sized units provide the most generally applicable and revealing descriptions of the phonologies of languages.

Thus 'the phonetic sciences have proceeded in a somewhat schizophrenic state of mind, knowing that phones aren't really there, but at the same time they have to be there. The hope is generally that at some stage the relationship between segmentalised descriptions and the continuum of speech performance will become clear and well specified.

This paper draws together a number of cases where it seems that the tension between the discretely segmental description and the observable speech event is high enough to make the resolution of their relationship a priority.

DISCRETENESS ON TWO AXES

The traditional phone-based model of phonetic description implies discreteness on two axes.

Firstly, the phone symbols from left to right in a transcription imply a temporal sequence of discrete phonetic events. The strongest interpretation of this, with for instance all acoustic cues to a segment ending simultaneously and abruptly at a boundary with a following segment, is clearly falsified even by casual observation of spectrograms. Perhaps the weakest interpretation is one which allows overlapping in the realisation of phones, but still expects their implied sequencing to be respected in that the realisation of phone n will not extend later in time than that of phone n+1 nor earlier than that of phone n-1 (see Fig.1). For instance, if in
cases will work on coarticulation. The utterances [ski] and [sku] have the velar or is not cued until the vowel following is already in its place, and [d] is either present in an utterance or it is not, and if it is, it is wholeheartedly [d] and not something which vacillates between being a [d] and being a (b). Speech performance is thus implied to be categorical.

The existence of partially assimilated forms is supported in a similar experiment by Kerswill (1985). The gradual nature of assimilation in production is at variance with the paradigmatic discreetness of phone based representations. In principle, articulation could be categorial in that a speaker either made a gesture sufficient to create a given configuration of the vocal tract, or did not make it. Instead, articulation appears to be gradual – in that perception there are also signs of assimilation to the alveolar ridge (partial assimilation). The occurrence of greater morpho-phonetic underspecification, though not directly determined, by speech rate.

Neuralisation It appears that it is not only connected speech processes which put phone-segmental descriptors under strain. The question which seems current is whether the category change implied by this type of description reflects the facts of speech performance (articulation, or perception) since, as pointed out above, a phrase-segmental representation inherently implies discrete-ness sequentially and paradigmatically.

To find out if assimilation involves a discrete change in production we can compare form versus form, or form versus with forms which underlyingly contain the segment which is assimilated in the assimilation. For instance, when a speaker assimilates the place of articulation of the final consonant to the following velar in the road collapsed, is the utterance then phonetically identical in every respect to the realisation of the rogue collapsed?

This question has recently begun to be studied using electropalato-graphy (EPG). For instance, Barry (1985) found that where a word-final alveolar precedes a word-final velar, then the EPG display for a matched utterance with an underlying velar word-finally (complete assimilation) or, crucially, in many tokens it shows that no closure is completed on the alveolar ridge, but nevertheless the sides of the tongue make contact far forward along the sides of the palate in a ‘buttonhole’ gesture towards the alveolar ridge (partial assimilation). The occurrence of greater morpho-phonetic underspecification, though not directly determined, by speech rate. Therefore, the implied sequencing has been respected. If it is not, and if it is, it is whole—

It has been argued that the underlying word-final voicing contrast is in the preceding vowel, sequential discreetness of segments is not violated, but that neutralisation in these languages is as a transcription for the pair be?

Neuralisation It appears that it is not only connected speech processes which put phone-segmental descriptors under strain. The question which seems current is whether the category change implied by this type of description reflects the facts of speech performance (articulation, or perception) since, as pointed out above, a phrase-segmental representation inherently implies discrete-ness sequentially and paradigmatically.

To find out if assimilation involves a discrete change in production we can compare form versus form, or form versus with forms which underlyingly contain the segment which is assimilated in the assimilation. For instance, when a speaker assimilates the place of articulation of the final consonant to the following velar in the road collapsed, is the utterance then phonetically identical in every respect to the realisation of the rogue collapsed?

This question has recently begun to be studied using electropalato-graphy (EPG). For instance, Barry (1985) found that where a word-final alveolar precedes a word-final velar, then the EPG display for a matched utterance with an underlying velar word-finally (complete assimilation) or, crucially, in many tokens it shows that no closure is completed on the alveolar ridge, but nevertheless the sides of the tongue make contact far forward along the sides of the palate in a ‘buttonhole’ gesture towards the alveolar ridge (partial assimilation). The occurrence of greater morpho-phonetic underspecification, though not directly determined, by speech rate. Therefore, the implied sequencing has been respected. If it is not, and if it is, it is whole—

It has been argued that the underlying word-final voicing contrast is in the preceding vowel, sequential discreetness of segments is not violated, but that neutralisation in these languages is as a transcription for the pair be?

Neuralisation It appears that it is not only connected speech processes which put phone-segmental descriptors under strain. The question which seems current is whether the category change implied by this type of description reflects the facts of speech performance (articulation, or perception) since, as pointed out above, a phrase-segmental representation inherently implies discrete-ness sequentially and paradigmatically.

To find out if assimilation involves a discrete change in production we can compare form versus form, or form versus with forms which underlyingly contain the segment which is assimilated in the assimilation. For instance, when a speaker assimilates the place of articulation of the final consonant to the following velar in the road collapsed, is the utterance then phonetically identical in every respect to the realisation of the rogue collapsed?

This question has recently begun to be studied using electropalato-graphy (EPG). For instance, Barry (1985) found that where a word-final alveolar precedes a word-final velar, then the EPG display for a matched utterance with an underlying velar word-finally (complete assimilation) or, crucially, in many tokens it shows that no closure is completed on the alveolar ridge, but nevertheless the sides of the tongue make contact far forward along the sides of the palate in a ‘buttonhole’ gesture towards the alveolar ridge (partial assimilation). The occurrence of greater morpho-phonetic underspecification, though not directly determined, by speech rate. Therefore, the implied sequencing has been respected. If it is not, and if it is, it is whole—

It has been argued that the underlying word-final voicing contrast is in the preceding vowel, sequential discreetness of segments is not violated, but that neutralisation in these languages is as a transcription for the pair be?
perhaps, as summarised in Fig. 1, whether even proper sequencing is preserved in the speech signal.

Do speakers behave as if segments represent categorial choices? Apparently not; in environments with the potential for place of articulation assimilation a gradation of assimilation occurs.

Categories may be a function of hearing, rather than speaking. The continuum of behaviour from no place assimilation through partial to complete assimilation may turn out to yield a categorial perceptual boundary somewhere in the 'partial' region. But on the other hand it is possible that no perceptual boundary will emerge because, as with the cod card case, listeners can't exploit the acoustic details.

A consideration of the limits of segmental description, then, inevitably leads to consideration of the status of the categories which phone-segments imply, and of the representations which they comprise. If the disparity between production and perception which is hinted at by work cited here is confirmed, the general conception of phonetic analysis will have to be radically revised and its relation to aspects of speech performance made explicit.

REFERENCES


SCOTT, D.R. 1984 More on the /t/:/d/ distinction in American alveolar flaps. JASA 75 (Suppl. 1), S66.